

THE UNIVERSITY
of EDINBURGH

This thesis has been submitted in fulfilment of the requirements for a postgraduate degree (e.g. PhD, MPhil, DClinPsychol) at the University of Edinburgh. Please note the following terms and conditions of use:

This work is protected by copyright and other intellectual property rights, which are retained by the thesis author, unless otherwise stated.
A copy can be downloaded for personal non-commercial research or study, without prior permission or charge.
This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the author.
The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the author.
When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given.

The Scottish Epilepsy-related Deaths Study (SEDS)

Gashirai Kimeon Mbizvo



THE UNIVERSITY
of EDINBURGH

Doctor of Philosophy

The University of Edinburgh

2021

To Val,

Thank you for being my compass.

CONTENTS

ABSTRACT 8

LAY SUMMARY 10

DECLARATION 12

FUNDING 14

PERMISSIONS 14

ACKNOWLEDGEMENTS 15

OUTPUTS AND AWARDS DURING THIS PHD STUDENTSHIP 19

NATIONAL PRIZES 19
GRANTS 20
PUBLICATIONS FROM THIS THESIS 21
ABSTRACTS DERIVED FROM WORK IN THIS THESIS 21
OTHER PUBLICATIONS FORMING PART OF ACADEMIC AND PERSONAL DEVELOPMENT DURING THE PHD 22
PRESENTATIONS 24

PUBLIC ENGAGEMENT ACTIVITIES SURROUNDING THIS PHD PROJECT 26

ABBREVIATIONS 30

CHAPTER 1. INTRODUCTION 32

1.1 EPILEPSY AND MORTALITY 32
1.1.1 Epilepsy: an overview *32*
1.1.2 The evolving concept of epilepsy-related mortality *35*
1.1.2.1 Advancing the conversation beyond SUDEP: the need for a multiple-cause-of-death analysis 35
1.1.2.2 Identifying the proportion of potentially avoidable epilepsy-related deaths: the need for a novel method 39
1.1.2.3 Identifying those at high risk of epilepsy-related death: the need for a risk prediction model 40
1.1.3 Identifying epilepsy-related deaths in Scotland *42*
1.1.3.1 Rationale 42
1.1.3.2 The Scottish administrative healthcare data model 43
1.1.3.3 Validating the diagnostic accuracy of administrative healthcare data for epilepsy 46
1.1.3.4 Permissions required for administrative data research in Scotland 48
1.2 SUMMARY 49
1.3 AIMS OF THIS THESIS 50

CHAPTER 2. EPILEPSY-RELATED AND OTHER CAUSES OF MORTALITY IN PEOPLE WITH EPILEPSY: A SYSTEMATIC REVIEW OF SYSTEMATIC REVIEWS 53

2.1 INTRODUCTION 53
2.2 MANUSCRIPT 55
2.3 CONCLUSION 56

CHAPTER 3. THE ACCURACY OF USING ADMINISTRATIVE HEALTHCARE DATA TO IDENTIFY EPILEPSY CASES: A SYSTEMATIC REVIEW OF VALIDATION STUDIES 60

3.1 INTRODUCTION 60
3.2 PROTOCOL 62
3.3 SUMMARY 63

3.4 MANUSCRIPT 66
3.5 CONCLUSION 67

CHAPTER 4. VALIDATING THE ACCURACY OF ADMINISTRATIVE HEALTHCARE DATA IDENTIFYING EPILEPSY IN DECEASED ADULTS: A SCOTTISH DATA LINKAGE STUDY 69

4.1 INTRODUCTION 69
4.2 MANUSCRIPT 71
4.3 CONCLUSION 72

CHAPTER 5. A NATIONAL STUDY OF EPILEPSY-RELATED DEATHS IN SCOTLAND: TRENDS, MECHANISMS, AND AVOIDABLE DEATHS 73

5.1 INTRODUCTION 73
5.2 METHODS 76
5.2.1 Setting 76
5.2.2 Study design, data sources and participants 78
5.2.3 Enrichment linking 79
5.2.4 Variables and data sources 80
5.2.4.1 Identifying the mechanisms of epilepsy-related death using MCOD data 80
5.2.4.2 Identifying potentially avoidable epilepsy-related deaths 80
5.2.5 Bias and study size 81
5.2.6 Statistical analysis 81
5.2.6.1 Age-standardised mortality rate (MR) 81
5.2.6.2 Age-specific mortality ratios 82
5.2.6.3 MCOD data 82
5.2.6.4 Avoidable deaths 82
5.2.7 Approvals 83
5.3 RESULTS 83
5.3.1 Frequency and demographics of epilepsy-related deaths 83
5.3.2 Annual trend in epilepsy-related death MR 86
5.3.3 Trend in age-specific mortality ratios 87
5.3.4 Mechanisms involved with epilepsy-related death 89
5.3.5 Potentially avoidable epilepsy-related deaths in the young adults 96
5.4 DISCUSSION 98
5.4.1 Summary of findings 98
5.4.2 Research in context 98
5.4.3 Strengths and limitations 101
5.5 CONCLUSION 102

CHAPTER 6. A CASE-CONTROL STUDY DEVELOPING THE SCOTTISH EPILEPSY DEATHS STUDY SCORE (SEDS SCORE) AS A RISK PREDICTION MODEL FOR EPILEPSY-RELATED DEATHS 104

6.1 INTRODUCTION 104
6.2 METHODS 105
6.2.1 Study design summary 105
6.2.2 Data sources, and participants 106
6.2.2.1 Case ascertainment 106
6.2.2.2 Control ascertainment 106
6.2.2.3 Confirming epilepsy diagnosis 107
6.2.2.4 Data abstraction, outcome and predictors 107
6.2.3 Statistical analysis 108
6.2.3.1 Sample size 108
6.2.3.2 Missing data 109
6.2.3.3 Model development 109
6.2.4 Approvals 111
6.3 RESULTS 111
6.3.1 Baseline characteristics of cases and controls 111
6.3.2 Missing data 112
6.3.3 Correlation 116
6.3.4 Univariable analysis 117
6.3.5 Multivariable analysis 120
6.4 DISCUSSION 123
6.4.1 Risk factors 123
6.4.2 Risk prediction model 125
6.4.3 Strengths and limitations 127
6.5 CONCLUSION 129

CHAPTER 7. DISCUSSION 130

7.1 OVERVIEW 130
7.2 ESTIMATING THE GLOBAL BURDEN OF ALL-CAUSE AS WELL AS EPILEPSY-RELATED MORTALITY IN PWE 130
7.2.1 Study observations 130
7.2.2 Areas for future research 131
7.3 ESTIMATING THE OVERALL ACCURACY OF USING ADMINISTRATIVE HEALTHCARE DATA TO IDENTIFY EPILEPSY CASES 132
7.3.1 Study observations 132
7.3.2 Areas for future research 132

7.4 Identifying the optimal algorithms for identifying epilepsy cases in deceased adults in Scotland using administrative data 134
7.4.1 Study observations 134
7.4.2 Areas for future research 135
7.5 Using administrative data from adults to investigate the trends and mechanisms of epilepsy-related deaths in Scotland, highlighting the proportion that were potentially avoidable 136
7.5.1 Study observations 136
7.5.2 Areas for future research 137
7.6 Identifying multiple risk factors for the occurrence of epilepsy-related death in Scottish adults and developing a pragmatic risk prediction model for identifying which individuals are at the highest risk 138
7.6.1 Study observations 138
7.6.2 Areas for future research 140
7.7 Final conclusions 141
REFERENCES 143
APPENDIX 1 – SUPPLEMENTARY MATERIAL TO CHAPTER 2 170
Appendix A: Search strategy 170
Appendix B: data abstraction pro forma 171
Appendix C: List of excluded records 172
Appendix D: List of the primary studies contributing evidence 174
Appendix E: AMSTAR quality scoring of included systematic reviews 184
APPENDIX 2 – SUPPLEMENTARY MATERIAL TO CHAPTER 3 186
Appendix S1: Search strategies for MEDLINE and Embase 186
Appendix S2: Data abstraction tool 188
Appendix S3: Adapted QUADAS-2 assessment form 190
Appendix S4: Kappa statistic interrater agreement: study selection 193
Appendix S5: Reference list of excluded studies 194
Appendix S6: QUADAS-2 quality assessment items 207
Appendix S7: Study algorithm lookup table 211
Appendix S8: Summary of results from countries with a single study 221
Appendix S9: Meta-analysis 224
APPENDIX 3 – SUPPLEMENTARY MATERIAL TO CHAPTER 4 233
Appendix A - Scottish mortality data 233
Appendix B – Antiepileptic drug filters 234
Appendix C – Primary care data 236
APPENDIX 4 – SUPPLEMENTARY MATERIAL TO CHAPTER 5 238
Appendix S1 – Primary care recruitment 238
Appendix S2 – Epilepsy Read codes used 238
Appendix S3 – Comorbidity Read codes used 239
Appendix S4 – Regression analyses 266
APPENDIX 5 – SUPPLEMENTARY MATERIAL TO CHAPTER 6 277
Appendix S1 - Scottish Epilepsy-related Deaths Study (SEDS) data collection pro formas 278

APPENDIX S2 – DISCRIMINATION ESTIMATES 285

Abstract

Introduction: Epilepsy is a global condition associated with premature death. Although people with epilepsy (PWE) can die of causes unrelated to their epilepsy, many are thought to suffer increased risk of death because of the epilepsy itself, its treatment, or comorbidities related to epilepsy. Together, these are termed 'epilepsy-related deaths'. These remain poorly understood despite many being potentially avoidable. This thesis aims to perform an in-depth assessment of epilepsy-related deaths.

Methods: We first undertake a systematic review describing the comparative risks, causes, and risk factors for all-cause mortality in PWE, highlighting the burden that were likely to have been epilepsy-related deaths. We then identify how accurately administrative healthcare data capture epilepsy globally by systematically reviewing studies assessing this, before undertaking our own diagnostic accuracy study of multiple linked administrative epilepsy datasets in adults (aged 16 years and over) deceased between 2009 and 2016 across Scotland. The optimal case-ascertainment strategy is then used to describe Scotland-wide trends and mechanisms of epilepsy-related deaths during this study period, highlighting the proportion that were potentially avoidable. Finally, we undertake an age- and sex-matched case-control study comparing a sample of adults who suffered epilepsy-related death, identified in the linked datasets, to living adults with epilepsy in Scotland, identified from a research database and epilepsy clinics.

Results: The administrative data accurately capture epilepsy and show that all-cause mortality is increased in PWE, has failed to reduce over time, and

that young adults aged 16 to 54 years are at the highest risk when compared to age-mated mortality figures in the general population. Many of these deaths are likely to have been epilepsy-related. Although sudden unexpected death in epilepsy (SUDEP) is a common epilepsy-related cause of death in young adults, the non-SUDEP epilepsy-related causes may be more common. These include aspiration pneumonia, drowning, and antiepileptic drug (AED) poisoning. Many epilepsy-related deaths are potentially avoidable. The risk factors for suffering epilepsy-related death include living in the most deprived areas of Scotland, recent hospital attendances for seizures, and inherited or congenital aetiologies for developing epilepsy. A novel risk prediction model, the Scottish Epilepsy-related Deaths Score (SEDS Score), is indicative that any combination of these risk factors may be associated with very high risks of suffering epilepsy-related death, with a combination of all three reaching a nearly 20-times increased risk.

Conclusion: Epilepsy-related deaths are potentially a major public health problem, failing to reduce over time, affecting young adults more than would be expected, and with many being potentially avoidable. The SEDS Score may be a useful clinical tool for capturing those at the highest risk of epilepsy-related death, and it will require external validation. It is reasonable to investigate epilepsy-related mortality using linked administrative healthcare data.

Lay summary

Epilepsy is common and, sadly, it can result in early loss of life. In this study, we try to understand deaths caused by epilepsy in greater detail. In order to do this, we first review studies that have already investigated deaths in people with epilepsy. These indicate that the problem of early death in people with epilepsy happens in many countries around the world. Moreover, these deaths do not seem to have reduced over the years.

To help understand the problem more, we need to investigate deaths caused by epilepsy in Scotland, as this has not been done before in a national study. We do this by reviewing anonymous data routinely collected around the country to help healthcare professionals study disease. The data that we study are held securely and only approved healthcare professionals responsible for this study can access them. We show that these data identify people with epilepsy accurately. We go on to show that during our study period of 2009 to 2016, there were many deaths caused by epilepsy in Scottish adults aged over 16 years. Young adults, aged between 16 and 54 years, were at the highest risk of such deaths, when their deaths were compared to the number we would expect in people of the same age in the general population. Many causes of death were identified, some of which were surprising as they were more common than we would normally expect. Our findings suggest that many of the deaths caused by epilepsy may well be avoidable, although more research is needed in this area. However, this result is encouraging, particularly when contextualised with the fact a tremendous amount of work is being done in Scotland to see how health data can save lives.

Our data show that being from poorer areas of Scotland is associated with an increased risk of dying because of epilepsy. We also show that having the sort of poorly controlled epilepsy that requires frequent hospital attention also increases the risk of these deaths. Another important factor that could increase the risk of deaths caused by epilepsy would be having an underlying inherited condition causing the epilepsy. These factors can be combined to help doctors figure out who is at the highest risk of deaths caused by epilepsy. This is something we have tried to do in the current study, and our work may therefore help other doctors looking after epilepsy patients. It will be helpful to have other researchers also try to do what we have done in other countries to see if they can show the same sorts of results.

Declaration

This thesis has been composed by myself, Gashirai Mbizvo, and the work has not been submitted for any other degree or professional qualification. I declare that the work is my own unless otherwise stated, where I have indicated that the work has formed part of a jointly authored research publication originating from this work. In all such cases, I can confirm I have made a substantial contribution to the work as first author and indicate the contribution of all authors involved in the statement below. For all published manuscripts included in this thesis, I can confirm that the current University of Edinburgh Guidelines on Including Publications in Postgraduate Research Theses have been followed (www.ed.ac.uk/files/atoms/files/publications_in_thesis.pdf), including obtaining the publisher's formal permission to include each manuscript presented in the thesis.

The work presented in Chapter Two was published in *Epilepsy Research* as Mbizvo GK, Bennett K, Simpson CR, Duncan SE, Chin RFM. Epilepsy-related and other causes of mortality in people with epilepsy: A systematic review of systematic reviews. *Epilepsy Res* 2019; **157**: 106192. Author contributions (student in bold): Study conception (**GKM**), study design (**GKM**, RFMC), data acquisition (**GKM**, KB), data analysis (**GKM**, KB), risk of bias assessments (**GKM**, KB), manuscript drafting (**GKM**), manuscript revisions (all authors), final approval of submitted manuscript (all authors).

The work presented in Chapter Three was published in *BMJ Open* (protocol) and *Epilepsia* (manuscript), as Mbizvo GK, Bennett K, Simpson CR, Duncan SE, Chin RFM. Accuracy and utility of using administrative healthcare

databases to identify people with epilepsy: a protocol for a systematic review and meta-analysis. *BMJ Open* 2018; **8**(6): e020824. Author contributions (student in bold): Study conception (**GKM**), study design (**GKM**, RFMC), data acquisition (**GKM**, KB), data analysis (**GKM**), risk of bias assessments (**GKM**, KB), manuscript drafting (**GKM**), manuscript revisions (all authors), final approval of submitted manuscript (all authors). Mbizvo GK, Bennett KH, Schnier C, Simpson CR, Duncan SE, Chin RFM. The accuracy of using administrative healthcare data to identify epilepsy cases: A systematic review of validation studies. *Epilepsia* 2020;61(7):1319-35. Author contributions (student in bold): Study conception (**GKM**), study design (**GKM**, RFMC), data acquisition (**GKM**, KB), data analysis (**GKM**, CS), risk of bias assessments (**GKM**, KB), manuscript drafting (**GKM**), manuscript revisions (all authors), final approval of submitted manuscript (all authors).

The work presented in Chapter Four was published in *Epilepsy Research* as Mbizvo GK, Schnier C, Simpson CR, Duncan SE, Chin RFM. Validating the accuracy of administrative healthcare data identifying epilepsy in deceased adults: A Scottish data linkage study. Epilepsy Res 2020; 167: 106462. Author contributions (student in bold): Study conception (**GKM, RFMC**), study design (**GKM**, RFMC), data acquisition (**SD**), data analysis (**GKM**, CS), manuscript drafting (**GKM**), manuscript revisions (all authors), final approval of submitted manuscript (all authors).

Funding

This PhD was kindly supported charitably by Epilepsy Research UK (R44007) and the Juliet Bergqvist Memorial Fund. The funders played no role in the design or conduct of this work.

Permissions

Chapters One, Two, Three, and Seven did not require prior approval. The studies in Chapters Four, Five and Six were approved by South East Scotland Research Ethics Committee 2 (15/SS/0165, IRAS 181131), and the Scottish Public Benefit and Privacy Panel for Health and Social Care (PBPP).

Acknowledgements

I am eternally grateful to have a world-class team of supervisors, led academically by Professor Richard Chin and clinically by Dr Susan Duncan. I have learned, over the years, that completing a PhD is as much about personal development as it is about scientific progress, and I feel I have been supported well in both of these areas by my lead supervisors. I owe an especially great deal of thanks to them for their patience in giving me the space to complete this PhD over an extended period that has been interrupted by my clinical commitments to the COVID-19 crisis. I also owe them thanks for pushing me to try and publish my work as I go, and for also providing all of the necessary support for me to do so. Professor Colin Simpson, my co-supervisor, has also been hugely helpful and supportive in providing outside guidance on managing various strategic aspects of the PhD, and it has been invaluable to be able to draw from his extensive experience in epidemiology.

I am incredibly grateful to the funders of this PhD, without whose support the work would not have been possible. This commenced with funding from Epilepsy Research UK (ERUK), followed a few years later by funding from the Juliet Bergqvist Memorial Fund. ERUK have remained hugely supportive throughout my PhD, and they have provided a helpful platform for public engagement for this work. I am grateful particularly to Maxine Smeaton and Caoimhe Bennett. The unwavering support I have received from the Bergqvist family has also been invaluable, and I am hugely grateful to them for this. I hope that they are proud of what I have achieved with their kind support. I

would also like to acknowledge support received from Epilepsy Action in producing our systematic review of epilepsy-related mortality.

I would like to make a special mention about the support I have received from Professor Siddharthan Chandran as an academic mentor. He has helped me to believe in myself and to be confident in what I can achieve in a way that has made me aspire to aim high and try to follow as impactful an academic career path as I can. I remain hugely grateful for his ongoing mentorship. I would like to further add thanks to Professor Tony Marson, who quite unexpectedly forwarded me the job advert for this PhD some years ago and encouraged me to apply, highlighting all of the benefits of an academic career, my suitability for it, and speaking very highly of the supervisory team and centre. Without this kind push, I would have never even thought to apply for a PhD or move to Scotland.

Dr Christian Schnier is an invaluable member of the research team and I am hugely grateful for his technical support with the statistical packages. I consider him no less than a genius. I have presented complex statistical problems to him that are not in his area on a number of occasions and he has managed to figure them out and show me how to do them. This PhD thesis would definitely be less interesting without Christian, and I am certain our Chapter Three revision to *Epilepsia* would have been rejected had it not been for his technical support with the meta-analysis. I am hugely grateful to Professor Cathie Sudlow for having the foresight to kindly introduce me to Christian very early on in my studies. I am also thankful to the Centre for Medical Informatics team for their regular research networking opportunities, particularly those that lead

me to the HealthyR Course, which helped provide me with the skills to transform some of the analyses in my chapters. I am also grateful to Professor Christopher Weir and Dr Jacqueline Stephen for their statistical support. Furthermore, I would like to thank the eDRIS, SHARE, Albasoft, and NRS teams for their assistance with data control and acquisition, particularly Sian Nowell, Shobna Vasishta, Louise Dow, Dave Kelly, Colin Fischbacher, and Julie Ramsay. I am also grateful to Kenneth Scott for his support with obtaining ethics and R&D approvals.

I am thankful to the team of consultant neurology colleagues around Scotland who have kindly assisted with access to patient medical records for this project: Kate Taylor (NHS Highland), Katy Murray (NHS Forth Valley), Linda Gerrie and Graham Mackay (NHS Grampian), Kathleen White and Ian Morrison (NHS Tayside), Saif Razvi (NHS Ayrshire and Arran), Myles Connor (NHS Borders), Ondrej Dolezal (NHS Dumfries and Galloway), Russell Hewett (NHS Greater Glasgow and Clyde), Martin Zeidler (NHS Fife), and John Paul Leach (NHS Lanarkshire).

I am also grateful to the Muir Maxwell Epilepsy Centre (MMEC) team for support with everyday research life, particularly Jane Andrews for helping with several administrative tasks during the PhD. I also thank Karen Chalmers for her administrative support. I would also like to thank the Centre for Clinical Brain Sciences (CCBS) team.

I would like to acknowledge and thank the team of senior academics who helped me prepare for funding interviews early on during my PhD in 2016: Professors Cathie Sudlow, Aziz Sheikh, Richard Knight, Charles ffrench-

Constant, William Whiteley, and David Hunt. I am also grateful to Tim Wilkinson and Danielle Leighton for the peer-to-peer research support they have provided me.

Finally, I would like to thank my mum, Dr Elizabeth Mbizvo, and my dad, Professor Michael Mbizvo, for always believing in me and supporting me with everything that I do. They are my most cherished role models, both academically and personally, and I am never unware of how lucky I am to have parents like them, and of how incredibly hard they have worked to get me to where I am today. I am also forever grateful to my wonderful partner Val for essentially propping me up in love throughout this PhD, particularly during some of the most difficult times.

It is clear to me that my PhD work has been a major collaboration between a fantastic network of people, who I am eternally grateful to, particularly my co-authors. Therefore, in the spirit of this gratitude, the third person “we” will be used throughout the scientific work reported in this thesis, instead of “I”.

Outputs and awards during this PhD studentship

National Prizes

1. Runner up: Clinical Neurology Prize at the European Academy of Neurology Tournament for Neurologists in Training 2020. *Project title:* Secular trends in adult epilepsy-related and potentially avoidable mortality in Scotland: a nationwide population-based study

2. Runner up: Case of the Quarter Prize 2019 - Royal College of Physicians of Edinburgh. *Project title:* Ictal asystole: A diagnostic and management conundrum.

3. Winner: Liversedge Prize 2019 - Best platform presentation at the North of England Neurological Association (NENA) meeting. *Project title:* The Scottish Epilepsy Deaths Study (SEDS): identifying avoidable epilepsy-related deaths

4. Winner: Charles Symonds Prize 2019: Best platform presentation, at the Association of British Neurologists (ABN) Annual Meeting, held in Edinburgh. *Project title:* The Scottish Epilepsy Deaths Study (SEDS): identifying avoidable epilepsy-related deaths

5. Winner: The John D Matthews Postgraduate Award 2019. This Scholarship is given to the PhD student who has demonstrated the greatest academic excellence in neurology during the previous year. *Project title:* The Scottish Epilepsy Deaths Study (SEDS): identifying avoidable epilepsy-related deaths

6. Winner: Best Oral Platform Presentation Prize at the International League Against Epilepsy (ILAE) British Chapter conference 2018. *Project Title:* Secular trends in adult epilepsy-related deaths: a nationwide population-based study using validated routine administrative data

7. Winner: Best Overall Presentation Prize at the Scottish Association of Neurological Sciences (SANS) Annual Meeting. *Project Title:* The Scottish Epilepsy Deaths Study (SEDS)

Grants

1. Brain Travel Award recipient 2019: £1,000 to attend ILAE/IBE 33rd International Epilepsy Congress, Bangkok, Thailand. *Project title:* Epilepsy-Related Deaths Are A Major Public Health Problem: Results from A National Study Using Linked Administrative Health Data to Determine the Burden of Epilepsy-Related Deaths and The Proportion That Are Potentially Avoidable

2. ABN Conference Bursary recipient 2019: £100 to attend the Annual Meeting in Edinburgh. *Project title:* The Scottish Epilepsy Deaths Study (SEDS): identifying avoidable epilepsy-related deaths

3. ILAE Conference Bursary 2018: €965 to attend the 13th European Congress on Epileptology in Vienna. *Project title:* Secular trends in sudden unexpected death in epilepsy (SUDEP) and other epilepsy-related deaths in adults: a nationwide population-based study

Publications from this thesis

1. **Mbizvo GK**, Bennett K, Simpson CR, Duncan SE, Chin RFM. Epilepsy-related and other causes of mortality in people with epilepsy: A systematic review of systematic reviews. *Epilepsy Res* 2019; **157**: 106192
2. **Mbizvo GK**, Bennett K, Simpson CR, Duncan SE, Chin RFM. Accuracy and utility of using administrative healthcare databases to identify people with epilepsy: a protocol for a systematic review and meta-analysis. *BMJ Open* 2018; **8**(6): e020824.
3. **Mbizvo GK**, Bennett KH, Schnier C, Simpson CR, Duncan SE, Chin RFM. The accuracy of using administrative healthcare data to identify epilepsy cases: A systematic review of validation studies. *Epilepsia* 2020; **61**(7): 1319-35.
4. **Mbizvo GK**, Schnier C, Simpson CR, Duncan SE, Chin RFM. Validating the accuracy of administrative healthcare data identifying epilepsy in deceased adults: A Scottish data linkage study. *Epilepsy Res* 2020; **167**: 106462.
5. **Mbizvo GK**, Schnier C, Simpson CR, Chin RFM, Duncan SE. A national study of epilepsy-related deaths in Scotland: Trends, mechanisms, and avoidable deaths. *Epilepsia* 2021; **62**(11): 2667-84.

Abstracts derived from work in this thesis

1. **Mbizvo G**., Schnier C., Simpson C.R., Duncan S.E., Chin R.F. Secular trends in adult epilepsy-related and potentially avoidable mortality in Scotland: A nationwide population-based study. Eur. J. Neurol.. 2020;**27**(Supplement 1):61.

2. **Mbizvo G.K**., Schnier C., Simpson C.R., Duncan S.E., Chin R.F. Epilepsy-related deaths are a major public health problem: Results from a national study using linked administrative health data to determine the burden of epilepsy-related deaths and the proportion that are potentially avoidable. *Epilepsia*. 2019;**60**(Supplement 2):73-74.
3. **Mbizvo GK**, Schnier C, Simpson CR, et al 11.42 The Scottish epilepsy deaths study (SEDS): identifying avoidable epilepsy-related deaths. *Journal of Neurology, Neurosurgery & Psychiatry* 2019;90:e9-e10.
4. **Mbizvo G.K**., Simpson C.R., Duncan S.E., Chin R.F. Secular trends in sudden unexpected death in epilepsy (SUDEP) and other epilepsy-related deaths in adults: A nationwide population-based study. *Epilepsia*. 2018;**59(**Supplement 3):S26.

Other publications forming part of academic and personal development during the PhD

1. **Mbizvo GK**, Larner AJ. Receiver operating characteristic plot and area under the curve with binary classifiers: pragmatic analysis of cognitive screening instruments. *Neurodegener Dis Manag*. 2021;11(5):353-60.
2. **Mbizvo GK**, Ziso B, Larner AJ. Epilepsy and prion diseases: A narrative review. *Epilepsy Behav* 2021; **115**: 107630.
3. Yong K., Eaton C., Walter V., **Mbizvo G.K**., Rhodes S., Chin R.F.M. Stimulant and non-stimulant drug therapy for people with epilepsy (PWE) and attention deficit hyperactivity disorder (ADHD): A Cochrane review. *Dev. Med. Child Neurol.*. 2021;63(SUPPL 1):88.

4. Schnier C, Duncan S, Wilkinson T, **Mbizvo GK**, Chin RFM. A nationwide, retrospective, data-linkage, cohort study of epilepsy and incident dementia. *Neurology* 2020; **95**(12): e1686-e93.
5. **Mbizvo GK**, Chandrasekar B, Nevitt SJ, Dixon P, Hutton JL, Marson AG. Levetiracetam add-on for drug-resistant focal epilepsy. *Cochrane Database of Systematic Reviews* 2020, Issue **6**. Art. No.: CD001901.
6. Schnier C., Duncan S., Wilkinson T., **Mbizvo G.K**., Chin R.F.M. A nationwide, retrospective, data-linkage, cohort study of epilepsy and incident dementia. *Neurology*. 2020;**95**(12):e1686-e1693.
7. **Mbizvo GK**, Derry C, Davenport R. Ictal asystole - a letter within a letter. *J Roy Coll Phys Edin* 2020; **50**(2): 207-8.
8. **Mbizvo, G. K**., Lentell, I. C., Leen, C., Roddie, H., Derry, C. P., Duncan, S. E. & Rannikmae, K. 2019. Epilepsia partialis continua complicated by disseminated tuberculosis and hemophagocytic lymphohistiocytosis: a case report. *J Med Case Rep,* 13**,** 191.
9. **Mbizvo G.K**., Derry C., Davenport R. Ictal asystole: A diagnostic and management conundrum. *J. R. Coll. Phys. Edinburgh*. 2019;**49**(2):128-131
10. **Mbizvo, G. K**., Lentell, I. C., Leen, C., Roddie, H., Derry, C. P., Duncan, S. E. & Rannikmae, K. 2019. 083 Facial twitching, breathlessness and fever: an unexpected and potentially lethal triad. *Journal of Neurology, Neurosurgery & Psychiatry* 2019;**90**:A27-A28.
11. Walter V, **Mbizvo GK**, Chin RFM. Stimulant and non-stimulant drug therapy for people with attention deficit hyperactivity disorder and epilepsy.

Cochrane Database of Systematic Reviews 2018, Issue 9. Art. No.: CD013136.

12. **Mbizvo GK,** Derry C. Diagnosis and management of sleep-related epilepsy in adults. *Practitioner*. Sept 2018;262(1818):11-16
13. **Gashirai M**, Chris D, Richard D. THUR 051 Ictal asystole: return to the history. *Journal of Neurology, Neurosurgery & Psychiatry* 2018;89:A6.
14. Ellis R.J.B., **Mbizvo G.K**., Jacob A., Doran M., Larner A.J. Relapsing polychondritis complicated by cognitive dysfunction: two distinct clinical phenotypes? *Int. J. Neurosci*.. 2017;**127**(2):124-134.
15. **Mbizvo GK**, Leighton DJ. When breath becomes air. *Pract Neurol*. 2017;**17**(4):335.

Presentations

Platform presentations

1. *Presenter:* Mbizvo GK. Secular trends in adult epilepsy-related and potentially avoidable mortality in Scotland: a nationwide population-based study. *Conference:* Virtual Congress of the European Academy of Neurology 2020
2. *Presenter:* Mbizvo GK. The Scottish Epilepsy Deaths Study (SEDS): identifying avoidable epilepsy-related deaths. *Conference:* North of England Neurological Association (NENA) 2019, Harrogate UK.
3. *Presenter:* Mbizvo GK. Secular trends in adult epilepsy-related deaths: a nationwide population-based study using validated routine administrative data *Conference:* ILAE British Chapter Annual Scientific Meeting 2018.

4. *Presenter:* Mbizvo GK. Secular trends in sudden unexpected death in epilepsy (SUDEP) and other epilepsy-related deaths in adults: a nationwide population-based study. *Conference:* 13th European Congress on Epileptology (ECE) 2018 in Vienna, Austria.
5. *Presenter:* Mbizvo GK. The Scottish Epilepsy Deaths Study. *Conference:* Scottish Epilepsy Group Annual Scottish Research Day on 2018 in Edinburgh UK
6. *Presenter:* Mbizvo GK. The Scottish Epilepsy Death Study. *Conference:* South East Scotland Regional Epilepsy Day 2018 in Edinburgh
7. *Presenter:* Mbizvo GK. The Scottish Epilepsy Deaths Study. *Conference:* Scottish Association of Neurological Sciences (SANS) Annual Meeting 2017

Poster presentations

1. *Presenter:* Mbizvo GK. Epilepsy-Related Deaths Are A Major Public Health Problem: Results from A National Study Using Linked Administrative Health Data to Determine the Burden of Epilepsy-Related Deaths and The Proportion That Are Potentially Avoidable. *Conference:* 33rd International Epilepsy Congress 2019 in Bangkok, Thailand
2. *Presenter:* Mbizvo GK. Secular trends in adult epilepsy-related deaths: a nationwide population-based study using validated routine administrative data *Conference:* Clinical Academics in Training Annual Conference 2018, Edinburgh, UK

Public engagement activities surrounding this PhD project

There have been three key areas of public engagement surrounding this PhD project.

First, I presented the study findings at Scottish Parliament, Holyrood on 28th March 2019. This was for a Cross Party Group Meeting convened to discuss the findings of this study. Those in attendance included members of the public with epilepsy, family members who have lost loved ones to epilepsy, charitable organisations representing epilepsy, members of the press, and Members of the Scottish Parliament (MSPs). This provided an opportunity for active discussion about the project with the public, allowing us to educate them about epilepsy-related deaths and also allowing us to draw from the experiences of the public to inform lessons learned from and conduct of the study. For example, one MSP highlighted their own experience of living with epilepsy and said that they could not remember the last time they had spoken to a medical professional about their epilepsy or medications. They observed that there is no proactive system to remind people to come in for review. This information would be supportive of some of our study findings in Chapter 5 and help strengthen an argument for the need for greater resources to support outpatient epilepsy care. Another member of the public asked about the feasibility of mental health screening in adults with epilepsy, drawing on lessons learned from such a pilot in paediatric epilepsy care in Scotland. This was a helpful suggestion with practical implications for areas of study for the project.

The second area of public engagement was through online multimedia, where I took on the role of social media and website representative for the Muir Maxwell Epilepsy Centre (MMEC) through 2017–19. This is an epilepsy research charity at The University of Edinburgh. I was, therefore, able to use these online multimedia platforms to send relevant updates about the project directly to members of the public, including highlighting ley summaries of published work.

The third area of public engagement has been through one of the project funders, ERUK. This is the largest epilepsy research charity in the UK. I presented the project to lay ERUK members and their families on 4th April 2019, and also wrote a publicly-available blog for the ERUK website about the work (*available at* https://epilepsyresearch.org.uk/our-research/research-blog/epilepsy-deaths-research-changing-policy).

Links to press coverage of the PhD project

This PhD project work was recently recommended in Nature Reviews Neurology, in a news article by the Senior Editor (available at www.nature.com/articles/s41582-020-0387-6). Links to other press coverage of the work are provided below.

- www.medscape.com/viewarticle/931365
- www.practicalneurology.com/news/epilepsy-related-deaths-have-not-reduced-over-time
- www.eurekalert.org/pub_releases/2020-05/sh-edc052020.php
- www.outlookindia.com/newsscroll/young-adults-more-likely-to-die-from-epilepsy-study/1845341

- www.indiatvnews.com/lifestyle/news-young-adults-more-likely-to-die-from-epilepsy-study-620376
- www.thehindubusinessline.com/news/science/young-adults-more-prone-to-die-of-epilepsy-study/article31668354.ece#
- www.pressreader.com/ireland/irish-daily-mirror/20200525/281934545152688
- www.socialnews.xyz/2020/05/25/young-adults-more-likely-to-die-from-epilepsy-study/
- www.medicalxpress.com/news/2020-05-epilepsy-related-deaths-common-young-adults.html
- www.santelog.com/actualites/epilepsie-les-deces-augmentent-pourtant-beaucoup-pourraient-etre-evites
- www.mdedge.com/neurology/article/223122/epilepsy-seizures/most-adult-epilepsy-related-deaths-could-be-avoided
- www.epilepsy.ie/content/study-suggests-epilepsy-deaths-are-not-decreasing-despite-advances-treatments
- https://bit.ly/3ejMsMM
- www.neurologyreviews-digital.com/neurologyreviews/july2020/MobilePagedArticle.action?articleId=1601258
- www.neurodiem.it/news/epilepsy-related-deaths-remain-common-especially-in-young-adults-say-5wiZQlZxeoIgg72ZQva5ez

- www.univadis.com/viewarticle/epilepsy-excess-mortality-often-disease-related-potentially-preventable-4228a7ac-95c2-36e8-ad50-83d0336e8563

Abbreviations

A&E - Accident and emergency

AEDs - Antiepileptic drugs

CHI - Community Health Index

CI - Confidence interval

DWP - Department for Work and Pensions

F25 - Epilepsy read codes

G40 - ICD-10 code for epilepsy

G41 - ICD-10 code for status epilepticus

GDPR - General Data Protection Regulation

GTCS - Generalised tonic-clonic seizure

HMRC - Her Majesty's Revenue and Customs

HSC-PBPP - Public Benefit and Privacy Panel for Health and Social Care

ICD - International Classification of Diseases system

MCOD - Multiple cause of death

MR - Mortality rate

NHS – National Health Service

NPV - Negative predictive value

NRS - National Records of Scotland

ONS - Office for National Statistics

OR - Odds ratio

PIS -Prescribing Information System

PPV - Positive predictive value

PWE – People with epilepsy

R56.8 - ICD-10 code for seizures

RDAM - Revised Definition of Avoidable Mortality

SEDS Score - Scottish Epilepsy-related Deaths Study Score

SHARE - Scottish Health Research Register

SIMD - Scottish Index of Multiple Deprivation

SMR - Standardised mortality ratio

SMR01/00 - Scottish Morbidity Record 01/00

SUDEP - Sudden unexpected death in epilepsy

TBI - Traumatic brain injury

TRIPOD - Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis

YPLL - Years of potential life lost

Chapter 1. Introduction

Student contributions: conception, online database searches, literature review, drafting and revision.

1.1 Epilepsy and mortality

1.1.1 Epilepsy: an overview

Epilepsy is a chronic noncommunicable disorder that affects over 70 million people worldwide.[1,2] Three percent of the global burden of disease, measured in disability-adjusted life years, is resultant from neurological disorders; of which a quarter are due to epilepsy.[3] This makes epilepsy one of the most common conditions to affect the human brain.[1] In high-income countries, the incidence of epilepsy is consistent across different regions, around 50/100,000 people per year (range 40–70/100,000).[1] The incidence is even higher in low-income countries, usually above 80–100/100,000 people per year.[1] In Europe, similar to in most global regions, age distribution of the incidence of epilepsy is bimodal, with two peaks of frequency: in childhood (<16 years of age) and in older adults (≥ 55 years of age) (Figure 1-1).[4] Regardless of the geographical location, the prevalence of active epilepsy is usually between four and 12 per 1,000.[1]



Figure 1-1: Age-specific incidence of epilepsy in Europe

Figure reproduced with permission from ILAE/IBE/WHO Global Campaign Against Epilepsy 2010[4]

Epilepsy is defined as: two unprovoked seizures occurring more than 24 hours apart; a single unprovoked seizure if recurrence risk is high (i.e., >60% over the next 10 years); or a diagnosis of an epilepsy syndrome.[5] This lasting predisposition to generate spontaneous epileptic seizures has several physical, psychological and social consequences.[1] From a physical perspective, accidents and injuries are more common in people with epilepsy (PWE) than in the general population, and this is more so in those with increasing seizure frequency.[6] From a psychological perspective, poor mental health is often associated with epilepsy, with the incidence of depression in epilepsy reported to be as high as 55%.[7] This relationship can be bidirectional, with depression leading to seizures through sleep deprivation, for example, and seizures leading to depression through both biological mechanisms and the social impact of epilepsy.[7] A recent case-control study of employability in PWE demonstrated that having epilepsy was more strongly associated with higher unemployment rates ($p < 0.0001$); job layoffs ($p = 0.001$); being unfit to

work ($p < 0.0001$); feeling shame for having the disease ($p = 0.014$); absence of partners ($p = 0.026$); and depression ($p = 0.004$).[8] It is not uncommon for PWE and their families to suffer from stigma and discrimination.[2] Additional socioeconomic issues for PWE relate to the impact of seizures on driving. In the United Kingdom (UK), similar to in many other countries, there is a legal requirement for a one-year period of seizure freedom for drivers with epilepsy who use a Group 1 car and motorcycle license.[9]

The diagnosis of epilepsy is primarily based on a patient's clinical history and eyewitness accounts of events that have occurred. Investigations such as electroencephalography and brain imaging can help support the diagnosis by demonstrating electrical or structural brain abnormalities that may increase the chances of events being ictal in origin. However, normal investigations do not preclude making a diagnosis of epilepsy, which remains primarily clinical.

The classification of epilepsy is made at three levels: seizure type (focal, generalised, unknown), epilepsy type (focal, generalised, combined generalised and focal, unknown), and by epilepsy syndrome.[10,11] At each stage, the aetiology and risk factors for developing epilepsy (structural, genetic, infectious, metabolic, immune and unknown), alongside the burden of comorbidities should be identified as these may have important implications for therapy.[1]

Antiepileptic drugs (AEDs) are the mainstay of therapy for epilepsy, with the aim of achieving seizure remission at the earliest opportunity without causing treatment-related adverse effects. Epilepsy is considered to be in remission for individuals who either had an age-dependent epilepsy syndrome but are now

past the applicable age, or those who have remained seizure-free for the last 10 years and off AEDs for at least the last five years.[5] Approximately two thirds of PWE achieve remission with AED use.[1] Epilepsy is drug-resistant in the remainder.

1.1.2 The evolving concept of epilepsy-related mortality

1.1.2.1 Advancing the conversation beyond SUDEP: the need for a multiple-cause-of-death analysis

PWE are at increased risk of premature death.[12-16] The first few observational studies reporting this were published in the 1930s–1960s mainly in the United States (US) and Europe,[17-22] and there has since been a large growth in the number of studies reporting premature mortality in PWE worldwide.[12] Despite mounting evidence and awareness of the excess mortality faced by PWE in general, much of the lay and academic interest in this area has been focused on sudden unexpected death in epilepsy (SUDEP).[23] SUDEP is defined as the sudden, unexpected, witnessed or unwitnessed, nontraumatic and nondrowning death of a PWE with or without evidence of a seizure, excluding documented status epilepticus, and in whom post-mortem examination does not reveal a structural or toxicological cause for death.[24]

The increased attention on SUDEP is justified: SUDEP is devastating particularly because of the unexpected manner in which such deaths occur, frequently during sleep in an otherwise healthy young individual.[25] Funding for SUDEP research has been on the rise, and this has translated into increased availability of information and awareness about SUDEP amongst both patients and caregivers.[23] SUDEP-related literature increased more than 4-fold from

2000–2004 (74 published articles) to 2010–2014 (324 published articles).[23] However, SUDEP is only one method by which death can occur in PWE.[23] In our clinical experience, this is not often discussed, even amongst specialists. It is clear that some of the remaining deaths in PWE may be entirely unrelated to epilepsy.[14] However, there is an emerging view that a substantial proportion of the deaths are likely to relate to the epilepsy itself, its treatment, or comorbidities worsened by epilepsy.[12,23] Together, these are termed 'epilepsy-related deaths', and SUDEP would be just one example of these.[23] There are several others, and a recently proposed classification system for epilepsy-related deaths organises them into four key groups:[23]

A. Deaths directly due to epilepsy;

- *SUDEP*
- *Status epilepticus*
- *Drowning*
- *Motor vehicle accidents*
- *Falls, burns, and other seizure-related accidents*

B. Deaths due to acute symptomatic seizures that would likely have gone on to become epilepsy (i.e. seizure recurrence risk would have been high if the person had survived);

- *Seizures with or without status epilepticus, occurring within one week of stroke, traumatic brain injury, anoxic encephalopathy, or brain surgery*
- *Subdural hematoma or central nervous system infection*
- *Active phase of multiple sclerosis or other autoimmune disorders*

C. Deaths indirectly due to epilepsy;

- *Aspiration pneumonia*
- *Suicide*
- *Cardiovascular disease that may be exacerbated or caused by AEDs*

D. Deaths occurring due to the underlying aetiology of epilepsy;

- *Brain tumour*
- *Stroke*
- *Metabolic/genetic and neurodegenerative diseases*

There is little research quantifying epilepsy-related mortality in such a broad and inclusive manner,[12] and there are many situations in which this has probably resulted in an underestimation of the potential impact of epilepsy-related mortality. For example, a large study in the UK looking at the long-term risk of premature mortality in PWE using The National General Practice Study of Epilepsy (NGPSE) identified that the leading cause of death in PWE was pneumonia, where PWE were 7.2–9.6-times more likely to die from this than people in general population.[26] However, the study then concluded that few people died from epilepsy-related causes. It is difficult to imagine how this could have been the case when having a seizure can lead to the development of pneumonia through aspiration of gastric content.[23,27,28] However, as was the case in this NGPSE study, studies investigating mortality commonly simplify assessing causes by counting only the underlying cause of death in the death certificate and excluding any contributory mechanisms listed alongside the underlying cause.[29-31] In this manner, even if epilepsy were listed alongside

pneumonia as a contributory mechanism by which death had occurred, it would not be counted. A better way to capture the various mechanisms by which epilepsy-related deaths may be occurring would be to conduct a multiple-cause-of-death (MCOD) analysis in which both the underlying and all contributory mechanisms listed on a death certificate are considered.[29-31] Such an analysis is yet to be reported in epilepsy mortality literature.[12]

Another example illustrates the potential impact of epilepsy-related mortality underestimation. A recent meta-analysis in the US demonstrated that SUDEP ranks second only to stroke as the leading neurological cause of total years of potential life lost (YPLL).[32] YPLL is a measure that takes into account the age at which deaths occur by giving greater weight to deaths at younger age and lower weight to deaths at older age.[33] It is calculated from the number of deaths in a population multiplied by the standard life expectancy at the age at which the deaths occurred.[33] The meta-analysis showed that stroke lead with 189,000 YPLL, and this was followed closely by 101,000 YPLL from SUDEP.[32] However, this study did not consider the impact including the other epilepsy-related causes of death would have had on these estimates. With an estimated annual rate of ~12/10,000 PWE,[23] it is possible that SUDEP may account for less than half of all epilepsy-related deaths when they are reclassified according to A–D above, as the annual mortality rate per 10,000 PWE was 5 from status epilepticus,[34] 2 from drowning,[35] 2–3 from suicide,[36,37] and 6 from unintentional injuries.[23,38] Together, these non-SUDEP epilepsy-related causes would have a higher mortality rate than SUDEP at ~16/10,000 PWE,

meaning that if they were considered alongside SUDEP, epilepsy would potentially become the leading neurologic cause of YPLL.[23]

1.1.2.2 Identifying the proportion of potentially avoidable epilepsy-related deaths: the need for a novel method

Many epilepsy-related causes of death are thought of as potentially avoidable.[15,39,40] Educating PWE regarding the risk of drowning, driving, the importance of AED compliance, as well discussing SUDEP risk could perhaps help avoid some deaths.[39] Others could be potentially be avoided by abolishing seizures and thereby preventing injuries, treating comorbidities such as epilepsy-related depression to help avoid suicides, and perhaps lowering the index of suspicion for pneumonia in PWE to rationalise early antimicrobial therapy.[12]

There is very little data quantifying what proportion of epilepsy-related deaths may be avoidable, particularly in adults.[12,41,42] This is likely due to the difficulties in establishing robust methods for identifying potentially avoidable deaths. Novel methods are required, perhaps taken from how potentially avoidable mortality has been quantified on a population-level outside of epilepsy literature.[43-45] For example, methods designed by The Office for National Statistics (ONS) in the UK have been used to quantify potentially avoidable mortality in Scotland as part of the assessment of health care, hospital admissions, and mortality by ethnicity.[43,45] They have also been used to quantify potentially avoidable deaths in children and young people with intellectual disabilities in Scotland.[44]

1.1.2.3 Identifying those at high risk of epilepsy-related death: the need for a risk prediction model

There are as yet no clinical tools available to identify which patients are at highest risk of suffering an epilepsy-related death.[12] Once again, this area is dominated primarily by SUDEP research. The first tool to be developed was a seven-item inventory (the SUDEP-7 inventory), which was proposed in 2010 by DeGiorgio *et al.* and revised in 2015 using a modified scoring methodology.[46,47] The SUDEP-7 inventory was based on a large prospective cohort study by Walczak *et al.* in 2001 identifying several highly predictive risk factors for SUDEP,[48] summarised in Figure 1-2.[47] Risk factors with low odds ratios (0–2) were not included.[47] Following on from the SUDEP-7 inventory was development of a three-item SUDEP-3 inventory, consisting of generalised tonic-clonic seizure (GTCS) frequency, seizure frequency, and intellectual disability.[49] This has been shown to outperform SUDEP-7 in predicting SUDEP.[49] More recently, a personalised prediction tool for SUDEP was developed by Jha *et al.* in which prognostic factors identified included focal and GTCS frequency, alcohol excess, younger age at epilepsy onset, and family history of epilepsy.[50] However, these risk prediction tools are all specific for SUDEP and cannot, therefore, be generalised to trying to predict epilepsy-related deaths more broadly.

SUDEP-7 RISK FACTOR INVENTORY					
SUDEP RISK FACTORS		ODDS RATIO/ CONFIDENCE INTERVAL	WEIGHTING LOG_E X O.R.	YES OR NO	SCORE
1	More than 3 Tonic-Clonic Seizures In Last Year	**8.1** [2.2-30] Walczak et al 2001	0 or 2		
2	1 or more Tonic-Clonic Seizures In Last Year	**2.4** [1.8-30.5] Walczak et al 2001	0 or 1		
3	One or more seizures of any type over the last 12 months	**2.2, 3.8, 4.6**** [0.3-14.4, .8-19.2, .5-40.9] Walczak et al 2001	0 or 1		
4	> 50 seizures of any type per month over the last 12 months	**11.5** [1.3-99.3] Walczak et al 2001	0 or 2		
5	Duration of Epilepsy ≥ 30 years	**13.9** [3.4-57.1] Walczak et al 2001	0 or 3		
6	Current use of three or more anti-epileptic drugs	**4.0** [1.4 – 11.7] Walczak et al 2001	0 or 1		
7	Mental Retardation, I. Q. < 70, or too impaired to test	**5.0** [1.3-19.3] Walczak et al 2001	0 or 2		
Total Weighted Score (0- 12)					
**Multiple Odds ratios reflect the following average frequencies of any seizures in the last month over a 12-month period: ≤1, > 1 to ≤ 15, >15 to ≤ 50, from Walczak et al 2001.					

Figure 1-2: SUDEP-7 Risk Inventory, derived from the large prospective cohort study reported by Walczak 2001[48]

Figure reproduced with permission from DeGiorgio 2010[47]

The SUDEP and Seizure Safety Checklist was also proposed recently.[51] This lists risk factors linked to epilepsy mortality including (but not restricted to) SUDEP as active seizures or injury, GTCS, status epilepticus, prolonged seizures, nocturnal seizures or lack of surveillance, poor medicine adherence, medication changes, alcohol and/or substance misuse, depression or other psychiatric disorders, the presence of intellectual disability, pregnancy, unclear

seizure frequency or treatment history, >15 years epilepsy duration or early-onset before 16 years of age, young adult age (20–40 years), and male sex.[51] Although extremely valuable as a clinical guide to prompt clinicians to think about epilepsy-related mortality when seeing patients, the SUDEP and Seizure Safety Checklist does not actually quantify an individual's risk of epilepsy-related death as no multivariable regression modelling was used to assess the proposed variables against one another, which were instead sourced from previous studies and put together through expert consensus.[51] There remains need for the development of a pragmatic risk stratification tool that can be used to identify individuals at high risk of suffering epilepsy-related death and quantify the specific risks.

1.1.3 Identifying epilepsy-related deaths in Scotland

1.1.3.1 Rationale

There are approximately 600,000 PWE in the UK, and nearly 60,000 PWE live in Scotland.[52] A nationwide study quantifying epilepsy-related deaths is yet to be undertaken in Scotland.[12] From a public health perspective, this is a pressing issue as Scotland has a lower life expectancy than the rest of the UK and Western Europe.[53] It also has a larger proportion of its population living in rural and remote communities compared with the rest of the UK.[53] This inevitably presents challenges in epilepsy healthcare provision.[43] A study of epilepsy-related deaths in such a Scottish population is likely to highlight valuable results when contextualised against neighbouring England, where epilepsy has been shown to be the fourth most common neurological cause of death, with a greater proportion of the epilepsy-related deaths in England occurring in areas of higher levels of deprivation.[54]

1.1.3.2 The Scottish administrative healthcare data model

The study of epilepsy-related mortality on a national scale requires both a nationwide administrative healthcare data infrastructure and accurate case-ascertainment methods for epilepsy and causes of death within that infrastructure. Administrative healthcare datasets consist of information routinely collected about patients when they use healthcare services. Such data can include clinical information, for example diagnosis, treatment, care service attendance, or causes of death. The data can also include demographic information, such as age, gender, home postcode, and area of socioeconomic deprivation inhabited.

Healthcare to the five-and-a-half million population of Scotland is administered by a single state-provided National Health Service (NHS), which is managed by Scottish Government and allocated through 14 regional Scottish Health Boards (Figure 1-3).[55]

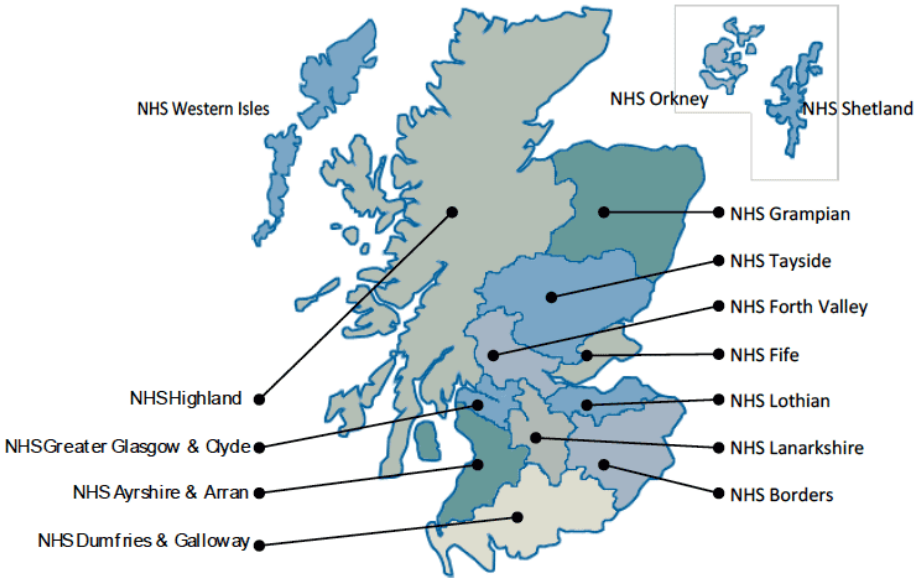


Figure 1-3 Regional Scottish NHS Health Boards

Figure reproduced with permission from McCallum 2014[56]

Scottish NHS care is funded by general taxation and is therefore free to patients at the point of use, including all medicines.[55] One of the advantages of such a model is that nearly all persons seeking healthcare in Scotland use the NHS, and there is very little private healthcare provided (<2% of care, information from which is fed back into NHS records).[55] To access NHS care, patients in Scotland are first provided with a unique and lifelong patient identification number (called a Community Health Index (CHI) number) upon registration with a general practitioner. This makes it possible for the entire Scottish population to be studied in a data linkage research environment using CHI numbers as the central identification spine upon which various administrative datasets are added.[43,55,57] The data are linked by the electronic Data Research and Innovation Service (eDRIS) at Farr Institute Scotland, and made available to approved researchers as deidentified datasets in a secure analytical platform (a data Safe Haven) – see figure 1-4.[57] Administrative economic and social data which do not contain CHI numbers are also available to supplement health data in order to enhance the understanding of socioeconomic influences on health.[55] These are linked to health data using probabilistic linkage techniques that match the data by other identifying information (e.g. names, addresses or dates of birth).[55]

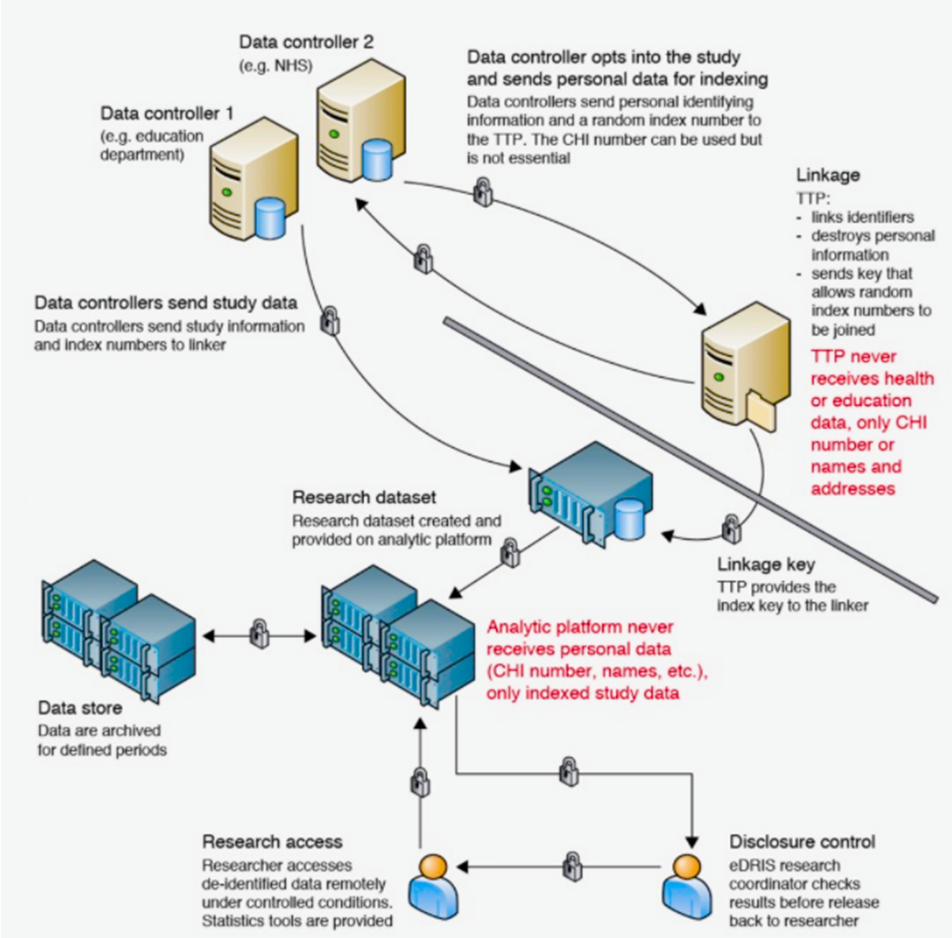


Figure 1-4: Scottish data linkage processes

Abbreviations: CHI – Community Health Index; eDRIS – eData Research and Innovation Service; NHS – National Health Service; TTP – Trusted third party.

Figure reproduced with permission from Pavis 2015[55]

There are many examples of administrative datasets used routinely for data linkage research in Scotland (Figure 1-5). These have been collected at the national level for over 45 years, with the Information Services Division of National Services Scotland (ISD) responsible for ensuring completion, quality, and comparability across geographic regions and health services.[55] The Health Boards hold further clinical information from medical records that can be used by researchers to enhance the clinical phenotype of study

participants. This data linkage model has resulted in Scotland possessing one of the most well-developed and complete administrative data infrastructures in the world,[55] integrating well with Health Data Research UK (HDR UK) – the UKs national institute for health data science.[58]

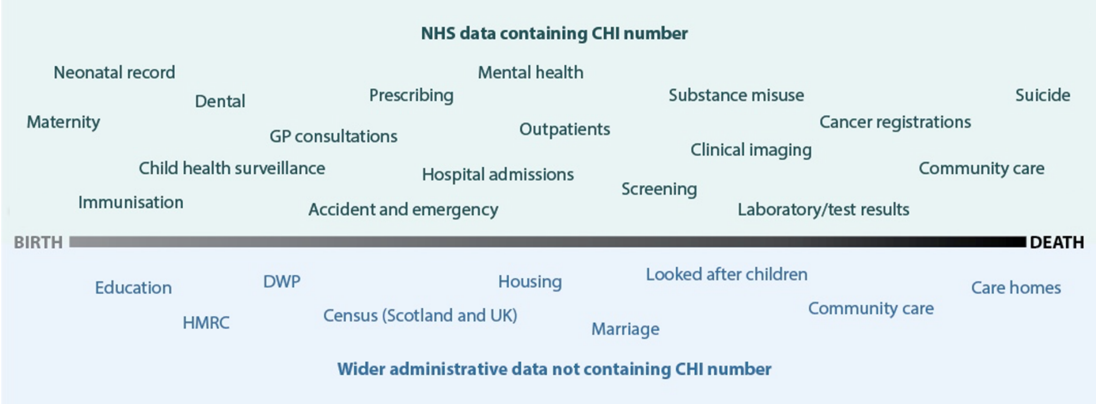


Figure 1-5: Examples of national-level administrative datasets in Scotland

Abbreviations: NHS – National Health Service; CHI – Community Health Index (unique NHS patient identification number; DWP – Department for Work and Pensions; HMRC - Her Majesty's Revenue and Customs

Figure reproduced with permission from Pavis 2015[55]

1.1.3.3 Validating the diagnostic accuracy of administrative healthcare data for epilepsy

The International Classification of Diseases 10 system (ICD-10) is used to code causes of death derived from death certificates to create a nationwide administrative mortality dataset held at the National Records of Scotland (NRS).[59] Scotland has a world-leading mortality review system in place in which these causes of death are routinely scrutinised for accuracy and amended, where necessary.[60,61] Within the ICD-10 system, the G40 code is used for epilepsy, G41 for status epilepticus, and R56.8 for seizures.[62] ICD-10 Scottish hospital admission diagnoses are also available in a nationwide manner, held within the Scottish Morbidity Record 01 (SMR01) dataset.[63] The

SMR00 dataset holds nationwide information on hospital outpatient attendance and clinic speciality.[64] Primary care diagnoses are coded using the Read code system. Read codes are a coded thesaurus of clinical terms used in primary care to create electronic health records.[65] *F25* Read codes are used for epilepsy.[62] A nationwide primary care dataset is as yet unavailable in Scotland but can be set up regionally using published methods.[66]

The diagnostic accuracy of any administrative healthcare dataset requires validation prior to scientific use within a research project.[67-69] This is done to identify the optimal case ascertainment method for the disease to be studied and to increase transparency about the validity of any conclusions drawn from that particular study. Although advantaged by being able to capture entire populations in an inexpensive manner, one of the challenges of using administrative data is that the disease coding can be inaccurate owing to several factors, such as human coding errors, incomplete hospital discharge summaries, and inaccurate coding reflective of genuine diagnostic uncertainty amongst clinicians.[67-69] Combining diagnostic coding data with additional information, such as drug prescriptions, can enhance case ascertainment accuracy. Various combinations of codes and drugs are tested to identify the optimal case-ascertainment algorithm, as assessed using positive predictive value (PPV) and sensitivity, as well as negative predictive value (NPV) and specificity, where possible.[62] The optimal coding strategy to identify epilepsy within administrative Scottish healthcare datasets is yet to be determined. This uncertainty is underpinned by lack of systematic review evidence for the

optimal coding strategies for capturing epilepsy within administrative data in general.

1.1.3.4 Permissions required for administrative data research in Scotland

All studies using administrative Scottish datasets require Research Ethics Committee approval, as well as approval from the Public Benefit and Privacy Panel for Health and Social Care (PBPP).[70] This is a patient advocacy panel that acts as a national Caldicott Guardian to scrutinise applications for access to NHS Scotland health data. The role of PBPP is to ensure that the research projects have sufficiently considered the public benefit and privacy implications of the proposed work and adhere to the relevant Data Protection Legislation.[70] Approved researchers also require appropriate Information Governance and General Data Protection Regulation (GDPR) training to access the datasets.[71]

1.2 Summary

Epilepsy is a common neurological condition, associated with an increased risk of premature death. Some of the deaths may be unrelated to epilepsy. However, a substantial proportion are thought to related to the epilepsy itself, its treatments, or comorbidities associated with epilepsy. These are called ‘epilepsy-related deaths’. The various mechanisms by which epilepsy-related deaths occur remain unclear. It is also unclear what portion of epilepsy-related deaths are potentially avoidable. Furthermore, work is also needed to identify which individuals are at high risk of suffering epilepsy-related death.

Administrative healthcare data contain routinely collected information about patients when they use healthcare services. They provide a potentially powerful tool for investigating epilepsy-related deaths on a nationwide scale if epilepsy can be captured accurately within these datasets. Validating the case-ascertainment accuracy of administrative data in capturing epilepsy and establishing the optimal algorithms for doing so are necessary prerequisite steps before epilepsy-related mortality can be confidently investigated using administrative data.

1.3 Aims of this thesis

The aims of this thesis are:

1. To estimate the global burden of all-cause as well as epilepsy-related mortality in PWE;
2. To estimate the overall accuracy of using administrative healthcare data to identify epilepsy cases;
3. To identify the optimal algorithms for identifying epilepsy cases in deceased adults in Scotland using administrative data;
4. To use administrative data from adults to investigate the trends and mechanisms of epilepsy-related deaths in Scotland, highlighting the proportion that were potentially avoidable;
5. To identify multiple risk factors for the occurrence of epilepsy-related death in Scottish adults and develop a pragmatic risk prediction model for identifying which individuals are at the highest risk.

In Chapter Two, we deal with the aim to estimate the global burden of all-cause as well as epilepsy-related mortality in PWE by undertaking a systematic review overview describing the comparative risks, causes, and risk factors for all-cause mortality in PWE at any age, highlighting the burden that were likely to have been epilepsy-related deaths. This chapter therefore provides a general background on the current literary understanding surrounding epilepsy-related deaths and which areas require further study. A novel approach of reclassifying the data according to whether or not deaths were likely to have been epilepsy-related is taken using a recently published classification system.[23] Although administrative data are used to capture

deceased epilepsy populations in many of the studies reviewed, none validate how accurately such datasets actually capture epilepsy.

Chapters Three and Four focus on ascertaining how accurately administrative data can capture epilepsy. In Chapter Three, we undertake the first global systematic review of studies that have validated the accuracy of using administrative healthcare data to identify epilepsy cases. This leads on to Chapter Four, in which we undertake our own study validating the accuracy of using administrative data to identify epilepsy cases in deceased adults in Scotland, with a focus on identifying the optimal case-ascertainment algorithms involving epilepsy cause-of-death coding. This is the first reported study to validate administrative epilepsy data in deceased adults.

Chapter Five is a nationwide administrative data linkage study investigating trends and mechanisms of epilepsy-related deaths in Scotland, highlighting the proportion that were potentially avoidable. We use the optimal case-ascertainment algorithms found in Chapter Four to identify the total number of adult epilepsy-related deaths in Scotland, whether these have changed over time, and how they compare to risk of mortality in the general population by age. We then describe all of the mechanisms by which epilepsy-related deaths occurred in the age groups at increased risk using an MCOD analysis.[29-31] We estimate what proportion of these deaths were potentially avoidable using a method provided by the ONS for identifying potentially avoidable mortality within deceased populations.[43-45]

Chapter Six aims to identify multiple risk factors for the occurrence of epilepsy-related death in Scottish adults and develop a pragmatic risk prediction model

for identifying which individuals are at the highest risk. This is an age- and sex-matched case-control study comparing adults who suffered epilepsy-related deaths to living adults with epilepsy in Scotland. Deceased cases of epilepsy-related death are sampled from those identified in Chapter Five. Controls living with epilepsy in Scotland are recruited from a large database of research volunteers and from epilepsy clinics. We analyse routine clinical information gathered from medical records to describe multiple potential risk factors for suffering epilepsy-related death and to develop a pragmatic risk prediction model for identifying those at increased risk.

In Chapter Seven, we summarise findings from the five studies, discuss their overall validity and implications, and we outline the directions for future research.

Chapter 2. Epilepsy-related and other causes of mortality in people with epilepsy: A systematic review of systematic reviews

2.1 Introduction

A comprehensive review of literature is required in order to estimate the global burden of all-cause as well as epilepsy-related mortality in PWE, thereby acting as a foundation to help refine the methodological approaches taken for the remainder of the aims in this thesis. Whilst a systematic review is strengthened by the ability to compile evidence in a systematic manner, studies can be missed within an individual systematic review if, for example, the search strategy does not cover all available electronic databases, or if aims of the systematic review are narrow compared to the breadth of data available on a subject. There can also be limitations in the quality of an individual systematic review, placing limitations on the validity of conclusions drawn. A more comprehensive approach is to undertake a systematic review of the available systematic reviews on a subject. This is a relatively new method of evidence synthesis first proposed as part Cochrane Reviews in 2011.[72] This method has the advantage of allowing the different search strategies used within different systematic reviews to be, in effect, combined; thereby bringing together a greater body of evidence. Such an approach also allows the quality of multiple systematic reviews to be evaluated, potentially improving the validity of conclusions drawn over and above those made within the individual systematic reviews or primary studies evaluated.

In the current chapter, we undertake a systematic review of six epilepsy mortality systematic reviews, bringing together an analysis of 103 primary

studies evaluating mortality in PWE. This is a larger number of primary studies evaluated than in any previous systematic review of mortality in PWE and allows us to more confidently discuss what the potential burden of epilepsy-related deaths was amongst the all-cause deaths captured. This systematic review was published in *Epilepsy Research* in 2019.[12] The published manuscript has been used as the body of this thesis chapter, with permission to do so granted from publishers. References listed in the published manuscript are integrated into the thesis references.[13-16,23,25,35,36,39,41,51,73-118]

Student contributions: I came up with an idea to undertake a systematic overview of the systematic review literature on epilepsy-related deaths globally as I felt this would be a unique and effective method to try. I then designed the searchers, screened studies, abstracted data, and undertook risk of bias assessments, with the support of a second reviewer undertaking these steps in parallel. I analysed the data then drafted and successfully submitted the work for publication.

2.2 Manuscript

Epilepsy Research 157 (2019) 106192

Contents lists available at ScienceDirect

Epilepsy Research

journal homepage: www.elsevier.com/locate/epilepsyres

Review article

Epilepsy-related and other causes of mortality in people with epilepsy: A systematic review of systematic reviews

Gashirai K. Mbizvo[a,d,*], Kyle Bennett[a], Colin R. Simpson[b,c], Susan E. Duncan[a,d], Richard F.M. Chin[a,e]

[a] Muir Maxwell Epilepsy Centre, Centre for Clinical Brain Sciences, The University of Edinburgh, Edinburgh, UK
[b] Faculty of Health, Victoria University of Wellington, Wellington, New Zealand
[c] Usher Institute of Population Health Sciences and Informatics, The University of Edinburgh, Edinburgh, UK
[d] Department of Clinical Neurosciences, Western General Hospital, Edinburgh, UK
[e] Royal Hospital for Sick Children, Edinburgh, UK

ARTICLE INFO

Keywords:
SUDEP
Death
Seizures
Overview of reviews
Cohort study
Case-control study

ABSTRACT

Background: This systematic review of epilepsy mortality systematic reviews evaluates comparative risks, causes, and risk factors for all-cause mortality in people with epilepsy (PWE) to specifically establish the burden of epilepsy-related deaths.

Methods: MEDLINE and Embase were searched from conception to 26/12/2018 for systematic reviews evaluating all-cause mortality in PWE of any age. Independent study selection, data extraction and quality assessment were performed. Deaths were separated into epilepsy-related and unrelated using a recently published classification system. Outcomes included standardized mortality ratio (SMR) and mortality rate (MR) in a primary analysis of comparative risks, causes, and risk factors for all-cause and epilepsy-related mortality. A narrative synthesis of review findings was used to present results, including from a secondary analysis of individual epilepsy-related death risk factors.

Results: Six moderate or high-quality systematic reviews were included in the primary analysis, evaluating 103 observational studies. All-cause mortality remained similarly high between 1950 and present (median SMR range 2.2–3.4). Africa had the highest SMR (median 5.4, range 2.6–7.2). SMRs were also higher for children < 18 years (median 7.5, range 3.1–22.4) than adults (median 2.6, range 1.3–8.7), and for epilepsy-related (median 3.8, range 0.0–82.4,) than unrelated causes (median 1.7, range 0.7–17.6). Structural brain disease conferred the greatest risk for all-cause mortality (SMR range 24.0–41.5). Common epilepsy-related causes included alcohol, drowning, pneumonia, and suicide. In secondary analysis of nine additional systematic reviews, epilepsy-related death risk factors were reported for sudden unexpected death in epilepsy (SUDEP), drowning and suicide.

Conclusions: Premature all-cause mortality remains a major problem in PWE globally, particularly in children and young adults, with most being epilepsy-related and potentially preventable. SUDEP is only one of several other common and important epilepsy-related causes of death.

1. Introduction

There are 70 million people with epilepsy (PWE) worldwide, and epilepsy contributes to 0.7% of the global burden of disease (Fazel et al., 2013; Mbizvo et al., 2012). Although PWE are subject to the same causes of death as the population without epilepsy, mortality in PWE is greater because they also die of epilepsy, its treatment, or its co-morbidities: so-called “epilepsy-related deaths” (Devinsky et al., 2016). Avoidable deaths in PWE tend to be those that are epilepsy-related because the greater risk of death over the general population is abolished by, for example, achieving seizure freedom through establishing effective treatment strategies (Neligan and Sander, 2011; Neligan and Bell, 2009). However, primary studies and systematic reviews reporting mortality in PWE are yet to focus on epilepsy-related mortality *specifically* perhaps because until the recently published Call to Action in recognizing and preventing epilepsy-related mortality by Devinsky

* Corresponding author at:Muir Maxwell Epilepsy Centre, University of Edinburgh, 20 Sylvan Place Edinburgh, EH9 1UW, UK.
E-mail addresses: gashirai.mbizvo@ed.ac.uk (G.K. Mbizvo), s1103639@sms.ed.ac.uk (K. Bennett), c.simpson@ed.ac.uk (C.R. Simpson), susan.duncan@ed.ac.uk (S.E. Duncan), rchin@exseed.ed.ac.uk (R.F.M. Chin).

https://doi.org/10.1016/j.eplepsyres.2019.106192
Received 6 May 2019; Received in revised form 15 August 2019; Accepted 25 August 2019
Available online 27 August 2019
0920-1211/ © 2019 Elsevier B.V. All rights reserved.

Table 1
Classification of epilepsy-related deaths by Devinsky et al. (2016).

Deaths directly due to epilepsy	• SUDEP • Status epilepticus • Drowning • Car accidents • Falls • Burns • Other seizure-related accidents
Deaths due to acute symptomatic seizures	• With or without status epilepticus, occurring within one week of stroke, traumatic brain injury, anoxic encephalopathy, or intracranial surgery • First identification of subdural hematoma or central nervous system infection • Active phase of multiple sclerosis/autoimmune disorders
Deaths indirectly due to epilepsy	• Aspiration pneumonia • Suicide • Cardiovascular disease caused by antiepileptic drugs
Deaths due to underlying neurologic disease	• Brain tumor • Stroke • Metabolic [in which we included alcohol]*/genetic/neurodegenerative disease

Abbreviations: SUDEP = sudden unexpected death in epilepsy || *Alcohol included owing to alcohol-related structural brain changes resulting in a strong and consistent association between chronic alcohol consumption and unprovoked seizures (i.e. epilepsy) outside of withdrawal seizures (Samokhvalov et al., 2010).

et al. (2016), there were no systematic methods for classifying the deaths in PWE into those that are most likely to be epilepsy-related. We have now capitalized on these methods by conducting a systematic review of the systematic reviews that have investigated mortality in PWE in order to ascertain: 1) the quality of individual systematic reviews; 2) the rate and comparative risks for all-cause mortality; 3) causes of death and the proportion that are epilepsy-related; and 4) risk-factors for all-cause and epilepsy-related mortality. We took the novel approach of classifying the data according to whether or not deaths were likely to be epilepsy-related using this recent Devinsky classification system (Devinsky et al., 2016) (Table 1). We also aimed to elucidate the relevant areas for future research.

2. Methods

2.1. Protocol and registration

An *a priori* study design was provided through registration of a protocol on PROSPERO (CRD42018102544 available from https://goo.gl/2UpNQM). The research questions and inclusion criteria were established before conduct of the review (Shea et al., 2007b,a; Shea et al., 2009). The review follows the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) system of reporting (Moher et al., 2009). Ethical approval was not required.

2.2. Eligibility criteria

We applied the following inclusion criteria:

i *Study design:* systematic reviews, i.e. a systematic search strategy was applied to one or more electronic databases *and* findings were summarized either narratively or in a formal meta-analysis (Clarke, 2008);
ii *Population of interest:* the reviews covered PWE who have died at any age, anywhere in the world;
iii *Evaluation of interest:* the reviews reported rate or comparative risk, all-cause mortality, or different risk factors for mortality in PWE;
iv *Outcomes of interest*: the reviews reported standardized mortality ratio (SMR), relative risk (RR), odds ratio (OR), hazard ratio (HR), mortality rate (MR), rate ratio, case fatality (CF), or proportionate mortality ratio (PMR).

We applied the following exclusion criteria:

i Systematic reviews focusing only on a specific individual cause of death in PWE unless the cause of death was epilepsy-related (Table 1) and the systematic review ascertained risk factors for this epilepsy-related cause of death. This exclusion meant we retained the ability to calculate the proportion of all-cause mortality attributable to epilepsy whilst not missing information on the risk factors for individual epilepsy-related causes of death (which may be reported in separate systematic reviews to those on all-cause mortality);
ii Systematic reviews of mortality in highly selected subpopulations of PWE only (e.g. intellectual disability or those undergoing epilepsy surgery) if not consistent with exclusion i. above. This was aimed at increasing generalizability of the review findings (Nevalainen et al., 2014; Watila et al., 2018);
iii Systematic reviews unable to capture any new primary studies/outcomes not already captured by other included systematic reviews. This would reduce the risks of reporting duplicate information.
iv Systematic reviews focusing only on status epilepticus, febrile convulsions, or infantile spasms, as these patient cohorts do not necessarily have epilepsy.

2.3. Information sources, literature searches and study selection

MEDLINE and Embase (including Embase grey literature) were searched from their conception to 26/12/2018 (date last searched - 30/12/18) for potentially eligible systematic review articles (Moher et al., 2009). The full search strategy is detailed in Appendix A. Reference lists of potentially eligible systematic reviews were also hand searched to identify any other potentially eligible systematic reviews. Two study authors (GKM and KB) independently screened titles, abstracts, or full-length articles against the inclusion/exclusion criteria and selected the eligible articles for inclusion. Any disagreements in article selection were resolved by consensus or, where necessary, adjunction by a third reviewer (RFMC). There were no language restrictions and, where necessary, translations were sought (Moher et al., 2009).

2.4. Data abstraction

Two study authors (GKM and KB) independently abstracted data pertaining to study design, population demographics, and mortality rate, causes, and risk factors from included systematic review articles using a piloted data abstraction form (Appendix B). Any differences in the abstracted study results between the two authors were resolved by mutual review of the relevant study article. Individual study data were sought from within the systematic reviews and their supplementary

material tables, where available, and any missing data were sought from the original study publications.

2.5. Quality assessment of reviews

A MeaSurement Tool to Assess systematic Reviews (AMSTAR) (Shea et al., 2007b,a; Shea et al., 2009) was used to assess systematic review quality. This validated tool assesses the degree to which review methods have avoided bias by evaluating them against 11 distinct criteria. Systematic reviews with an AMSTAR score of ≤3, 4–7, or 8–11 are considered low-, moderate-, or high-quality, respectively (Ryan et al., 2014). Two authors (GKM and KB) independently assessed the quality of included reviews against the AMSTAR tool. Disagreements in scoring were resolved by consensus or, where necessary, adjunction by a third reviewer (RFMC).

2.6. Summary measures and Synthesis of results

Two review authors (GKM and KB) used Devinsky's classification system of epilepsy-related deaths (Devinsky et al., 2016), summarized in Table 1, to independently organize the causes of death identified within systematic reviews according to whether or not they were likely to be epilepsy-related. Any disagreements were resolved by consensus or, where necessary, adjunction by a third reviewer (RFMC). Meta-analysis across systematic reviews was not attempted because of the already well-described heterogeneity known amongst epilepsy mortality data (Levira et al., 2016; Shackleton et al., 2002; Watila et al., 2018), which has precluded previous attempts at meta-analysis (Levira et al., 2016; Shackleton et al., 2002; Watila et al., 2018) and resulted in cautioning against this (Keezer et al., 2015; Logroscino and Hesdorffer, 2005). Therefore, we performed a narrative synthesis of review findings to present results (Clarke, 2008; Corry et al., 2015; Lunny et al., 2017; Smith et al., 2011). Primary outcome was SMR, the most frequently used outcome measure across epilepsy mortality studies. Secondary outcomes were mortality rate, relative risk, hazard ratio, case fatality, odds ratio, proportionate mortality ratio, and rate ratios. These outcomes were used to summarize information in a primary analysis around the rate and comparative risks for all-cause mortality, causes of death and the proportion that are epilepsy-related, and risk-factors for all-cause and epilepsy-related mortality. A preliminary screen of the data for distribution, using histograms and Shapiro–Wilk testing, indicated that they were not normally distributed ($p < 0.05$ on Shapiro–Wilk test). Therefore, medians and ranges were reported for the primary analysis (Levira et al., 2016; Thurman et al., 2017). Subgroup analysis would be used to explore the results from any low-quality systematic reviews (AMSTAR quality assessment) (Shea et al., 2007b,a; Shea et al., 2009) as compared to those from moderate- and high-quality systematic reviews combined. Our secondary analysis consisted of a narrative summary of systematic reviews that identified risk factors for an individual epilepsy-related cause of death. All analyses were performed and graphical images created using IBM SPSS Statistics Version 24, Armonk, NY: IBM Corp.

3. Results

The MEDLINE and Embase searches retrieved 47 and 238 records, respectively. Fig. 1 summarizes how these were handled. Six systematic reviews were included in the primary analysis (Escalaya et al., 2015; Levira et al., 2016; Nevalainen et al., 2014; Shackleton et al., 2002; Thurman et al., 2017; Watila et al., 2018). These will be referred to as Watila (2018), Levira (2016), Thurman (2017), Escalaya (2015), Nevalainen (2014), and Shackleton (2002) in the remaining text. There were 192 studies reported across the six systematic reviews: 63 in Watila (2018), 18 in Levira (2016), 46 in Thurman (2017), 5 in Escalaya (2015), 37 in Nevalainen (2014), and 23 in Shackleton (2002). 142 of these studies featured in more than one systematic review (with 25 studies overlapping two systematic reviews, 20 studies overlapping three systematic reviews, and eight studies overlapping four systematic reviews). For overlapping studies, we extracted their results from a single systematic review only in order to avoid reporting the same data more than once. Therefore, excluding duplicate data, the primary analysis reports results from 103 studies (Appendix D). 50 of these (49%) did not overlap across systematic reviews and therefore contribute independent data. The secondary analysis is a narrative synthesis of risk factors for epilepsy-related deaths from nine systematic reviews investigating a single epilepsy-related death in each. Seven systematic reviews assessed risk factors for sudden unexpected death in epilepsy (SUDEP) (Abdel-Mannan et al., 2018; Ali et al., 2017; Harden et al., 2017; Liebenthal et al., 2015; Monte et al., 2007; Tellez-Zenteno et al., 2005; Young et al., 2015). One assessed risk factors for drowning (Bell et al., 2008) and a related review dealt with suicide (Bell et al., 2009). A reference list of excluded records is provided in Appendix C (Moher et al., 2009).

3.1. Primary analysis – systematic reviews assessing all-causes and risk factors of mortality in people with epilepsy

3.1.1. Characteristics of included systematic reviews

Table 2 summarizes the baseline characteristics of systematic reviews included in the primary analysis. Studies were largely sourced from MEDLINE, Embase, and Latin American and Caribbean Health Sciences Literature (LILACS). There were no age restrictions for the populations covered by each systematic review. All studies were observational (cohort or case-control designs). They covered both incident cases (mortality in people recruited into a study when newly diagnosed with epilepsy) and prevalent cases (mortality in people already known to have epilepsy when recruited), as well as both hospital- and community-based source populations. Watila (2018), Levira (2016) and Shackleton (2002) discovered significant levels of heterogeneity between studies. Therefore, they did not pool results into a meta-analysis, instead providing a narrative synthesis of results that included medians and ranges. Thurman (2017) also provided a narrative synthesis of results with medians and ranges. Nevalainen (2014) discovered significant levels of heterogeneity between studies but still decided to pool results into a meta-analysis, and that approach was later criticized elsewhere by Thurman et al. (Keezer et al., 2015).

3.1.2. Quality of included systematic reviews

Watila (2018) and Nevalainen (2014) were high-quality, scoring 9 out of possible maximum score of 11 on the AMSTAR scale. The remaining reviews were moderate quality, scoring 7 out of 11 for both Levira (2016) and Thurman (2017), and 5 out of 11 for both Escalaya (2015) and Shackleton (2002). The few points lost were mainly for failure to mention a protocol or *a priori* published research objectives (six reviews), failure to provide a list of the excluded studies (five reviews), failure to disclose that grey/unpublished literature were sought (four reviews), and failure to assess the likelihood of publication bias (four reviews, Appendix E) (Shea et al., 2007b,a; Shea et al., 2009).

3.1.3. Overall risk of mortality

Fig. 2 demonstrates a panel of box plots indicating the overall median and range of SMRs stratified according to various subgroups of interest. Fig. 2A shows that mortality results were broadly similar between systematic reviews, with median SMR ranging between 2.3–3.6. This indicates that the overall mortality was similar for the two high-quality reviews (Watila, 2018 and Nevalainen, 2014) and the remaining moderate quality reviews. Two major study outliers identified in Nevalainen (2014) reported SMRs of 22.4 (Ackers et al., 2011) and 22.2 (Moseley et al., 2013). These studies were both in children (age < 18 years). Fig. 2B indicates that whilst both children and adults were at increased risk of death compared to the general population, this risk tended to be higher for children (SMR range 3.1–22.4, median 7.5) than

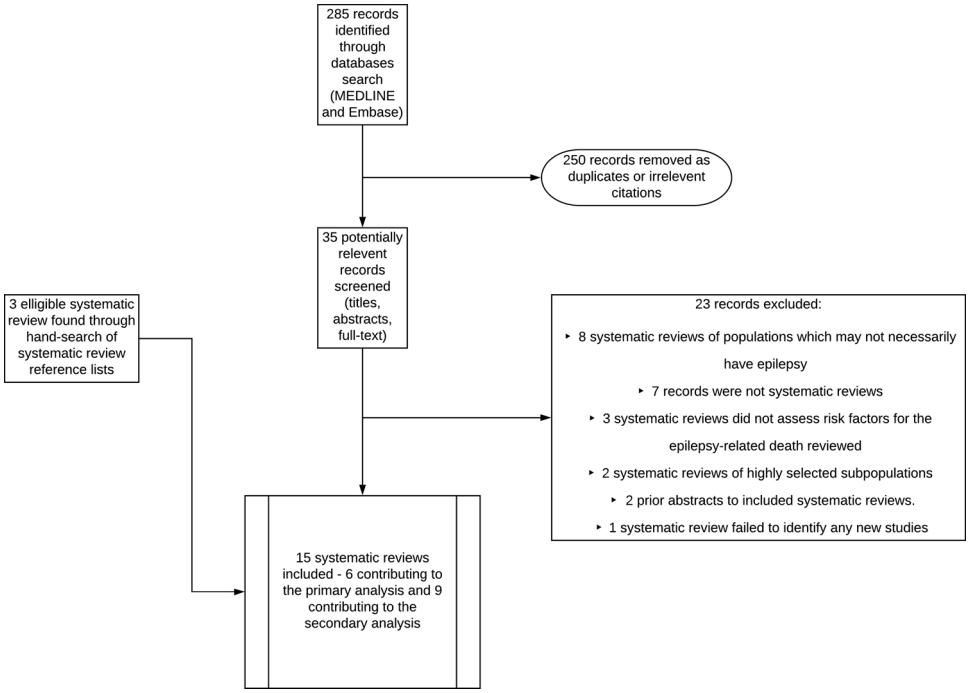


Fig. 1. Flow diagram of study selection process.
A reference list of excluded studies is provided in Appendix C.

for adults (SMR range 1.3–8.7, median 2.6). Fig. 2C shows that SMR results appeared broadly similar between community-based (median SMR 2.6) and hospital-based studies (median 3.1), with a tendency toward higher SMR results in hospital-based studies after the two pediatric community outliers were taken into account. Fig. 2D shows that SMR results appeared broadly similar between incident or prevalent epilepsy cohorts, with median SMR ranging 2.2–2.9 across these groups. Fig. 2E indicates that there has been little change in the overall mortality of epilepsy since the 1950s, with median SMR remaining increased between 2.2–3.4. It also demonstrates that the number of epilepsy mortality studies conducted per decade has been steadily increasing, with only one study per decade in the 1930s–1940s, rising up to 24–28 studies per decade since the year 2000. In Fig. 2F, continents demonstrate three levels of increased risk, with the highest being in Africa (SMR range 2.6–7.2, median 5.4), followed by Eastern Asia and North America (median SMR range 3.6–3.7), and then Europe, South Asia, and South America (median SMR range 2.5–2.6).

Only Levira (2016) provided overall mortality as an MR. This was presented as the median MR for community-based studies with quality scores > 80% (seven studies), with a rate of 19.8 deaths per 1,000 PWE per year (range 9.7–45.1), a weighted mean follow-up period of 5.8 years (range 1.5–10), and a total of 6,665 person-years observation. The median MR for community-based studies with lower quality scores was 13.1 deaths per 1,000 PWE per year (range 7.9–34.9), with a weighted mean follow-up period of 6.0 years (range 4.5–14) and total of 32,850 person-years observation. All of the hospital-based studies had quality scores < 50%. Their median MR was 7.1 deaths per 1,000 PWE per year (range 1.6–25.1), with a weighted mean follow-up period of 12.4 years (range 3–30) and total of 33,990 person-years observation.

For studies not reporting SMR or MR, Watila (2018) reported an overall HR range of 1.9–11.9, and Levira (2016) reported an overall CF of 8.1% (range 3.3–31.6%) for high-quality community-based studies, and 5.4% (range 1.3–75.3%) for hospital-based studies.

3.1.4. Causes of death in people with epilepsy

Four systematic reviews investigated cause of death (Escalaya et al., 2015; Levira et al., 2016; Thurman et al., 2017; Watila et al., 2018). Between these, 20 causes of death were reported. These came from various sources (Fig. 3). Verbal autopsies and family reports were more common in low- and middle-income countries, whilst high-income countries tended to use death certificates and death registries.

3.1.5. Epilepsy-related mortality

The mortality likely to be related to epilepsy was largely reported using SMR, MR and PMR. Visual inspection of the SMR data revealed that PWE were at increased risk of death from all of the epilepsy-related causes, with three levels of risk (Fig. 4A). The highest risk was from alcohol, brain tumors, fires, drowning, and falls, where median SMR ranged between 5.2–24.6. This was followed by risk from pneumonia, suicide, and injuries, where median SMR ranged between 3.2–3.9. Finally, cerebrovascular disease and transport accidents had a median SMR around 2.5. SUDEP and status epilepticus were not reported using SMR.

For PMR (Fig. 4B), visual inspection of the data also revealed three main groups of risk. The highest risk group died of causes that are likely to be indirectly due to epilepsy or due to underlying neurological disease (Devinsky et al., 2016), i.e. brain tumor, pneumonia and cerebrovascular disease, where PMR ranged between 19.7–50.0%. This was followed by status epilepticus, SUDEP, and drowning, where PMR ranged between 12.1–14.8%. The final group, each with PMR less than 10%, were suicide, burns, transport accidents, and falls.

SUDEP and status epilepticus were also investigated using MR (Fig. 4C). Rates tended to be higher for SUDEP (range 0.20–6.30, median 1.32 per 1,000 PWE per year) than for status epilepticus (range 0.09–0.97, median 0.32 per 1,000 PWE per year). More studies investigated SUDEP (32 studies) than any other cause of death.

One study from Thurman (2017) reported epilepsy-related causes of death using OR. They identified highest risk from falls, drowning, and suicide, with an OR range between 3.7–8.5 (Fazel et al., 2013). Lowest risk was from transport accidents and assault, ranging between 1.4–2.8 (Fazel et al., 2013). In Watila (2018), an MR rate ratio of 15.96 was reported for suicide (Christensen et al., 2015), alongside an HR of 2.54 (Nevalainen et al., 2013). HR was also 3.05 for accidents (Nevalainen et al., 2013), and 1.0–3.4 for alcohol (Nevalainen et al., 2012, 2013).

Table 2
Characteristics of included systematic reviews.

Study characteristics		Watila (2018)	Levira (2016)	Thurman (2017)	Escalaya (2015)	Nevalainen (2014)	Shackleton (2002)
Review objectives	Identify rate of mortality in people with epilepsy	✗	✓	✗	✓	✗	✗
	Identify comparative risk of mortality in people with epilepsy	✓	✓	✓	✓	✓	✓
	Assess cause of death in people with epilepsy	✓	✓	✓	✓	✗	✗
	Identify risk factors for mortality in people with epilepsy	✗	✓	✓	✓	✓	✗
Populations covered	Low- and middle-income countries	✓	✓	✗	✓	✓	✓
	High-income countries	✓	✗	✓	✗	✓	✓
Search strategy	MEDLINE	✓	✓	✓	✓	✓	✓
	Embase	✓	✓	✓	✓	✗	✗
	Ovid	✗	✗	✓	✗	✗	✓
	LILACS	✗	✓	✗	✓	✓	✗
	Scopus	✗	✗	✗	✗	✓	✗
	Web of Science	✗	✗	✗	✗	✓	✗
	PsycINFO	✗	✗	✗	✗	✓	✗
	Cochrane Library	✗	✗	✗	✗	✓	✗
	CINAHL	✗	✗	✗	✗	✓	✗
	IngentaConnect	✗	✗	✗	✗	✓	
	Cumulative Index Medicus	✗	✗	✗	✗	✗	✓
	Exerpta Medica	✗	✗	✗	✗	✗	✓
	Epilepsy Bibliography	✗	✗	✗	✗	✗	✓
	WHO Global Health Library	✗	✗	✗	✗	✓	✗
	African Index Medicus	✗	✗	✗	✗	✓	✗
	IMEMR	✗	✗	✗	✗	✓	✗
	IMSEAR	✗	✗	✗	✗	✓	✗
	WPRIM	✗	✗	✗	✗	✓	✗
	IBECS	✗	✗	✗	✗	✓	✗
	SciELO	✗	✗	✗	✗	✓	✗
	MedCarib	✗	✗	✗	✗	✓	✗
Study design	Cohort studies	✓	✓	✓	✓	✓	✓
	Case-control studies	✗	✓	✓	✗	✗	✗
Quality assessment tools used	Watila (2018)	Modified Newcastle–Ottawa scale and the ILAE standards for epidemiology					
	Levira (2016)	Ranked studies either 0/5, 10, 15, or 20 out of 20 across five quality domains					
	Thurman (2017)	Ranked studies 1, 2, 3, or 4 in increasing quality strength across the same five quality domains as Levira (2016)					
	Escalaya (2015)	Did not report a study quality assessment method					
	Nevalainen (2014)	Newcastle-Ottawa Scale for quality assessment and GRADE system evaluation of strength of evidence.					
	Shackleton (2002)	Did not report a study quality assessment method					

Abbreviations: LILACS - Latin American and Caribbean Health Sciences Literature; CINAHL - Cumulative Index to Nursing and Allied Health Literature; IMEMR - Index Medicus for the WHO Eastern Mediterranean Region; IMSEAR - Index Medicus for South-East Asia Region; WPRIM - Western Pacific Region Index Medicus; SciELO - Scientific Electronic Library Online; MedCarib - Caribbean Network of Health Sciences Libraries; ILAE - International League Against Epilepsy; GRADE - Grading of Recommendations Assessment, Development and Evaluation.

3.1.6. Mortality unrelated to epilepsy

Mortality likely to be unrelated to epilepsy was largely reported using SMR and PMR. For SMR (Fig. 5A), visual inspection of the data revealed two broad categories of risk. The highest category was risk of death from congenital anomalies or toxic effects (carbon monoxide and pesticides), with median SMR ranging between 9.2–12.1. The lower category was risk of death from 'other' diseases (not otherwise specified), neoplasia, digestive diseases, and cardiovascular disease, with median SMR ranging between 1.5–2.0.

PMR was high across all of the causes (other diseases, unknown causes, and sepsis), ranging between 15.2–33.5% (Fig. 5B).

3.1.7. Comparison of epilepsy-related and unrelated causes

Fig. 5C combines the results of studies showing epilepsy-related and unrelated causes using the primary outcome, SMR (23 studies). SMR results tended to be higher for epilepsy-related causes (median 3.8, range 0.0–82.4), than for unrelated causes (median 1.7, range 0.7–17.6), suggesting more PWE are likely to die prematurely from epilepsy-related causes of death. Epilepsy-related causes constituted 60% (12 out of 20) of the reasons PWE died.

3.1.8. Risk factors for all-cause mortality in people with epilepsy

Risk factors were reported using SMR, PMR, MR, RR, and HR. Inspection of SMR results reveals that the highest risk to a person with epilepsy appears to come from their comorbidities, type of epilepsy, treatment adherence, and their age at onset of epilepsy (Fig. 6). Equally important are a patient's age, and how long they have had epilepsy (Fig. 7A–D).

In comorbidities (Fig. 6), brain tumors, intellectual disabilities/cerebral palsy, and encephalopathies conferred particularly high risk, with an SMR ranging between 24.0–41.5. The types of epilepsy with a particularly high risk were congenital or developmental epilepsies (18.6-times increased median risk), and symptomatic epilepsies (3.3-times increased median risk). SMR was high for both poor and good treatment adherence groups, 8.0 and 7.4, respectively. These were self-reported outcomes, investigated by one study only. Whilst both childhood-onset and adult-onset epilepsies were associated with increased risk of death compared to the general population, this risk tended to be higher for the childhood-onset epilepsies (SMR range 5.7–46.4, median 7.3) than for adult-onset epilepsies (SMR range 1.3–10.7, median 6.0). The type of seizures, frequency of seizures, and gender appeared to be all around the same level of increased risk as each other. Within those groups, the peak SMR record came from generalized seizures (2.7-times increased median risk), increased seizure frequency (2.4-times increased median risk), and being male (2.3-times increased median risk).

In terms of age, children (< 18 years old) and young adults (≤ 50

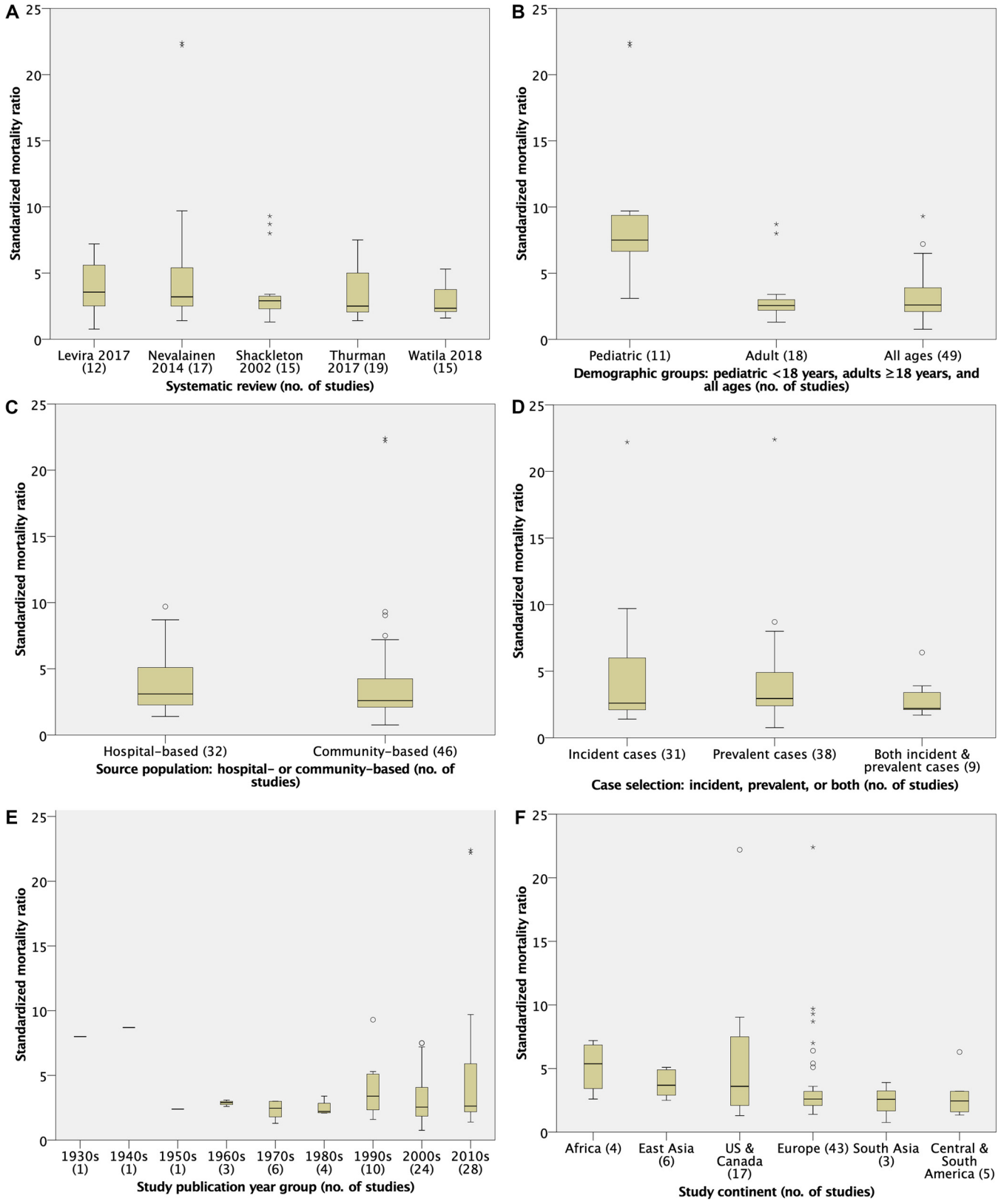


Fig. 2. Box plot panel showing overall mortality in people with epilepsy.
The box plots show the standardized mortality ratio (SMR) stratified by various groups. (A) shows SMR for the included systematic reviews in descending median order. Results from the five studies included in Escalaya et al. (2015) are not shown to avoid duplication with Levira et al. (2016), who reported these same studies. (B) shows SMR for studies in adults and children. (C) shows SMR for hospital and community studies. (D) shows SMR for incident and prevalent cases. (E) shows SMR over time by decade of study publication. (F) shows SMR by continent within which studies were conducted, in descending median order. The number of studies used to generate each box plot is shown in brackets on the x-axis. $^{*}/^{O}$ = outlier studies. Abbreviations: US = United States.

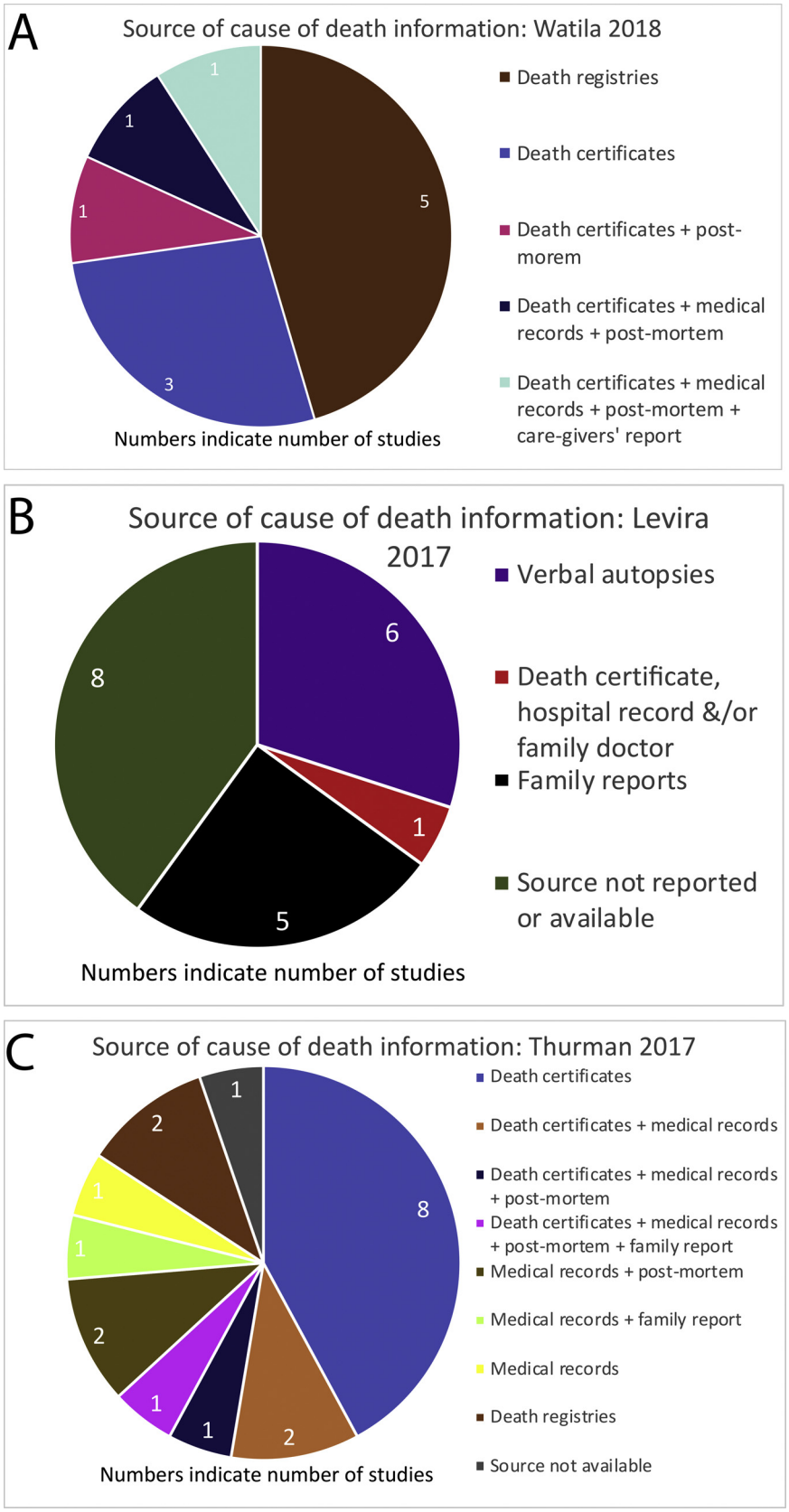


Fig. 3. Pie charts showing sources of cause of death information.
Three of the systematic reviews assessing cause of death as an outcome are shown. Results from the five studies included in Escalaya et al. (2015) are not shown to avoid duplication with Levira et al. (2016), who reported these same studies.

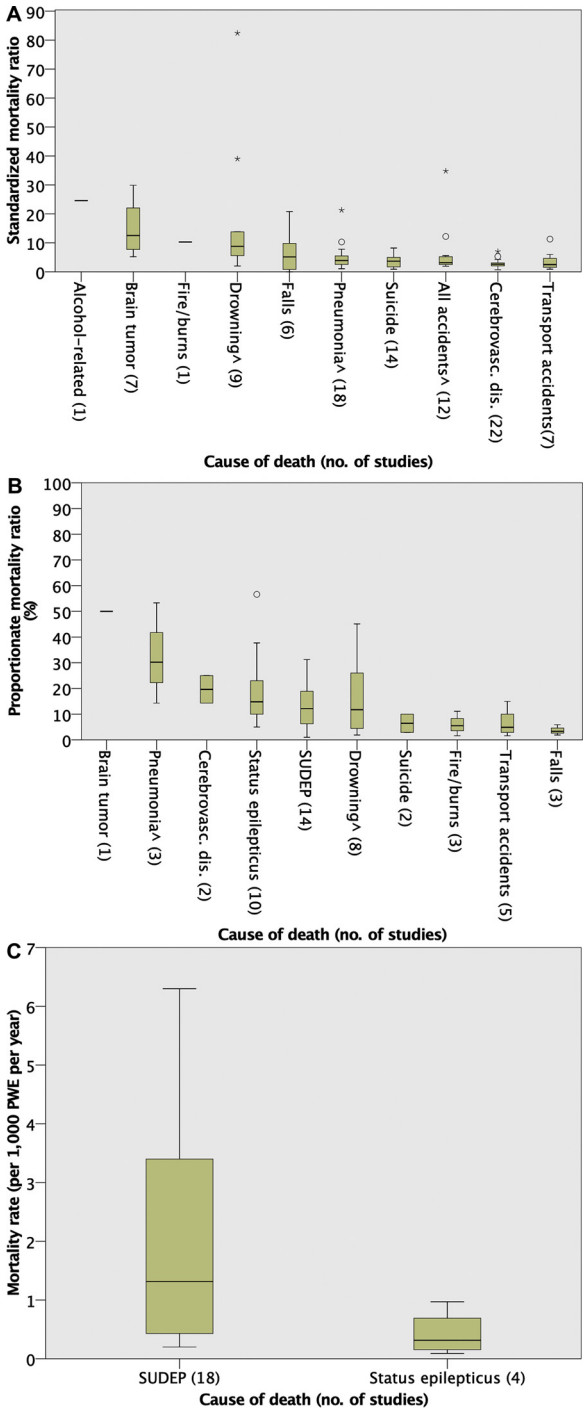


Fig. 4. Panel of box plots showing epilepsy-related causes of death.
(A) standardized mortality ratio (SMR), (B) proportionate mortality ratio (PMR), (C) mortality rate (MR) for the different causes, in descending median order. The number of studies used to generate each box plot is shown in brackets on the x-axis. ^Drowning also includes submersion and suffocation; Pneumonia also includes respiratory causes; All accidents also includes injuries. */O = outlier studies. Abbreviations: Cerebrovasc. dis. = cerebrovascular disease; PWE = people with epilepsy.

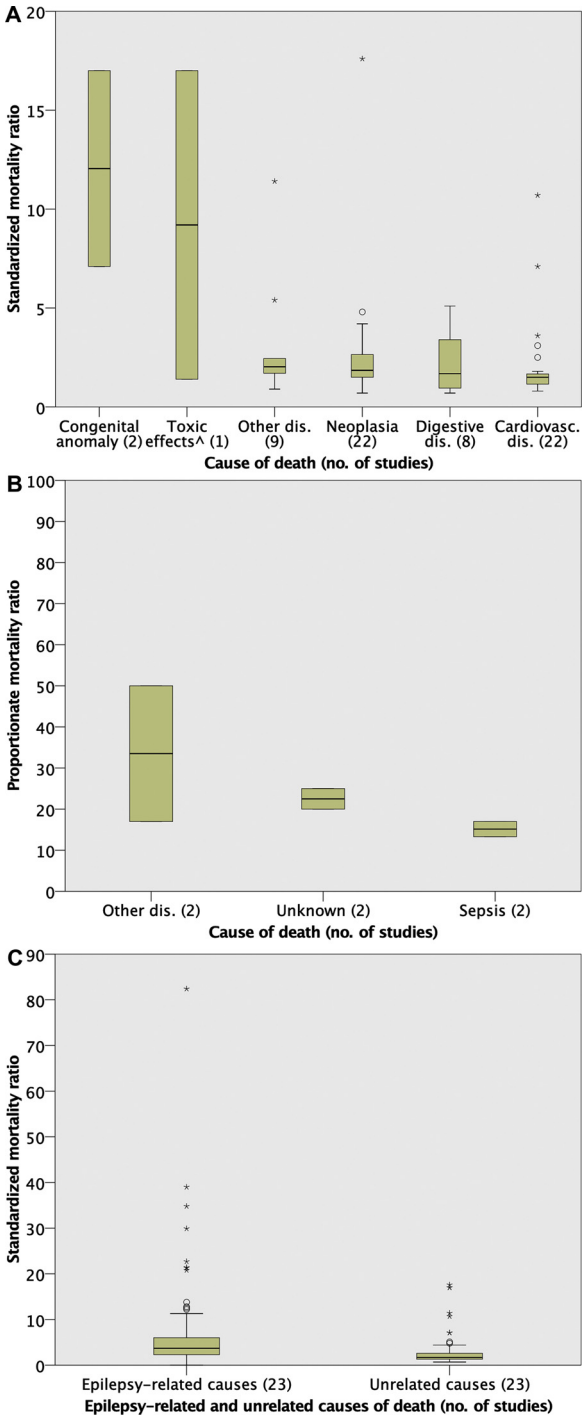


Fig. 5. Box plots panel showing mortality unrelated to epilepsy and comparison of epilepsy-related and unrelated causes.
(A) standardized mortality ratio (SMR), (B) proportionate mortality ratio (PMR), (C) epilepsy-related and unrelated causes, in descending median order. The number of studies used to generate each box plot is shown in brackets on the x-axis. ^Toxic effects includes poisoning. */O = outlier studies. Abbreviations: Other dis. = other diseases; Digestive dis. = digestive diseases; Cardiovasc. dis. = cardiovascular disease.

years) were at particularly high risk of death (Fig. 7A–C), with SMRs ranging between 3.5–14.4. SMRs decreased with age. Risk also appeared high in the first four years after diagnosis (8.6-times increased risk), and 10–14 years after diagnosis (23.8-times increased risk) (Fig. 7D).

PMRs, MRs, HRs and RRs showed a broadly similar picture to SMR in terms of risk (Fig. 7E–H), with the greatest risk signals associated with symptomatic epilepsy (Figs. 7E and 6 G), male gender (Fig. 7E), comorbidities (particularly encephalopathies, intellectual disability/cerebral palsy, and brain tumors, Fig. 7F), and the period soon after epilepsy diagnosis (Fig. 7H). PMR results suggested increased risk associated with focal seizures compared to generalized seizures (Fig. 7E),

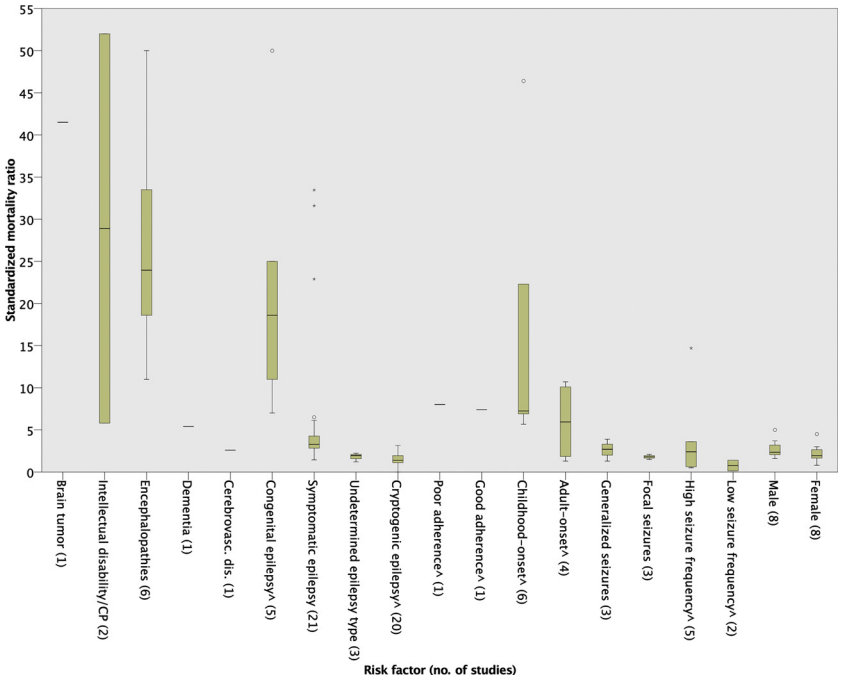


Fig. 6. Box plot showing mortality risk factors identified by standardized mortality ratio (SMR).
Risk factor categories are: **Comorbidities** = brain tumor, intellectual disability/cerebral palsy, encephalopathies, dementia, and cerebrovascular disease; **Epilepsy type** = congenital epilepsy, symptomatic epilepsy, undetermined epilepsy type, and cryptogenic epilepsy; **Compliance** = poor adherence and good adherence; **Onset age** = childhood-onset and adult-onset; **Seizure type** = generalized seizures and focal seizures; **Seizure frequency** = high seizure frequency and low seizure frequency; **Gender** = male and female. Plots are in descending median order within categories. The number of studies used to generate each box plot is shown in brackets on the x-axis. ^Congenital includes developmental epilepsy; Cryptogenic includes idiopathic epilepsy; Adherence was a self-reported outcome; Childhood-onset was < 18 years; Adult-onset was ≥ 18 years; High seizure frequency was > 1/week or refractory; Low seizure frequency was < 1/week or seizure-free. */^{O} = outlier studies. Abbreviations: CP = cerebral palsy; Cerebrovasc. dis. = cerebrovascular disease; Other dis. = other diseases; Digestive dis. = digestive diseases; Cardiovasc. dis. = cardiovascular disease.

in paradox to the findings of studies that used SMR to assess type of seizures as a risk factor (Fig. 6). Levira (2016) reported a study that investigated treatment adherence by monitoring medication possession, demonstrating increased risk of death in association with nonadherence to antiepileptic drugs (AEDs) compared to adherence (rate ratio 3.37, 95% CI 1.84–6.16) (Ngugi et al., 2014).

3.1.9. Risk factors for epilepsy-related mortality

From the four systematic reviews that investigated cause of death (Escalaya et al., 2015; Levira et al., 2016; Thurman et al., 2017; Watila et al., 2018), risk factors were reported for all-cause mortality. The reviews made no attempts to further subclassify risk factors according to whether or not they were in association with epilepsy-related versus unrelated causes of death.

3.2. Secondary analysis – systematic reviews of risk factors for individual epilepsy-related causes of death

The epilepsy-related causes of death for which there were systematic reviews of their risk factors were SUDEP, drowning, and suicide.

3.2.1. SUDEP

SUDEP was the epilepsy-related cause of death with the greatest number of systematic reviews of its risk factors. This was done across seven systematic reviews. Four of these assessed multiple risk factors for SUDEP. The earliest, Tellez-Zenteno et al. (2005), was of moderate quality (Appendix E) and included 36 studies (cohort and case-control studies) of SUDEP at all ages. They assessed for the positive presence of a risk factor amongst SUDEP cases within included studies. The most consistent risk factors were seizures preceding death (with most patients found dead in bed), and subtherapeutic AED levels (80% of studies). Other positive risk factors were youth (defined as age 15–30 years) in 50% of studies, high seizure frequency (> 15 per month) in 60%, a high number of AEDs (> 2) in 50%, and a long duration of epilepsy (> 15 years) in 50% of studies. Monte et al. (2007) was also of moderate quality (Appendix E). The authors recognized conflicting data on SUDEP risk factors following Tellez-Zenteno (2005) and sought to address this using risk factor total studies ratios (RFT), which were values based on the results in studies analyzing the risk factor corrected for the total number of included SUDEP cases in the different studies. RFT scores of ≥1.0, ≥0.5, and < 0.5 were considered to indicate a strong, weak, and absent risk factor for SUDEP, respectively. The strong risk factors were young age (< 45 years), early onset of seizures (< 45 years at diagnosis), the presence of generalized tonic-clonic seizures (GTCS), male sex and being in bed. Harden et al. (2017) then developed a clinical practice guideline summary for SUDEP. In this moderate-quality systematic review (Appendix E), they looked at SUDEP in all ages from 70 studies (clinical trials and observational studies) and identified GTCS as the greatest risk factor (OR 10, 95% CI 14–17). Others were increased frequency of GTCS: OR 15.46 (95% CI 9.92–24.10) for > 3 GTCS per year and OR 5.07 (95% CI 95% 2.94–8.76) for 1–2 GTCS per year, not adding in another AED to refractory patients (OR 6, 95% CI 2–20), and not being seizure-free for 1–5 years (OR 4.7, 95% CI 1.4–16). Protective factors were nocturnal supervision (OR 0.4, 95% CI 0.2–0.8), and use of a nocturnal listening device (OR 0.1, 95% CI 0–0.3). Abdel-Mannan et al. (2018) then recognized that there had not been a focus on children in SUDEP systematic reviews. In this high-quality systematic review (Appendix E), they identified 108 pediatric cases (age ≤18 years) of SUDEP from 22 studies including cohorts, case-controls, case series, and case reports. Increased SUDEP was associated with GTCS (affecting 88% of participants), chronic brain lesions (affecting 61%), male sex (51% males), and cognitive or developmental delay (affecting 45%).

Other SUDEP systematic reviews focused on an individual risk factor for SUDEP. Young et al. (2015), a low-quality systematic review (Appendix E), looked at intellectual disability as a risk factor. From 23 studies assessing this risk factor, 60% (14 studies) found it was positively associated with SUDEP. In a moderate-quality systematic review (Appendix E), Liebenthal et al. (2015) investigated the association between a prone body position and SUDEP in all ages from 253 cases in 25 studies (case reports and case series). 73.3% (95% CI 65.7–80.9%) died in a prone position, compared to 26.7% (95% CI 16.3–37.1%) dying in a nonprone position ($p < 0.001$). A prone position was found in 85.7% (95% CI 74.6–93.3%) of those aged ≤40 years compared to 60% (95% CI 5 38.7%, 78.9%) of patients > 40 years (OR 3.9, 95% CI 1.4–11.4%). In their moderate-quality systematic review (Appendix E), Ali et al. (2017) investigated the associated of sleep with SUDEP, identifying 880 cases from 67 studies (a mixture of case reports and case series in participants of all ages). SUDEP was significantly associated with sleep

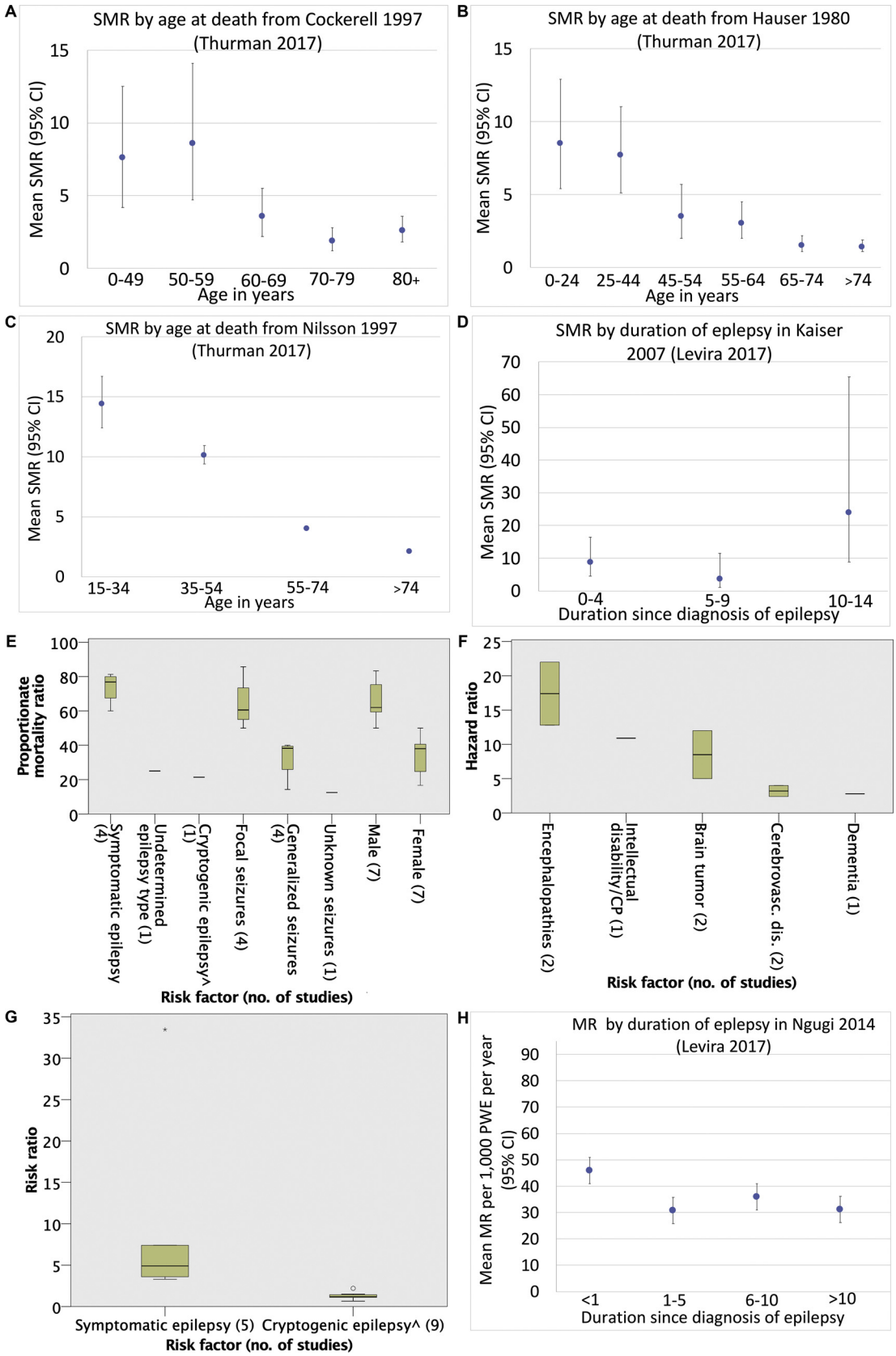


Fig. 7. Scatter and box plots showing additional mortality risk factors.
(A–C) mean standardized mortality ratios (SMRs) according to age, as reported by three studies from Thurman et al. (2017); (D) SMR according to length of time since epilepsy diagnosis; (E) risk factors as reported using proportionate mortality ratios (PMR) categorized by epilepsy-type (symptomatic, undetermined, cryptogenic), seizure-type (focal, generalized, unknown), and gender. Plots are in descending median order within categories; (F) hazard ratios (HR) in descending median order; (G) risk ratio; (H) mortality rate. The number of studies used to generate each box plot is shown in brackets on the x-axis. ^Cryptogenic includes idiopathic epilepsy. */^{O} = outlier studies. Abbreviations: SMR = standardized mortality ratio; CP = cerebral palsy; MR = mortality rate; PWE = people with epilepsy; CI = confidence interval.

compared to wakefulness ($p < 0.001$), and patients ≤40 years were more likely to have SUDEP in sleep than those older (OR 2.0, 95% CI 1.0–3.8). 87.6% (95% CI 81.1–94.2%) of those experiencing SUDEP in sleep were in a prone position compared to 52.9% (95% CI 4.7–81.1%) during wakefulness. Patients with night-time seizures were 6.3 times more likely to have SUDEP when in a prone position than those with diurnal seizures (OR 6.3, 95% CI 2.0–19.5).

3.2.2. Drowning

In their moderate-quality systematic review, Bell et al. (2008) assessed risk factors for drowning in PWE of all ages, identifying 88 drowning deaths from 51 cohorts. Drowning was more closely associated with an established diagnosis of epilepsy (prevalent epilepsy SMR 18.0, 95% CI 12.9–24.4) than newly diagnosed epilepsy (incident epilepsy SMR 5.4, 95% CI 1.5–13.9). Other positive risk factors were learning disability (SMR 25.7, 95% CI 16.6–38.0), institutional care (SMR 96.9, 95% CI 26.4–248), and temporal lobe excision (SMR 41.1, 95% CI 11.2–105). Risks of drowning were highest in Africa.

3.2.3. Suicide

In another moderate-quality systematic review, Bell et al. (2009) assessed risk factors for suicide in PWE of all ages, identifying 190 suicides from 76 cohorts. Suicide was most closely associated with temporal lobe excision (SMR 13.9, 95% CI 8.93–20.74), temporal lobe epilepsy (SMR 6.57, 95% CI 1.79–16.8), epilepsy surgery (SMR 6.37, 95% CI 3.06–11.72), and an established diagnosis of epilepsy (prevalent epilepsy SMR 4.81, 95% CI 3.08–7.16). Developmental disability was associated with lower suicide risks (SMR 0.34, 95% CI 0.07–0.99).

4. Discussion

4.1. Areas of novelty

There are three main areas of novelty within the current systematic review. First, systematically reviewing the systematic reviews of mortality in PWE has allowed us to draw conclusions using a primary analysis from a much larger sample of evidence than any previous systematic review of this topic, since we have captured 103 studies (40 more studies than was captured by the next largest review (Watila et al., 2018)). This is important because as a method, systematic review should aim to evaluate an entire body of evidence surrounding a topic (Higgins and Green, 2011). Inspection of Table 2 reveals that some of the likely reasons for lack of overlap in included studies from the systematic reviews are variation in primary outcomes (Nevalainen, 2014 and Shackleton, 2002, for example, did not assess causes of death), variation in the number of electronic databases searched (Watila, 2018; Levira, 2016; Thurman, 2017 and Escalaya, 2015 searched 2–3 whilst Nevalainen, 2014 searched 16), and variation in included study designs (Levira, 2016 and Thurman, 2017 included both cohort and case-control studies, whilst the remainder included only cohort studies). The second area of novelty is that we have applied a recently published classification system (Devinsky et al., 2016) to allow us to quantify epilepsy-related and unrelated deaths. This has not been done before, and it has allowed us to include within the epilepsy-related dataset, causes not originally classed as epilepsy-related within the included systematic reviews, e.g. alcohol, pneumonia, and suicide. Third, our secondary analysis gives an overview of the risk factors for epilepsy-related mortality, identifying that these have failed to be captured by systematic reviews of all-cause mortality in which epilepsy-related causes of death were included. Finally, we have provided the first independent quality assessment of the epilepsy mortality systematic reviews.

Our most novel findings beyond those identified in the original reviews are that:

1) Despite presumed advances in epilepsy diagnostics and therapeutics, overall mortality in PWE has remained high and virtually unchanged in the decades since 1950. The reasons for this are unclear and this requires further investigation. One possible explanation is increased reporting of epilepsy on death certificates over the last 10–20 years due to improved physician education occurring at the same time as improvements in epilepsy care. The former would act to increase reported rates whilst the latter would act to reduce actual rates, resulting in a net lack of change in final mortality rate figures (Neligan et al., 2010);
2) Overall mortality is highest in Africa and there are also no studies in Australasia, the Middle East, South East Asia, and Russia. Some of the possible reasons for higher mortality in Africa include that the condition remains stigmatized widely there (Belhocine et al., 2004), there are limited health personnel and facilities for diagnosis and treatment, and there are major treatment gaps, with as high as 80–85% of PWE not diagnosed or treated (Donner et al., 2001);
3) Epilepsy-related mortality predominates causes of death, with an SMR of more than twice that of unrelated causes and accounting for 60% of the reasons PWE died;
4) Whilst SUDEP contributes to a high proportion of epilepsy-related causes of death, other epilepsy-related causes are likely to be similarly lethal including alcohol, drowning, pneumonia, and suicide (these may be preventable (Devinsky et al., 2016; Hanna et al., 2002; Neligan and Sander, 2011));
5) Risk factors for epilepsy-related deaths specifically, as a group, remain uncharacterized. However, risk factors for individual epilepsy-related deaths have been looked at largely for SUDEP, but also for drowning and suicide. For SUDEP, youth (age under 30–45 years) and the presence of GTCS appear to be the risk factors most commonly reported as significant across systematic reviews of multiple SUDEP risk factors, and reviews of individual SUDEP risk factors add further information that not only is SUDEP risk greatest in the prone position, but that this is more so in youth. Between them, the systematic reviews also identify other pertinent SUDEP risk factors as nocturnal seizures or sleep (particularly in the prone position), increased seizure frequency (particularly generalized seizures) or lack of seizure freedom, intellectual disability/developmental delay, undertreatment (subtherapeutic AED levels or incomplete treatment of refractory epilepsy), and male sex. Nocturnal supervision and listening devices are suggested to be protective, although this is only mentioned in one systematic review. SUDEP risk factors are similar for children and adults although there are additional reports of chronic brain lesions and developmental delay as significant risk factors in children. Drowning appears to be most closely associated with chronic epilepsy, learning difficulty, temporal lobe surgery, and institutional care. The latter is surprising, given that institutional care is assumed to involve greater supervision for patients. Although Africa conferred the highest risks of drowning for PWE, the World Health Organization reports that drowning rates in the general population worldwide are highest in Africa anyway (World Health Oraganization, 2018). Suicide appears to also be associated with temporal lobe epilepsy pathology/intervention and chronic epilepsy, whilst developmental delay is protective for this. The reasons for a link to temporal lobe pathology remain unclear (Bell et al., 2009). Similarly, it is unlikely that lower suicide risk is due to greater supervision in those with developmental delay because risk of drowning remains high in institutional care, where many of such patients are.

4.2. Implications and areas for future research

The primary analysis reveals that the overall risks of mortality were consistently high across systematic reviews, compared to the general population, in line with the two-to-threefold increase frequently quoted in the literature (Neligan and Bell, 2009). This result was unaffected by systematic review quality. The consistent finding that children and

young adults had the highest risk of death may be explained by a high incidence of congenital or symptomatic epilepsies in children (Levira et al., 2016), which we found to be pertinent risk factors for premature death. It may also be explained by comparatively lower mortality rates in the general population at a younger age (Levira et al., 2016). This highlights one of the limitations of comparative measures, such as SMR. They can mask absolute increases in the rate of mortality of a study population if equal to the general population's (Thurman et al., 2017), even if the causes of death in the study and general population differ. Thus, we propose that more work needs to be done to ascertain and report on absolute mortality rates, rather than studies and systematic reviews only reporting comparative measures. Absolute MR was reported only in one systematic review, which focused on data from low- and middle-income countries alone (Levira et al., 2016). The absolute MR burden for PWE in high-income countries remains to be clarified.

Overall mortality did not appear different between hospital- and community-based studies. This was likely to be because by excluding studies of highly selected subpopulations (such as those consisting of only surgical patients or only patients with intellectual disability), the hospital sample became more equally representative of the community sample (Nevalainen et al., 2014). This helps highlight how epilepsy still remains a major cause of death even in the community. Therefore, it will remain important for future studies investigating mortality in PWE to include both hospital and community populations.

This is the first systematic review to make an attempt to quantify epilepsy-related mortality using a formal classification system (Devinsky et al., 2016). Whilst researchers can never be absolutely certain that a death was epilepsy-related, even in the event of a post-mortem (Devinsky et al., 2016), the classification system helps us to functionally categorize deaths into those that are most likely to be epilepsy-related. Given the variation in defining epilepsy-related mortality in literature thus far (Devinsky et al., 2016), our use of the classification system in this manner, if developed and continued, should help standardize case ascertainment for epilepsy-related deaths within future primary studies and systematic reviews. One of the rationales for a need for greater consistency and quantification of epilepsy-related mortality, specifically, is that it is the epilepsy-related proportion of deaths (rather than the unrelated deaths) that are felt to most likely be preventable (Devinsky et al., 2016; Hanna et al., 2002; Neligan and Sander, 2011).

The main epilepsy-related causes of death we identified were in relation to alcohol, brain tumors, SUDEP, accidents (fires, drowning, falls), pneumonia, and suicide. There is more literature on SUDEP than any other epilepsy-related cause of death (Devinsky et al., 2016). Whist increased attention on SUDEP is warranted, our report suggests that equal attention is needed toward other epilepsy-related causes, which may be similarly dangerous. For example, we found that whilst SUDEP contributed to 12% of deaths on PMR, drowning also contributed to 12%, and pneumonia contributed to 30%. It may be that national patient and family education programs may help prevent the risks of drowning. The increased risk of pneumonia in the epilepsy population is likely to be related to gastric aspiration during seizures (Devinsky et al., 2016). This may require better screening and early treatment for those at risk. The increased risk of suicide in also worrying and occurred in multiple older and newer studies worldwide (Chamorro-Munoz et al., 2017; Chang et al., 2012; Day et al., 2005; Ding et al., 2006, 2013; Granbichler et al., 2015; Mohanraj et al., 2006; Mu et al., 2011; Rafnsson et al., 2001; Shackleton et al., 1999; White et al., 1979). It may be that more needs to be done to manage mental health problems in PWE. We know that mood disorders represent a frequent psychiatric comorbidity in PWE, but they often remain unrecognized and untreated (Mula and Schmitz, 2009). The potential association of alcohol and increased mortality requires further investigation given that this cause had the highest SMR of all epilepsy-related causes but was reported only in one study. We included alcohol in the epilepsy-related cause of death group because beyond withdrawal seizures, there is a strong and consistent association between alcohol consumption and unprovoked seizures (i.e. epilepsy), which demonstrates a dose-response relationship between the amount of alcohol consumed daily and the probability of onset of epilepsy (Samokhvalov et al., 2010).

The risk factors most strongly associated with increased all-cause mortality were comorbid diseases, particularly those affecting the brain (tumors, intellectual disability or cerebral palsy, and encephalopathies), congenital or symptomatic epilepsies, poor adherence to treatment (although there was a limited number of studies to confirm this), young age or childhood-onset of epilepsy, the period soon after epilepsy diagnosis, and male sex. There were similarities between these and risk factors for SUDEP reported in the secondary analysis, suggesting a degree of overlap between risk factors for SUDEP and risk factors for all-cause mortality in PWE. This indicates that perhaps risk prevention strategies for SUDEP might also reduce all-cause mortality. Further work is still needed to investigate risk factors for other epilepsy-related causes beyond SUDEP, which has received the greatest literary attention, given that the other epilepsy-related deaths are also common. Only downing and suicide have been reviewed in terms of risk factor profiles. Evaluating the risk factors for epilepsy-related death together as a group would allow for the development of a risk-index scoring tool to identify PWE at high risk of epilepsy-related death (Devinsky et al., 2016) in a similar fashion to how the SUDEP and Seizure Safety Checklist identifies PWE at high risk of SUDEP (SUDEP Action, 2018).

4.3. Limitations

There are several limitations to the current systematic review. Firstly, although the Devinsky classification system is set up to identify epilepsy-related deaths in PWE, it is not a perfect system and will likely need streamlining in future. For example, it may be argued that deaths due to acute symptomatic seizures should be removed as a category, as some of these patients may not go on to develop epilepsy. We did not remove this category as ours is the first review to use the original classification system, therefore it paves the way for literary debate on which categories should now be amended/removed going forward. Related to this, some proposed epilepsy-related causes may not reflect the possibility of an epilepsy-independent mechanism of death. For example, PWE may die naturally from their brain tumor, stroke, and even from a drowning, traffic accident, or fall unrelated to their epilepsy. Therefore, the classification system should remain viewed as one in which the weight of probability is that these are likely to be epilepsy-related deaths in PWE but that this is not always certain. However, the clinical importance of such a classification system remains clear. In the instances where these deaths are indeed epilepsy-related (and many are), then these should have been preventable (Devinsky et al., 2016; Hanna et al., 2002; Neligan and Sander, 2011). Therefore, the classification system remains a useful tool to help clinicians and researchers start to capture and highlight information on other potential epilepsy-related deaths beyond SUDEP. A second limitation is that we were unable to address the problem of confounding. There may have been several significant confounders between systematic reviews, particularly between the low- or middle-income countries and the higher-income countries. Confounding is usually managed using multivariate meta-regression, but this was not possible in a narrative synthesis of findings. There would have been difficulty in dealing with confounding at source, as three of the included primary analysis systematic reviews were narrative and did not use formal techniques to handle confounding (Escalaya et al., 2015; Levira et al., 2016; Thurman et al., 2017). This highlights the difficulties other authors have identified in trying to summarize the epilepsy mortality data uniformly (Logroscino and Hesdorffer, 2005). Related to this, the reviews highlighted significant heterogeneity between studies, precluding valid use of meta-analysis to provide summary estimates (Keezer et al., 2015). Differences in source population and case selection are the most likely explanations for the heterogeneity (Shackleton et al., 2002). There are limitations in

the utility of SMR, which assesses mortality within a study population compared to a general population stratified by age and sex. Different studies will use different age bands to stratify, and there may be differing baseline mortality rates between the different general populations used by studies. This means that any comparison of SMRs between studies should be made and interpreted with caution. This is another reason often cited for not performing meta-analysis across studies that have used SMR (Logroscino and Hesdorffer, 2005). Epilepsy mortality data is limited in this way for all systematic reviews of this subject, as most studies have reported SMRs alone. We were also unable to assess the risk of publication bias. Finally, we have presented causes of death as though they were definitive. However, we note that ascertaining cause of death accurately can be challenging (Devinsky et al., 2016), especially when it comes to identifying SUDEP, where post-mortem is required. None of the studies in low- and middle-income countries used post-mortem, relying instead on verbal autopsies, which can be inaccurate (Levira et al., 2016). Furthermore, only seven of the studies in high-income countries used post-mortem. Once again, epilepsy mortality data is limited in this way for all systematic reviews of this subject, as the heterogeneity in how epilepsy mortality studies have been conducted worldwide remains. Even within the systematic review in which analysis was limited to studies from low- and middle-income countries alone (Levira et al., 2016) and in the review in which analysis was limited to those from high-income countries (Thurman et al., 2017), heterogeneity remained present. Future studies should aim to combine information from several sources, such as death certificates, information from medical records, interview with witnesses and responsible physicians, and post-mortem results. Furthermore, consensus should be reached on how such studies should be conducted in future in order to help reduce heterogeneity.

5. Conclusions

Despite the limitations, several implications of this review are clear. There are high-quality systematic reviews of mortality in PWE, covering the majority of the globe. However, when looked at alone, each review has been unable to identify all of the available studies. When the results of the systematic reviews are combined, they indicate that despite advances in diagnostics and therapeutics over time, mortality has remained high for PWE compared to the general population, particularly in children and young adults. Future studies should aim to ascertain not only these comparative mortality outcomes, but also the absolute mortality rates, particularly in high-income countries. Our review also indicates that there is an increased burden of epilepsy-related deaths. These present in several potentially preventable manners, so future studies and reviews of mortality in PWE would benefit from stratifying results into epilepsy-related and unrelated deaths. Risk factors for all-cause mortality and SUDEP are similar, and further work is required to investigate the risk factor profile of other epilepsy-related causes of death.

Funding

This work was supported charitably by Epilepsy Research UK (R44007), the Juliet Bergqvist Memorial Fund, and Epilepsy Action.

Ethics approval, informed consent, trial registration

Not required.

Declaration of Competing Interest

None.

Acknowledgement

We also thank Jacqueline Stephen from The University of Edinburgh for providing statistical support.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.eplepsyres.2019.106192.

References

Abdel-Mannan, O., et al., 2018. A systematic review of sudden unexpected death in epilepsy (SUDEP) in childhood. Epilepsy Behav. 90, 99–106.
Ackers, R., et al., 2011. Mortality rates and causes of death in children with epilepsy prescribed antiepileptic drugs: a retrospective cohort study using the UK General Practice Research Database. Drug Saf. 34 (5), 403–413.
Ali, A., et al., 2017. Association of sleep with sudden unexpected death in epilepsy. Epilepsy Behav. 76, 1–6.
Belhocine, M., de Boer, H.M., Mandlhate, C., 2004. Epilepsy in the WHO African region: bridging the gap. Out of the Shadows. (Accessed 04/07/2019). https://www.who.int/mental_health/management/epilepsy_in_African-region.pdf.
Bell, G.S., et al., 2008. Drowning in people with epilepsy - how great is the risk? Neurology 71 (8), 578–582.
Bell, G.S., et al., 2009. Suicide in people with epilepsy: how great is the risk? Epilepsia 50 (8), 1933–1942.
Chamorro-Munoz, M.I., et al., 2017. Epidemiological study of mortality in epilepsy in a Spanish population. Seizure 46, 19–23.
Chang, Y.H., et al., 2012. Risk of mortality among patients with epilepsy in southern Taiwan. Seizure-Eur. J. Epilepsy 21 (4), 254–259.
Christensen, J., et al., 2015. Long-term mortality in children and young adults with epilepsy–A population-based cohort study. Epilepsy Res. 114, 81–88.
Clarke, M., 2008. Systematic review of reviews of risk factors for intracranial aneurysms. Neuroradiology 50 (8), 653–664.
Corry, M., et al., 2015. A systematic review of systematic reviews on interventions for caregivers of people with chronic conditions. J. Adv. Nurs. 71 (4), 718–734.
Day, S.M., et al., 2005. Causes of death in remote symptomatic epilepsy. Neurology 65 (2), 216–222.
Devinsky, O., et al., 2016. Recognizing and preventing epilepsy-related mortality: a call for action. Neurology 86 (8), 779–786.
Ding, D., et al., 2006. Premature mortality in people with epilepsy in rural China: a prospective study. Lancet Neurol. 5 (10), 823–827.
Ding, D., et al., 2013. Premature mortality risk in people with convulsive epilepsy: long follow-up of a cohort in rural China. Epilepsia 54 (3), 512–517.
Donner, E.J., Smith, C.R., Snead 3rd, O.C., 2001. Sudden unexplained death in children with epilepsy. Neurology 57 (3), 430–434.
Escalaya, A.L., et al., 2015. Epilepsy and mortality in Latin America. Seizure 25, 99–103.
Fazel, S., et al., 2013. Premature mortality in epilepsy and the role of psychiatric comorbidity: a total population study. Lancet 382 (9905), 1646–1654.
Granbichler, C.A., et al., 2015. Cause-specific mortality in adult epilepsy patients from Tyrol, Austria: hospital-based study. J. Neurol. 262 (1), 126–133.
Hanna, N.J., et al., 2002. The National Sentinel Clinical Audit of Epilepsy-related Death: Epilepsy – Death in the Shadows. The Stationery Office.
Harden, C., et al., 2017. Practice guideline summary: Sudden unexpected death in epilepsy incidence rates and risk factors: Report of the Guideline Development, Dissemination, and Implementation Subcommittee of the American Academy of Neurology and the American Epilepsy Society. Neurology 88 (17), 1674–1680.
Higgins, J.P.T., Green, S., 2011. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0.
Keezer, M.R., Bell, G.S., Sander, J.W., 2015. Epilepsy-related clinical characteristics and mortality: a systematic review and meta-analysis. Neurology 84 (17), 1823.
Levira, F., et al., 2016. Premature mortality of epilepsy in low- and middle-income countries: a systematic review from the Mortality Task Force of the International League Against Epilepsy. Epilepsia 1–11.
Liebenthal, J.A., et al., 2015. Association of prone position with sudden unexpected death in epilepsy. Neurology 84 (7), 703–709.
Logroscino, G., Hesdorffer, D.C., 2005. Methodologic issues in studies of mortality following epilepsy: measures, types of studies, sources of cases, cohort effects, and competing risks. Epilepsia 46 (Suppl. 11), 3–7.
Lunny, C., et al., 2017. Toward a comprehensive evidence map of overview of systematic review methods: paper 1-purpose, eligibility, search and data extraction. Syst. Rev. 6 (1), 231.
Mbizvo, G.K., et al., 2012. Levetiracetam add-on for drug-resistant focal epilepsy: an updated Cochrane Review. Cochrane Database Syst. Rev. 9, CD001901.
Mohanraj, R., et al., 2006. Mortality in adults with newly diagnosed and chronic epilepsy: a retrospective comparative study. Lancet Neurol. 5 (6), 481–487.
Moher, D., et al., 2009. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. BMJ 339, b2535.
Monte, C.P., et al., 2007. Sudden unexpected death in epilepsy patients: Risk factors. A systematic review. Seizure 16 (1), 1–7.
Moseley, B.D., et al., 2013. Early onset epilepsy is associated with increased mortality: a

population-based study. Epilepsy Res. 105 (3), 410–414.
Mu, J., et al., 2011. Causes of death among people with convulsive epilepsy in rural West China: a prospective study. Neurology 77 (2), 132–137.
Mula, M., Schmitz, B., 2009. Depression in epilepsy: mechanisms and therapeutic approach. Ther. Adv. Neurol. Disord. 2 (5), 337–344.
Neligan, A., Bell, G.S., 2009. The mortality of epilepsy. In: Sander, J.W., Rugg-Gunn, F.J., Smalls, J.E. (Eds.), Epilepsy 2009: From Benchside to Bedside - a Practical Guide to Epilepsy, 12th edn. Chalfont St Peter: National Society for Epilepsy, pp. 347–351.
Neligan, A., Sander, J.W., 2011. Premature mortality in epilepsy: is it preventable? Expert Rev. Neurother. 11 (6), 767–770.
Neligan, A., et al., 2010. Temporal trends in the mortality of people with epilepsy: a review. Epilepsia 51 (11), 2241–2246.
Nevalainen, O., et al., 2012. Mortality by clinical characteristics in a tertiary care cohort of adult patients with chronic epilepsy. Epilepsia 53 (12), e212–4.
Nevalainen, O., et al., 2013. Long-term mortality risk by cause of death in newly diagnosed patients with epilepsy in Finland: a nationwide register-based study. Eur. J. Epidemiol. 28 (12), 981–990.
Nevalainen, O., et al., 2014. Epilepsy-related clinical characteristics and mortality: a systematic review and meta-analysis. Neurology 83 (21), 1968–1977.
Ngugi, A.K., et al., 2014. Premature mortality in active convulsive epilepsy in rural Kenya: causes and associated factors. Neurology 82 (7), 582–589.
Rafnsson, V., et al., 2001. Cause-specific mortality in adults with unprovoked seizures. A population-based incidence cohort study. Neuroepidemiology 20 (4), 232–236.
Ryan, R., et al., 2014. Interventions to improve safe and effective medicines use by consumers: an overview of systematic reviews. Cochrane Database Syst. Rev. 4.
Samokhvalov, A.V., et al., 2010. Alcohol consumption, unprovoked seizures, and epilepsy: a systematic review and meta-analysis. Epilepsia 51 (7), 1177–1184.
Shackleton, D.P., et al., 1999. Mortality in patients with epilepsy: 40 years of follow up in a Dutch cohort study. J Neurol Neurosurg Psychiatry 66 (5), 636–640.
Shackleton, D.P., et al., 2002. Survival of patients with epilepsy: an estimate of the mortality risk. Epilepsia 43 (4), 445–450.
Shea, B.J., et al., 2009. AMSTAR is a reliable and valid measurement tool to assess the methodological quality of systematic reviews. J. Clin. Epidemiol. 62 (10), 1013–1020.
Shea, B.J., et al., 2007a. External validation of a measurement tool to assess systematic reviews (AMSTAR). PLoS One 2 (12), e1350.
Shea, B.J., et al., 2007b. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. BMC Med. Res. Methodol. 7, 10.
Smith, V., et al., 2011. Methodology in conducting a systematic review of systematic reviews of healthcare interventions. BMC Med. Res. Methodol. 11 (1), 15.
SUDEP Action, 2018. SUDEP and Seizure Safety Checklist. (Accessed 07 November 2018). https://sudep.org/checklist.
Tellez-Zenteno, J.F., Ronquillo, L.H., Wiebe, S., 2005. Sudden unexpected death in epilepsy: evidence-based analysis of incidence and risk factors. Epilepsy Res. 65 (1–2), 101–115.
Thurman, D.J., et al., 2017. The burden of premature mortality of epilepsy in high-income countries: a systematic review from the Mortality Task Force of the International League Against Epilepsy. Epilepsia 58 (1), 17–26.
Watila, M.M., et al., 2018. Overall and cause-specific premature mortality in epilepsy: a systematic review. Epilepsy Behav.
White, S.J., McLean, A.E., Howland, C., 1979. Anticonvulsant drugs and cancer. A cohort study in patients with severe epilepsy. Lancet 2 (8140), 458–461.
World Health Oraganization, 2018. Drowning. Fact Sheets. (Accessed 03/07/2019). https://www.who.int/news-room/fact-sheets/detail/drowning.
Young, C., et al., 2015. Does intellectual disability increase sudden unexpected death in epilepsy (SUDEP) risk? Seizure-Eur. J. Epilepsy 25, 112–116.

2.3 Conclusion

In this chapter, our aim was to estimate the global burden of all-cause as well as epilepsy-related mortality in PWE. We learn that all-cause mortality in PWE is increased compared to the general population in nearly all countries tested, and has remained similarly high over time. We also show that epilepsy-related deaths probably contribute a substantial proportion of the deaths in PWE. Children and young adults with epilepsy appear to be at the highest risk of death. Whilst SUDEP is a common epilepsy-related cause of death, it appears to be one of many other epilepsy-related causes affecting PWE, some of which may be as common as SUDEP. This is something that will be investigated further in a later chapter of the thesis, using MCOD analysis.

Our systematic review further indicates that there has been little work attempting to identify risk factors for epilepsy-related deaths as a group, with most risk factors thus far identified for all-cause mortality in PWE or individual causes such as SUDEP. This will be explored later on in the thesis, in a chapter where we develop a risk stratification tool for epilepsy-related deaths. Such a tool would be useful as a substantial proportion of epilepsy-related deaths are thought to be avoidable,[15,39,40] which is something we will also try to quantify later in this thesis.

We recently presented the systematic review reported here at the Association of British Neurologists' (ABN) Annual Meeting 2020. We received helpful feedback that it would be useful to also show how the epilepsy-related causes of death identified were spread across different countries in order to gain

insight into whether there is a tendency for some causes to occur in certain countries more. This is now shown in Figures 2-1 and 2-2.

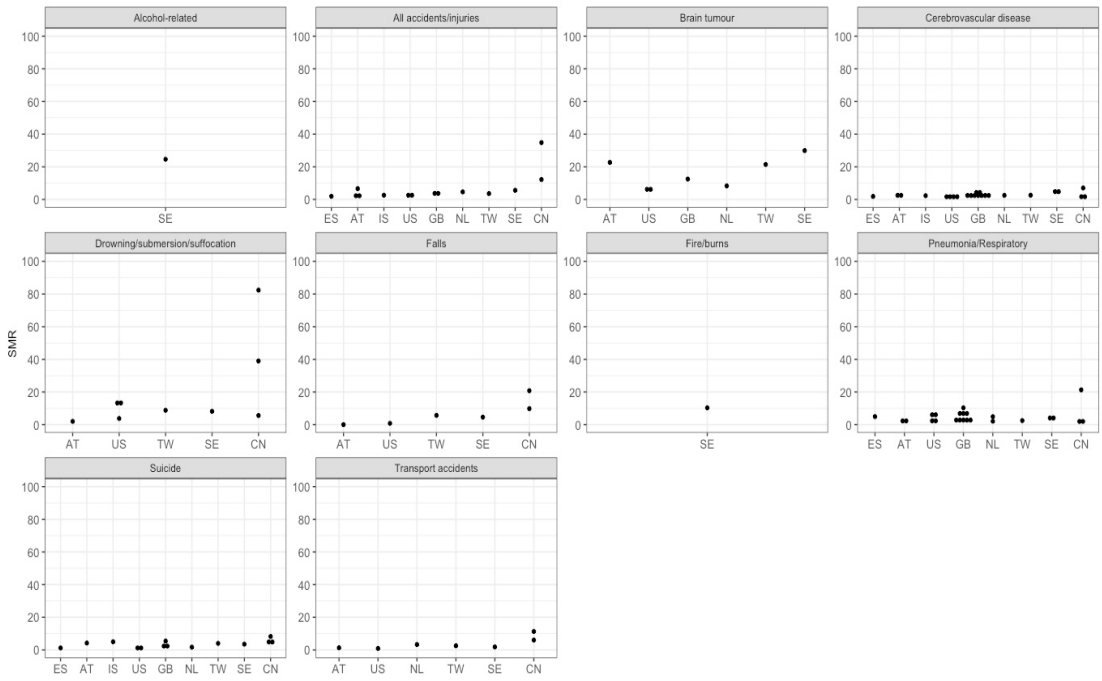


Figure 2-1 Dot plot panel of epilepsy-related causes of death by SMR and country

Abbreviations: SE – Sweden; ES – Spain; AT – Austria; IS – Iceland; US – United States of America; GB – Great Britain; NL – The Netherlands; TW – Taiwan; CN – China; SMR – Standardised mortality ratio

For any causes of death reported in China (accidents and injuries, cerebrovascular disease, drowning, falls, pneumonia, suicide, transport accidents), the peak SMRs for that cause were higher in China than anywhere else in the world reporting that particular cause. For example, the SMR range for epilepsy-related accidents and injuries in China was 12.20–34.80. SMRs ranged 1.90–6.60 across the reminder of the world. SMRs for drowning in China ranged from 5.60–82.40, and ranged from 2.00–13.80 across the remainder of the world. The two studies providing data for China involved rural populations.[82,99] The authors noted that many patients in the rural areas had

not been treated with standard AED therapy regimens, with the treatment gap owing mostly to low socio-economic status.[82] They suggested the increased risk of premature epilepsy-related death could be partly associated, therefore, with poverty. Also noted in relation to the markedly high drowning risks was that drowning in China mainly occurred in ponds and paddy fields, which differs from the bathtubs and swimming pools often mentioned in developed countries.[99] This suggests geography and occupational factors played a role in worsening epilepsy-related risks of drowning as many in rural China depend on the water resources for rice and freshwater fish farming.[99]

Figure 2-2 shows that PMRs tended to be reported in low-income populations and were generally high amongst all countries studied, with no obvious patterns between the countries.

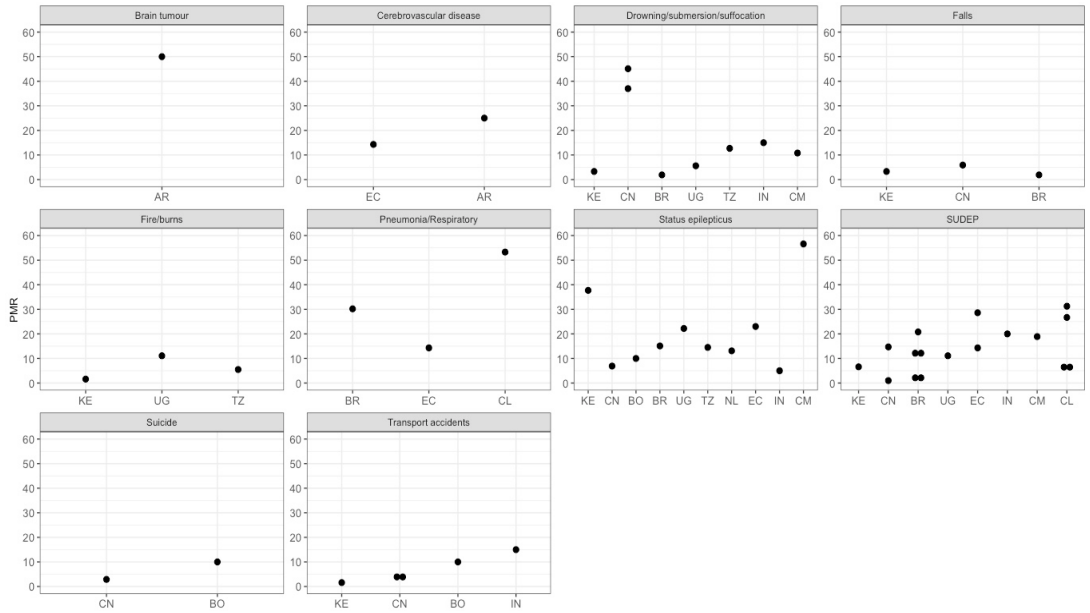


Figure 2-2: Dot plot panel of epilepsy-related causes of death by PMR and country

Abbreviations: AR – Argentina; EC – Ecuador; KE – Kenya; CN – China; BR – Brazil; UG – Uganda; TZ – Tanzania; IN – India; CM – Cameroon; BO – Bolivia; NL

– The Netherlands; CL – Chile; PMR – Proportional mortality ratio; SIDEP – Sudden unexpected death in epilepsy.

A general limitation of studies reporting mortality in PWE we noted was that there was little validation of the diagnostic accuracy of sources used to ascertain epilepsy cases. This is an important step in enhancing transparency of the work done, particularly where case-ascertainment involves using administrative datasets.[68,119] The next two chapters focus on clarifying the accuracy of using administrative data to identify epilepsy cases as these were the case-ascertainment sources for our in-depth analysis of epilepsy-related deaths in the remainder of the thesis.

Chapter 3. The accuracy of using administrative healthcare data to identify epilepsy cases: A systematic review of validation studies

3.1 Introduction

The first step in investigating epilepsy-related deaths on a population-level using administrative healthcare datasets is to assess the general accuracy of such datasets in identifying epilepsy cases. This is possible through systematically reviewing all of the available studies that have assessed the diagnostic accuracy of such datasets within their own populations. In 2015, Bossuyt *et al.* proposed guideline Standards for Reporting Diagnostic Accuracy (STARD) studies.[120] They indicate that such studies should have the administratively coded diagnoses compared against a reference standard in which the clinical diagnosis is confirmed (e.g. expert-assessed medical records) to see if the administrative diagnoses correctly reflect the reference standard. In this way, the study should report at least one measure of accuracy within PPV, sensitivity, NPV, specificity, or area under the curve (AUC). The guidelines recognise that diagnostic accuracy studies are not without risk of biases from, for example, different methods of participant recruitment, data collection, different administrative datasets used, and different analyses methods performed. For this reason, any systematic review assessing diagnostic accuracy studies should include risk of bias and study quality assessment.

An understanding of how accurately administrative data identify epilepsy through systematically reviewing all of the evidence will provide a strong

foundation for exploring the diagnostic accuracy of administrative Scottish data capturing epilepsy and identifying the optimal case-ascertainment algorithm for epilepsy within the remaining epilepsy-related mortality work of this thesis. In the current chapter, we report both the protocol and manuscript for a systematic review of studies assessing the accuracy of using administrative healthcare data to identify epilepsy cases. The protocol served to help develop methods undertaken in the main systematic review. The protocol was published in *BMJ Open* in 2018,[121] and the review was published in *Epilepsia* in 2020.[62] The published manuscripts have been used as the body of this thesis chapter, with permissions to do so granted from publishers. References listed in the published manuscripts are integrated into the thesis references.[65,67-69,94,96,111-113,120-178]

Student contributions: I came up with an idea to undertake a systematic review of studies validating the diagnostic accuracy of epilepsy coding globally as I hadn't seen this to have been done before. I then designed the searchers, screened studies, abstracted data, and undertook risk of bias assessments, with the support of a second reviewer undertaking these steps in parallel. I analysed the data then drafted and successfully submitted the work for publication.

3.2 Protocol

BMJ Open

Accuracy and utility of using administrative healthcare databases to identify people with epilepsy: a protocol for a systematic review and meta-analysis

Gashirai K Mbizvo,[1,2] Kyle Bennett,[1] Colin R Simpson,[3,4] Susan E Duncan,[1,2] Richard F M Chin[1,5]

To cite: Mbizvo GK, Bennett K, Simpson CR, *et al*. Accuracy and utility of using administrative healthcare databases to identify people with epilepsy: a protocol for a systematic review and meta-analysis. *BMJ Open* 2018;**8**:e020824. doi:10.1136/bmjopen-2017-020824

► Prepublication history for this paper is available online. To view these files, please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2017-020824).

Received 15 January 2018
Revised 12 March 2018
Accepted 15 May 2018

Check for updates

[1]Muir Maxwell Epilepsy Centre, University of Edinburgh Centre for Clinical Brain Sciences, Edinburgh, UK
[2]Department of Clinical Neurosciences, Western General Hospital, Edinburgh, UK
[3]Faculty of Health, Victoria University of Wellington, Wellington, New Zealand
[4]Usher Institute of Population Health Sciences and Informatics, University of Edinburgh School of Molecular Genetic and Population Health Sciences, Edinburgh, UK
[5]Department of Neurosciences, Royal Hospital for Sick Children, Edinburgh, UK

Correspondence to
Dr Gashirai K Mbizvo;
gashirai.mbizvo@ed.ac.uk

ABSTRACT

Introduction In an increasingly digital age for healthcare around the world, administrative data have become rich and accessible tools for potentially identifying and monitoring population trends in diseases including epilepsy. However, it remains unclear (1) how accurate administrative data are at identifying epilepsy within a population and (2) the optimal algorithms needed for administrative data to correctly identify people with epilepsy within a population. To address this knowledge gap, we will conduct a novel systematic review of all identified studies validating administrative healthcare data in epilepsy identification. We provide here a protocol that will outline the methods and analyses planned for the systematic review.

Methods and analysis The systematic review described in this protocol will be conducted to follow the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines. MEDLINE and Embase will be searched for studies validating administrative data in epilepsy published from 1975 to current (01 June 2018). Included studies will validate the International Classification of Disease (ICD), Ninth Revision (ICD-9) onwards (ie, ICD-9 code 345 and ICD-10 codes G40–G41) as well as other non-ICD disease classification systems used, such as Read Codes in the UK. The primary outcome will be providing pooled estimates of accuracy for identifying epilepsy within the administrative databases validated using sensitivity, specificity, positive and negative predictive values, and area under the receiver operating characteristic curves. Heterogeneity will be assessed using the I^2 statistic and descriptive analyses used where this is present. The secondary outcome will be the optimal administrative data algorithms for correctly identifying epilepsy. These will be identified using multivariable logistic regression models. 95% confidence intervals will be quoted throughout. We will make an assessment of risk of bias, quality of evidence, and completeness of reporting for included studies.

Ethics and dissemination Ethical approval is not required as primary data will not be collected. Results will be disseminated in peer-reviewed journals, conference presentations and in press releases.

PROSPERO registration CRD42017081212.

Strengths and limitations of this study

- The protocol describes what will be the first systematic review to conduct a worldwide assessment of the accuracy of administrative data in identifying epilepsy and the optimal disease-identification algorithms.
- This protocol also describes what will be the first systematic review to make an assessment of risk of bias, quality of evidence, and completeness of reporting for studies validating administrative healthcare data in epilepsy identification.
- The review described in this protocol will be limited to assessing the use of administrative data in diagnosing epilepsy within observational studies, which are more prone to bias than randomised controlled trials.
- A systematic review of the diagnostic accuracy of administrative data within randomised controlled trials in epilepsy remains to be completed and is out of the scope of the current review.

INTRODUCTION

Administrative healthcare databases are electronic data sources that consist of demographic, diagnostic and clinical information routinely collected about patients when they use a healthcare service.[1] They are often national and mandatory, and therefore they have the potential to provide a relatively cheap, widely available and less intrusive resource for medical research.[1] However, the accuracy of the information held in an administrative database needs to be validated before such use can be made. This is because the administrative data were not originally collected for research, but for other purposes such as assisting in health insurance claims. The clinical information held may therefore lack the rigour in accuracy that might be expected in scientifically collected data.

BMJ Open: first published as 10.1136/bmjopen-2017-020824 on 30 June 2018. Downloaded from http://bmjopen.bmj.com/ on June 30, 2021 at University of Edinburgh. Protected by copyright.

Furthermore, the data may be limited by inaccurate or incomplete hospital discharge letters or clinical coding transcription errors.[2]

The validation of administrative data involves comparing the diagnostic codes held within the administrative database against a reference standard (such as medical records) in order to quantify the number of instances in which the administrative diagnosis made matches the diagnosis in the reference standard (deemed to be the true diagnosis). In this way, the administrative database can be handled like a diagnostic test and measures of disease-identification accuracy calculated. These measures usually include the sensitivity, specificity and the positive or negative predictive value (PPV or NPV, respectively). Optimal disease-identification algorithms can also be determined by making relative comparisons of predictive values after adding in data from other variables recorded in an administrative database, such as drug combinations, investigations, and procedures.

There are many administrative databases worldwide in which the process of validation has been successfully performed for many diagnostic codes.[3–9] There are also examples of where the results of these have been pooled successfully into systematic review to increase confidence in the estimates made and scrutinise the quality of evidence, and this has led to changes in practice.[4] There has been limited systematic review of the validation of administrative databases in capturing epilepsy as a diagnosis. The only systematic review[10] on this subject included only studies from the USA or Canada and therefore excluded 121 studies because the data sources were not from these two countries. Furthermore, the 11 studies included were published between 2000 and 2010, making the conclusions nearly a decade old now. With health informatics now at the forefront of epidemiological disease surveillance, it is important to have an update on performance of the administrative disease-identification codes. Only one of the included studies evaluated the performance of the International Classification of Disease (ICD), 10th Revision (ICD-10) system in capturing epilepsy within administrative datasets[11]; the remainder evaluated the older ICD, 9th Revision (ICD-9) system.[10] The review also made no assessment of risk of bias, quality of evidence and completeness of reporting for included studies. This limits the confidence with which conclusions can be interpreted. There is now need for a more contemporary systematic review of the validation of administrative databases in capturing epilepsy. This should include evaluating performance of the ICD-10 system, as well as other non-ICD disease classification systems used, such as Read Codes in the UK.[12 13] The review should include studies from anywhere in the world in order to give clinicians and researchers a representative picture of the performance of administrative data in capturing epilepsy as a diagnosis and in order to allow more generalisable diagnostic algorithms to be suggested. Furthermore, the review should make an assessment of risk of bias, quality of evidence and completeness of reporting for included studies. These are the aims of the proposed systematic review described in this protocol. This will help researchers and clinicians better understand the accuracy of global estimates for incidence, prevalence and population characteristics in epilepsy, which have largely been made using administrative data.

Aims and objectives

The study hypotheses are:

1. Administrative data can correctly identify people with epilepsy within a population with a high degree of accuracy. We predict the PPV to be above 80%.
2. The optimal disease-identification algorithms for epilepsy within administrative datasets take into account diagnoses, investigations and drug combinations.[11 14–20]

The aim of the systematic review is to quantify the disease-identification accuracy and algorithm performance of administrative healthcare data in epilepsy. To this end, the research questions are:

1. How accurately do administrative data identify epilepsy within a population as measured by sensitivity, specificity, PPV, NPV, or area under the receiver operating characteristic curve (AUC) analysis (which are the approved accuracy measures described in the Standards for Reporting Diagnostic Accuracy (STARD) statement)?[21] This will be the primary outcome.
2. What are the optimal administrative data algorithms for correctly identifying epilepsy within a population? This will be the secondary outcome.

A preliminary feasibility search of the MEDLINE database via PubMed identifies at least nine studies validating diagnostic epilepsy codes held within administrative databases around the world that could be used to answer these research questions.[11 13–20]

METHODS AND ANALYSIS

This protocol follows the Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols (PRISMA-P) checklist.[22 23] The systematic review will follow the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) checklist.[24]

Eligibility criteria

We will include studies according to the following criteria:

Language: there will be no language restriction on full-length articles, although abstracts will need to be in English to allow the authors to screen them. We will seek translations for full-length articles not written in English that appear eligible in abstract. These will remain in the section for 'studies still awaiting classification' and will feature in subsequent updates to the review if not translated by the time of initial publication.

Setting: there will be no restrictions by study location worldwide. Where possible, we will show pooled accuracy and best algorithm data for administrative datasets from individual countries in addition to pooled global estimates of these measures.

BMJ Open: first published as 10.1136/bmjopen-2017-020824 on 30 June 2018. Downloaded from http://bmjopen.bmj.com/ on June 30, 2021 at University of Edinburgh. Protected by copyright.

Databases: the data sources will be routine administrative healthcare databases. This means that the data should have been routinely and passively collected without an a priori research question.[4] We will include databases containing diagnostic codes for epilepsy classified on the ICD system, where we will restrict this to studies using the ICD-9 onwards (active from 1975 to 1994).[25] This is because although the ICD system is currently in its 10th revision (ICD-10; active from 1994 to present) and the primary coding system is used by many countries around the world, a significant proportion of countries, particularly developing ones, still use the ICD-9 system.[11] ICD-9 code 345 and ICD-10 codes G40 and or G41 will be used to identify epilepsy diagnoses. We will provide summary measures of accuracy and best algorithms for any other disease classification systems used in studies separately, for example, the primary care Read Code system used in the UK.[12 13]

Study design: prospective or retrospective observational studies including cohort or case–control designs that are community-based/population-based or primary/secondary/tertiary care-based and have used administrative databases.

Participants: people with epilepsy of all ages. Where available, we will additionally report data for adults and children separately.

Observations and outcomes: studies will need to employ a validation process for diagnostic epilepsy codes, that is, estimate the disease-identification accuracy of the epilepsy codes held within the database using sensitivity, specificity, PPV, NPV or AUC analysis.[21] In this, true positives (TPs) and false negatives (FNs) will be considered as the patient has the disease and the administrative diagnosis is positive or negative, respectively. True negatives (TNs) and false positives (FPs) will be considered as the patient does not have the disease and the administrative diagnosis is negative or positive, respectively. Sensitivity will be considered as the ability of the administrative database to correctly identify those patients with the disease (TP/(TP+FN)). Specificity will be considered as the ability of the administrative database to correctly identify those patients without the disease (TN/(TN+FP)). PPV will be considered as how likely it is that a patient has the disease if the administrative diagnosis is positive (TP/(TP+FP)). NPV will be considered as how likely it is that this patient does not have the disease if the administrative diagnosis is negative (TN/(TN+FN)). AUC analysis will be considered as where TP and FP are plotted against each other in a perfect correlation as reference to show poor test accuracy, then accuracy measured as the area under the curve created by plotting the actual values against each other.[26]

Studies may also use diagnostic as well as other variables (eg, admissions, drugs or investigations) to calculate optimal disease-identification algorithms for epilepsy within the database. Studies will need to provide a clinical reference standard. An appropriate clinical reference standard will be medical records, clinical assessment, or a validated disease registry.[21]

Timeframe: studies conducted from 1 January 1975 to 01 June 2018. The year 1975 represents the advent of ICD-9.[25]

We will exclude studies according to the following criteria:

Data reported in systematic reviews unless we can identify the primary data, for example, by contacting authors of the original source.

Conference proceedings abstracts or studies not written in English where we are unable to obtain the meta-data from authors or full-length manuscript translations remain awaited, respectively.

Information sources

Studies will be identified from the following sources

- *Electronic databases:* we will search 01 January 1975–01 June 2018 for studies meeting the inclusion criteria within the MEDLINE (Ovid interface) and Embase (Ovid interface) databases. The search strategies are outlined in table 1.
- *Conference proceedings:* for conference abstracts that appear to meet the inclusion criteria but do not have a full-length article published, we will contact authors directly to request metadata.
- *Reference lists:* we will also identify studies meeting the inclusion criteria from the reference lists of included studies and relevant reviews identified through the electronic database searches.

Study records

Data management

Literature search results will be uploaded onto Review Manager 5, an internet-based software program that facilitates collaboration among reviewers. Citation titles and abstracts will be uploaded. GKM will then screen titles and abstracts to identify and exclude duplicate publications. Duplicate publications will be identified by comparing author names, study titles, sample sizes, outcomes used, and any other information held in the abstracts. All reviewer authors will have access to the systematic review process via the internet-based review software program, and this will create an audit trail of studies included/excluded, data analysis steps, and subsequent manuscript revisions. All data will be held within the management software and password protected.

Selection process

Once duplicates have been excluded, two review authors (GKM and KB) will independently screen the titles and abstracts yielded from the databases searches against the inclusion criteria. Where titles and abstracts indicate that a study may meet the inclusion criteria or where there is uncertainty about this, the full-length manuscripts will be

BMJ Open: first published as 10.1136/bmjopen-2017-020824 on 30 June 2018. Downloaded from http://bmjopen.bmj.com/ on June 30, 2021 at University of Edinburgh. Protected by copyright.

Table 1 Search strategies for MEDLINE and Embase

Ovid MEDLINE Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Ovid MEDLINE Daily and Ovid MEDLINE 1946 to 01 June 2018	Embase 1974 to 01 June 2018
1. Epilepsy, Complex Partial/or Epilepsy, Reflex/or Epilepsy, Absence/or Drug Resistant Epilepsy/or Epilepsy/ or Epilepsy, Rolandic/or Epilepsy, Partial, Motor/or Epilepsy, Benign Neonatal/or Epilepsy, Tonic-Clonic/or Epilepsy, Post-Traumatic/or Epilepsy, Partial, Sensory/or Epilepsy. mp. or Epilepsy, Temporal Lobe/or Epilepsy, Frontal Lobe/ or Myoclonic Epilepsy, Juvenile/or Epilepsy, Generalized/	1. reflex epilepsy/or photosensitive epilepsy/or grand mal epilepsy/or epilepsy.mp. or drug resistant epilepsy/ or experimental epilepsy/or severe myoclonic epilepsy in infancy/or childhood absence epilepsy/or benign childhood epilepsy/or catamenial epilepsy/or symptomatic epilepsy/ or startle epilepsy/or generalized epilepsy/or epilepsy/ or mesial temporal lobe epilepsy/or rolandic epilepsy/ or traumatic epilepsy/or myoclonic astatic epilepsy/ or temporal lobe epilepsy/or intractable epilepsy/ or focal epilepsy/or 'seizure, epilepsy and convulsion'/ or myoclonus epilepsy/or lateral temporal lobe epilepsy/ or frontal lobe epilepsy/
2. Databases, Factual/	2. factual database/or data base/
3. 'Reproducibility of Results'/	3. reproducibility/
4. Algorithms/or algorithm*.mp.	4. algorithm/or algorithm*.mp.
5. 3 or 4	5. 3 or 4
6. 2 and 5	6. 2 and 5
7. 1 and 6	7. 1 and 6
8. Administrative Claims, Healthcare/or administrativ*.mp. or insurance data*.mp. or claims data*.mp. or Veterans Health Administration.mp.	8. 'administrative claims (health care)"/or administrative*.mp. or insurance data*.mp. or claims data*.mp. or Veterans Health Administration.mp.
9. administrat* data*.mp.	9. administrat* data*.mp.
10. routin* data*.mp.	10. routin* data*.mp.
11. big data.mp.	11. big data.mp.
12. 8 or 9 or 10 or 11	12. 8 or 9 or 10 or 11
13. 1 and 12	13. 1 and 12
14. Algorithms/or algorithm*.mp.	14. algorithm/or algorithm*.mp.
15. 'Sensitivity and Specificity'/or 'Predictive Value of Tests'/ or predictive value.mp.	15. (specificity or sensitivity).mp. or 'sensitivity and specificity'/ or positive* predict* value*.mp. or negative* predict* value*. mp.
16. positive* predict* value*.mp.	16. area under the curve/or area* under* curve*.mp.
17. negative* predict* value*.mp.	17. roc curve/or receiver operating characteristic/or ROC curve*.mp.
18. sensitivity.mp.	18. 14 or 15 or 16 or 17
19. specificity.mp.	19. code*.mp. or 'Read code'/
20. area* under* curve*.mp. or Area Under Curve/	20. ICD-10.mp. or 'International Classification of Diseases'/or ICD-10/or disease classification/or ICD-9.mp. or ICD-9/
21. ROC Curve/or ROC curve*.mp.	21. 19 or 20
22. 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21	22. 18 and 21
23. code*.mp.	23. 1 and 22
24. (ICD-9 or ICD-10).mp. or 'International Classification of Diseases'/or Clinical Coding/or read code*.mp.	24. validation study/or validation Process/or validat*.mp. or validity/or predictive validity/or validity.mp.
25. 23 or 24	25. electronic medical record/or medical record/or medical record*.mp. or medical case note*.mp.
26. 22 and 25	26. Registries.mp. or register/
27. 1 and 26	27. 25 or 26
28. (validat* or validity).mp. or Validation Studies/	28. 24 and 27
29. Medical Records/or medical record*.mp. or medical case note*.mp.	29. 1 and 28
30. electronic health records.mp. or Medical Records Systems, Computerized/or Electronic Health Records/	30. 7 or 13 or 23 or 29
31. Registries.mp. or Registries/	31. limit 30 to yr='1975 -Current'
32. 29 or 30 or 31	32. animal/not human/
33. 28 and 32	33. 31 not 32
34. 1 and 33	
35. 7 or 13 or 27 or 34	
36. limit 35 to yr='1975 -Current'	
37. Animals/not Humans/	
38. 36 not 37	

downloaded and used to help decide. Where details in the manuscript are still insufficient for a decision to be made about eligibility, we will seek additional information from the study authors and automatically exclude studies where there is no response from authors after three weeks. We will record the reasons for excluding all excluded studies. The two review authors will compare their list of included and excluded trials and any disagreements

BMJ Open: first published as 10.1136/bmjopen-2017-020824 on 30 June 2018. Downloaded from http://bmjopen.bmj.com/ on June 30, 2021 at University of Edinburgh. Protected by copyright.

will be resolved by mutual discussion and, where necessary, adjunction by a third reviewer (RFMC/SED/CRS). Review authors will not be blind to the journal titles, study authors, or institutions.

Data collection process and items

Two review authors (GKM and KB) will independently abstract data about the primary and secondary outcomes from included studies using the data collection tool shown below. The additional information extracted maps onto the items contained within the STARD guidelines modified for epilepsy studies reporting diagnostic accuracy of administrative databases.[21] This will allow us to extract sufficient information to make a quality assessment of the completeness of included studies against the STARD checklist.[21] Any disagreements in the contents of data abstraction will be resolved by mutual discussion and, where necessary, adjunction by a third reviewer (RFMC/SED/CRS).

Data collection tool

- What is the study title?
- Who are the study authors?
- What is the year of study publication?
- What is the journal of publication?
- What country(s) was the study conducted in?
- Does the study explicitly identify as utilising 'administrative data' (yes/no)?
  - If not, how is this identified by the reviewer? For example, from descriptions given of the databases utilised and background knowledge about them or from correspondence with authors.
- Does the abstract provide a structured summary of study design, methods, results and conclusions (yes/no/unclear)?
- Does the introduction give a scientific and clinical background including intended use and clinical role of administrative data (yes/no/unclear)?
- Are the study objectives and hypotheses described (yes/no/unclear)?
  - If so, what are they?
- What was the intended study sample size and how was it determined?
- What is the study design?
- Was a study/participant flow diagram used?
- What are the eligibility criteria for participation in the validation cohort (ie, the cohort of patients to which the reference standard will be applied)?
- Where, when and how were potentially eligible validation cohort participants identified? Include within this:
  - What is the name of the administrative database(s) on which the validation cohort was identified?
  - What are the setting and location of the administrative database(s) from which the validation cohort was identified? For example, is it primary care, secondary care, tertiary care, outpatient care and emergency care?
  - What are the names of any hospitals/organisations affiliated with or using the administrative database routinely?
  - What is the size of the administrative database(s) on which the validation cohort was identified? That is, how many people/records does it hold in total?
  - What were the epilepsy ICD codes (or other disease classification system codes) used to identify the validation cohort within the administrative database(s)? That is, what are the diagnostic epilepsy codes that will be validated by the study?
  - What was the size of the validation cohort identified by the epilepsy codes? That is, give the number of participants identified by these diagnostic epilepsy codes.
  - Did the validation cohort include identifying a sample of people (1) without epilepsy and (2) with epilepsy 'mimicker codes'?
    - If so, give details of the codes used and the number of participants for each of these groups.
    - Samples of people without epilepsy are often used to help calculate the specificity and NPV of an administrative database. ' Mimicker codes' are often interrogated as the conditions may resemble epilepsy. These may include, for example, classical migraine (ICD-9 code 346.x and ICD-10 code G43.1), transient cerebral ischaemia (ICD-9 code 435 and ICD-10 code G45), syncope (ICD-9 code 780.2 and ICD-10 code R55) or convulsion (ICD-9 code 780.3 and ICD-10 code R56.0 or R56.8), which are intended to be used for organic convulsions but not for epilepsy.[11]
- What other information was obtained about an individual to help identify epilepsy on the administrative database? For example, describe if they linked an individual's ICD epilepsy codes with the investigations they underwent (such as an electroencephalogram (EEG)) or the antiepileptic drug (AED) they were taking.
- What were the demographic and clinical characteristics of the validation cohort? That is, age, gender, type of epilepsy, comorbidities and AEDs.
- Describe the reference standard. Include the following:
  - Name of the reference standard used.
  - What type of reference standard it was, for example, clinical assessment, medical records, or validated disease registry.
  - Any rational given for choosing this reference standard (if alternatives exist).
  - What were the number, training and expertise of persons reading the reference standard?
  - If more than one person read the reference standard, what were the measures of consistency given? For example, kappa statistic.
- Describe any methods used to blind persons reading the reference standard to how the validation cohort

BMJ Open: first published as 10.1136/bmjopen-2017-020824 on 30 June 2018. Downloaded from http://bmjopen.bmj.com/ on June 30, 2021 at University of Edinburgh. Protected by copyright.

were coded diagnostically on the administrative database; that is, how a person reviewing the medical records diagnosis of an individual was made unaware of their administrative ICD diagnosis.

- Describe any methods used to blind persons reading the diagnostic administrative data codes to results of the reference standard diagnoses.
- What method was used to estimate the disease-identification accuracy of the administrative database (ie, sensitivity, specificity, PPV, NPV or AUC)?
  - What were the results of this?
    - That is, provide figures for these estimates and, where possible, also the individual TP, FP, TN and FN figures.
    - Include the results of any cross-tabulation of the administrative data diagnoses results against the results of the reference standard diagnoses.
    - Describe the methods used and results of the measures used to estimate variability or precision of the diagnostic accuracy results (eg, 95% CI).
- What method was used to compare the diagnostic accuracy of variables within the administrative database? That is, describe the method used to determine an optimal diagnostic algorithm.
  - What were the results of this?
  - Describe the methods used and results of the measures used to estimate variability or precision of the algorithm estimates (eg, 95% CI).
- How were indeterminate or missing administrative database diagnoses or reference standard results handled?
- What were the time interval and any clinical interventions given between reference standard diagnosis and administrative database diagnosis?
- Describe any adverse events found from using the administrative database or reference standard.
- Summarise the study limitations described by the authors including any sources of potential bias, statistical uncertainty, generalisability limitation, and what they described as implications for practice.
- What were the study's sources of funding?

Systematic review and meta-analysis outcomes

The primary outcome will be providing pooled disease-identification accuracy estimates of the included administrative databases using sensitivity, specificity, PPV, NPV and AUC as the measures of accuracy. This will answer research question 1: how accurately do administrative data identify epilepsy within a population? We will provide an overall estimate of the accuracy of the ICD-9 coding system and that of the ICD-10 coding system in correctly identifying epilepsy cases. This will be done by pooling together individual accuracy estimates from all included studies in which ICD-9 or ICD-10 were used provided there is no significant heterogeneity between studies (as measured using the I^2 statistic). The preferred estimators will be means with standard errors (SEs) or medians with interquartile ranges (IQRs), dependent on distribution. We will also quote the 95% CIs for estimates. Where there is significant heterogeneity (I^2 statistic >50%), we will provide a descriptive analysis of the results and include ranges. Just as there may be heterogeneity introduced by making comparisons across different trial designs, we might expect there to be heterogeneity introduced by making comparisons across different healthcare systems. This is because there are likely to be differences in the accuracy of administrative healthcare data owing to differences in coding practice and/or the quality of reference standards between different healthcare systems. Therefore, we will also conduct subgroup analysis in which diagnostic accuracy results are pooled together for studies that have used the same or similar healthcare systems, for example, the National Health Service in the UK, Veterans Health Administration in the USA, healthcare systems with geographical overlap, and private-funded versus state-funded healthcare systems. We will also create subgroups in which results are pooled together within the following study design groups: prospective cohorts, retrospective cohorts, case–control studies, primary care studies, secondary care studies, tertiary care studies, paediatric studies (age <18 years), adult studies, and studies from the same country. Differences in the results for each subgroup may provide an important guide for future studies in the field, and they may also help to explain any statistical heterogeneity seen.

The secondary outcome will be the optimal administrative data algorithms for correctly identifying epilepsy within a population. For this, we will assign a dummy variable with a binary 0 = 'no' or 1 = 'yes' category to participants having the following:

A. A reference standard diagnosis of epilepsy (yes/no).
B. An administrative diagnosis code for epilepsy (yes/no).
C. Multiple administrative epilepsy diagnoses codes over time (yes/no).
D. Having previously had an EEG (yes/no).
E. Having previously had a computed tomography (CT) or magnetic resonance image (MRI) of the brain (yes/no).
F. Having previously had epilepsy surgery (yes/no);
G. Being on an individual AED (yes/no).
H. Being on two or more AEDs (yes/no).

Multivariable logistic regression models with A as the outcome variable and B–H individually and in combinations as the independent variables will be used in order to demonstrate the algorithm(s) best fitting the data across the included studies and to assess the significance of each variable's contribution to the model. The results of the logistic models will be displayed as sensitivity, specificity, PPV, NPV and AUC where possible, with 95% CIs and measures of interstudy heterogeneity provided using the I^2 statistic.

BMJ Open: first published as 10.1136/bmjopen-2017-020824 on 30 June 2018. Downloaded from http://bmjopen.bmj.com/ on June 30, 2021 at University of Edinburgh. Protected by copyright.

Table 2 Risk of bias and applicability judgements in QUADAS-2

Domain	Patient selection	Administrative database	Reference standard	Flow and timing
Description	Describe methods of patient selection: what is the review question?	Describe the administrative database and how it was used and interpreted: where available, include comment on how coding was done, by whom and whether there was reimbursement for coding.	Describe the reference standard and how it was conducted and interpreted: where available, include comment on quality of the reference standard, including the level of experience of clinicians making the diagnosis, access to diagnostic tests such as electroencephalography and telemetry and the thresholds/criteria used to make a diagnosis of epilepsy.	Describe any patients in the validation cohort who were not found within the reference standard or who were excluded from cross-tabulation of the administrative data diagnoses results against the results of the reference standard diagnoses: describe the time interval and any interventions between administrative database diagnosis and reference standard diagnosis.
Signalling questions (yes/no/unclear)	Was a consecutive or random sample of patients enrolled?	Were the administrative database diagnosis results interpreted without knowledge of the results of the reference standard diagnosis?	Is the reference standard likely to correctly classify the epilepsy?	Was there an appropriate interval between administrative database diagnosis and reference standard diagnosis?
	Was a case–control design avoided? Did the study avoid inappropriate exclusions?	If a diagnostic threshold was used, was it prespecified?	Were the reference standard results interpreted without knowledge of the results of the administrative database diagnosis?	Did all patients receive a reference standard? Did all patients receive the same reference standard? Were all patients included in the analysis?
Risk of bias: high/low/unclear	Could the selection of patients have introduced bias?	Could the conduct or interpretation of the administrative database have introduced bias?	Could the reference standard, its conduct or its interpretation have introduced bias?	Could the patient flow have introduced bias?
Concerns regarding applicability: high/low/unclear	Are there concerns that the included patients do not match the review question?	Are there concerns that the administrative database, its conduct or interpretation differ from the review question?	Are there concerns that epilepsy, as defined by the reference standard, does not match the review question?	

QUADAS-2, Quality Assessment of Diagnostic Accuracy Studies 2.

Risk of bias analysis

We will use the Quality Assessment of Diagnostic Accuracy Studies 2 (QUADAS-2)[27] tool to assess risk of bias within and across studies, modified for studies validating administrative data. This is summarised in the table 2 and will be completed independently by two review authors (GKM and KB) for each study, with disagreements resolved by mutual discussion and, where necessary, adjunction by a third reviewer (RFMC/SED/CRS). The tool consists of four key domains (see row 1) covering: (1) patient selection, (2) the administrative database, (3) the reference standard and (4) flow of patients through the study and timing of the administrative database and reference standard. Each domain is assessed in terms of the risk of bias (graded as high, low or unclear; see row 4) and the first three domains are also assessed in terms of concerns regarding applicability (see row 5). The description (see row 2) contains information used to support the risk of bias judgement. To help reach a judgement on the risk of bias, signalling questions are included (see row 3). These flag aspects of study design related to the potential for bias and aim to help reviewers make risk of bias judgements. If all signalling questions for a domain are answered 'yes', then risk of bias is judged 'low'. If any signalling question is answered 'no', then risk of bias is judged 'high'. If any signalling question is answered 'unclear', then risk of bias is judged 'unclear'. Applicability sections are structured in a similar way to the bias sections but do not include signalling questions. Review authors are asked

BMJ Open: first published as 10.1136/bmjopen-2017-020824 on 30 June 2018. Downloaded from http://bmjopen.bmj.com/ on June 30, 2021 at University of Edinburgh. Protected by copyright.

to record the information on which the judgement of applicability is made and then to rate their concern that the study does not match the review question.[27 28]

On completing the QUADAS-2 table, we will provide a risk of bias and applicability concerns graph demonstrating the review authors' judgements about each domain, presented as percentages across included studies. We will also provide a risk of bias and applicability concerns summary demonstrating review authors' judgements about each domain for each included study. We will use the Deek's test[29] to interrogate for publication bias. This test is specifically designed for detecting funnel plot asymmetry in reviews of diagnostic studies.[28]

Confidence in cumulative evidence

We will use the Grading of Recommendations Assessment, Development, and Evaluation (GRADE) approach to assess strength of the body of evidence.[30] The GRADE system classifies the quality of evidence into one of four grades:

1. *High*: further research is very unlikely to change our confidence in the estimate of effect.
2. *Moderate*: further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
3. *Low*: further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
4. *Very low*: any estimate of effect is very uncertain.[31]

A judgement is made on the individual studies used to provide the pooled effect estimates, and the quality of evidence is then downgraded by the cumulative presence of: (1) bias (see risk of bias analysis), (2) inconsistency (ie, heterogeneity present on I^2 statistic), (3) indirectness (ie, high concerns regarding applicability; see table 2), (4) imprecision (small sample sizes, wide CIs, and inadequately powered studies) and (5) publication bias (see Deek's test[29] comments).[32] GRADE classifications will be independently conducted by two review authors (GKM and KB) with any disagreements resolved by mutual discussion and, where necessary, adjunction by a third reviewer (RFMC/SED/CRS).

We will rate the completeness of reporting for each study out of 30 using the STARD 2015 checklist.[21] A score of 0–10, 11–20 and 21–30 will indicate a low, moderate and high quality of completeness of reporting, respectively.

PATIENT AND PUBLIC INVOLVEMENT

Patients and the public were not involved in development of the research question and outcome measures, nor the study design. The study does not involve patient recruitment, and patients were not involved in conduct of the study. We plan to liaise closely with patients, special interest groups, and charities in the dissemination of our results in printed and electronic media. Meta-data and information about the study will also be made available through our website (www.muirmaxwellcentre.com).

Ethics and dissemination

Ethical approval is not required as primary data will not be collected. Results will be disseminated in peer-reviewed journals, conference presentations, and in press releases. Meta-data and information about the study will also be made available through our website (www.muirmaxwellcentre.com) and via social media.

Contributors GKM and RFMC conceived the idea for the protocol and made the main contribution to planning and preparation of timelines for its completion. GKM and KB put together and tested various search strategies for the protocol and, after consultation with CRS, had these reviewed and approved by Marshall Dozier at the University of Edinburgh library whose support we acknowledge. GKM and RFMC planned the data extraction and statistical analysis, as well as of risk of bias, quality of evidence and completeness of reporting assessments. GKM designed the tables and wrote the first draft of the manuscript, which was then reviewed and amended by KB, CRS, SED and RFMC. All authors then approved the final written manuscript. RFMC is the guarantor for the work.

Funding This work was supported by Epilepsy Research UK (R44007) and the Juliet Bergqvist Memorial Fund.

Disclaimer The funders had no role in the design of the protocol, its preparation, analyses, interpretation of the data, manuscript preparation or decision to submit.

Competing interests None declared.

Patient consent Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

Author note Any future amendments of the protocol will be listed in this section along with a date, description, and rational for each amendment.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/

© Article author(s) (or their employer(s) unless otherwise stated in the text of the article) 2018. All rights reserved. No commercial use is permitted unless otherwise expressly granted.

REFERENCES

1. Garratt E, Barnes H, Dibben C. *Health administrative data: exploring the potential for academic research*. St Andrews: Administrative Data Liaison Service, 2010.
2. Loke YK. Use of databases for clinical research. *Arch Dis Child* 2014;99:587–9.
3. St Germaine-Smith C, Metcalfe A, Pringsheim T, *et al*. Recommendations for optimal ICD codes to study neurologic conditions: a systematic review. *Neurology* 2012;79:1049–55.
4. Abraha I, Montedori A, Eusebi P, *et al*. The current state of validation of administrative healthcare databases in Italy: a systematic review. *Pharmacoepidemiology and Drug Safety* 2012;21:400–00.
5. Traversa G, Bianchi C, Da Cas R, *et al*. Cohort study of hepatotoxicity associated with nimesulide and other non-steroidal anti-inflammatory drugs. *BMJ* 2003;327:18–22.
6. Hottes TS, Skowronski DM, Hiebert B, *et al*. Influenza vaccine effectiveness in the elderly based on administrative databases: change in immunization habit as a marker for bias. *PLoS One* 2011;6:e22618.
7. Lalmohamed A, Vestergaard P, Cooper C, *et al*. Timing of stroke in patients undergoing total hip replacement and matched controls: a nationwide cohort study. *Stroke* 2012;43:3225–9.
8. Schneeweiss S, Avorn J. A review of uses of health care utilization databases for epidemiologic research on therapeutics. *J Clin Epidemiol* 2005;58:323–37.
9. Krarup LH, Boysen G, Janjua H, *et al*. Validity of stroke diagnoses in a national register of patients. *Neuroepidemiology* 2007;28:150–4.

BMJ Open: first published as 10.1136/bmjopen-2017-020824 on 30 June 2018. Downloaded from http://bmjopen.bmj.com/ on June 30, 2021 at University of Edinburgh. Protected by copyright.

10. Kee VR, Gilchrist B, Granner MA, *et al*. A systematic review of validated methods for identifying seizures, convulsions, or epilepsy using administrative and claims data. *Pharmacoepidemiol Drug Saf* 2012;21 Suppl 1:183–93.
11. Jetté N, Reid AY, Quan H, *et al*. How accurate is ICD coding for epilepsy? *Epilepsia* 2010;51:62–9.
12. Benson T. The history of the read codes: the inaugural James Read memorial lecture 2011. *Inform Prim Care* 2011;19:173–82.
13. Fonferko-Shadrach B, Lacey AS, White CP, *et al*. Validating epilepsy diagnoses in routinely collected data. *Seizure* 2017;52:195–8.
14. Tu K, Wang M, Jaakkimainen RL, *et al*. Assessing the validity of using administrative data to identify patients with epilepsy. *Epilepsia* 2014;55:335–43.
15. Franchi C, Giussani G, Messina P, *et al*. Validation of healthcare administrative data for the diagnosis of epilepsy. *J Epidemiol Community Health* 2013;67:1019–24.
16. Reid AY, St Germaine-Smith C, Liu M, *et al*. Development and validation of a case definition for epilepsy for use with administrative health data. *Epilepsy Res* 2012;102:173–9.
17. Parko K, Thurman DJ. Prevalence of epilepsy and seizures in the Navajo Nation 1998-2002. *Epilepsia* 2009;50:2180–5.
18. Pugh MJ, Van Cott AC, Cramer JA, *et al*. Trends in antiepileptic drug prescribing for older patients with new-onset epilepsy: 2000-2004. *Neurology* 2008;70(22 Pt 2):2171–8.
19. Christensen J, Vestergaard M, Olsen J, *et al*. Validation of epilepsy diagnoses in the Danish National Hospital Register. *Epilepsy Res* 2007;75(2-3):162–70.
20. Tan M, Wilson I, Braganza V, *et al*. Development and validation of an epidemiologic case definition of epilepsy for use with routinely collected Australian health data. *Epilepsy Behav* 2015;51:65–72.
21. Bossuyt PM, Reitsma JB, Bruns DE, *et al*. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. *Clin Chem* 2015;61:1446–52.
22. Moher D, Shamseer L, Clarke M, *et al*. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Syst Rev* 2015;4:1.
23. Shamseer L, Moher D, Clarke M, *et al*. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. *BMJ* 2015;349:g7647.
24. Moher D, Liberati A, Tetzlaff J, *et al*. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ* 2009;339:b2535.
25. World Health Oraganisation. Classification of Diseases (ICD). 2016. www.who.int/classifications/icd/en/
26. Florkowski CM. Sensitivity, specificity, receiver-operating characteristic (ROC) curves and likelihood ratios: communicating the performance of diagnostic tests. *Clin Biochem Rev* 2008;29(Suppl 1):S83–7.
27. Whiting PF, Rutjes AW, Westwood ME, *et al*. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* 2011;155:529–36.
28. Leeflang MM, Deeks JJ, Takwoingi Y, *et al*. Cochrane diagnostic test accuracy reviews. *Syst Rev* 2013;2:82.
29. Deeks JJ, Macaskill P, Irwig L. The performance of tests of publication bias and other sample size effects in systematic reviews of diagnostic test accuracy was assessed. *J Clin Epidemiol* 2005;58:882–93.
30. Schünemann H, Broek J, Oxman A. *GRADE handbook for grading quality of evidence and strength of recommendation*. 3.2, 2009.
31. Mbizvo GK, Dixon P, Hutton JL, *et al*. Levetiracetam add-on for drug-resistant focal epilepsy: an updated Cochrane Review. *Cochrane Database Syst Rev* 2012;9:CD001901.
32. Singh S, Chang SM, Matchar DB, *et al*. Chapter 7: grading a body of evidence on diagnostic tests. *J Gen Intern Med* 2012;27(Suppl 1):47–55.

BMJ Open: first published as 10.1136/bmjopen-2017-020824 on 30 June 2018. Downloaded from http://bmjopen.bmj.com/ on June 30, 2021 at University of Edinburgh. Protected by copyright.

3.3 Summary

This protocol suggests a systematic review of diagnostic accuracy studies of administrative epilepsy datasets should assess not only ICD-9 and -10 codes, but also seek to assess any additional coding systems in place around the world. This would allow capture of, for example, the Read code system in the UK.[65] Read codes are a set of clinical descriptions that GP practices use to create electronic health records for patients within primary care. They are organised into chapters that cover a wide spectrum of information including demographics, clinical history, examination, medications, clinical procedures, occupation, and smoking status.

The aims from this protocol would be to estimate an overall accuracy for the identification of epilepsy within administrative data using this coded information, and to describe the optimal case-ascertainment algorithms; as assessed using PPV, sensitivity, NPV, specificity, or AUC. Risk of bias assessment would be included using the Quality Assessment of Diagnostic Accuracy Studies 2 (QUADAS-2) scoring system.[146] Two review authors would undertake study selection, data abstraction, and risk of bias assessments independently, with disagreements resolved by consensus or third reviewer adjunction.

Strengths of this study protocol include that it is extensive in the planned methodology, including illustrating a detailed electronic database search strategy that can be easily reproduced for the benefit of the main review and other researchers. The protocol attempts to deal with risk of bias, and is not restricted by country of study or disease coding system used. However, we

recognise that such an inclusive approach may increase the chances of heterogeneity between studies, and it would be important for the systematic review to try to deal with this appropriately. Whilst use of the I^2 statistic is suggested in the protocol for this purpose, other systematic reviews of validation studies in other conditions have also focused on descriptive assessment of heterogeneity based on evaluating differences in study characteristics and methods.[161,177] It may be that this will be a more suitable approach to take in the main review dependent on how many studies are identified from different countries and using different disease coding systems. Heterogeneity could be reduced by, for example, undertaking subgroup analysis by country or disease coding system.

Another limitation of the protocol is lack of a plan for statistical testing of interrater agreement for study selection and risk of bias assessment decisions. Cohen's kappa statistic can be used to test interrater agreement.[162,173] This would likely be helpful in a review of this topic as there are likely to be some differences in how diagnostic accuracy studies are performed around the world.

Finally, the protocol describes the assessment of epilepsy disease codes (ICD-10 G40–41, and ICD-9 345) but does not include how symptom codes (ICD-10 R56 and ICD-9 780.3, 780.39 for seizures) will be handled. These are likely to be reviewed within epilepsy validation studies, therefore it will be important for a systematic review to include estimation of the accuracy of these symptom codes also for epilepsy case-ascertainment.

In conclusion, this study protocol demonstrates a reproducible plan for systematically reviewing diagnostic accuracy studies of administrative epilepsy datasets. There are some additional features highlighted in the summary which would also need to be considered in such a systematic review.

3.4 Manuscript

Check for updates

Received: 30 July 2019 | Revised: 28 April 2020 | Accepted: 29 April 2020

DOI: 10.1111/epi.16547

CRITICAL REVIEW AND INVITED COMMENTARY

Epilepsia®

The accuracy of using administrative healthcare data to identify epilepsy cases: A systematic review of validation studies

Gashirai K. Mbizvo[1] | Kyle H. Bennett[1] | Christian Schnier[2] | Colin R. Simpson[2,3] | Susan E. Duncan[1,4] | Richard F.M. Chin[1,5]

[1]Muir Maxwell Epilepsy Centre, Centre for Clinical Brain Sciences, The University of Edinburgh, Edinburgh, UK

[2]Usher Institute of Population Health Sciences and Informatics, The University of Edinburgh, Edinburgh, UK

[3]School of Health, Faculty of Health, Victoria University of Wellington, Wellington, NZ

[4]Department of Clinical Neurosciences, Western General Hospital, Edinburgh, UK

[5]Royal Hospital for Sick Children, Edinburgh, UK

Correspondence
Gashirai K. Mbizvo, Muir Maxwell Epilepsy Centre, Child Life and Health, 20 Sylvan Place, EH9 1UW, Edinburgh, UK.
Email: gashiraimbizvo@hotmail.com

Funding information
Epilepsy Research UK, Grant/Award Number: R44007; The Juliet Bergqvist Memorial Fund, Grant/Award Number: N/A

Abstract

Our objective was to undertake a systematic review ascertaining the accuracy of using administrative healthcare data to identify epilepsy cases. We searched MEDLINE and Embase from 01/01/1975 to 03/07/2018 for studies evaluating the diagnostic accuracy of routinely collected healthcare data in identifying epilepsy cases. Any disease coding system in use since the International Classification of Diseases, Ninth Revision (ICD-9) was permissible. Two authors independently screened studies, extracted data, and quality-assessed studies. We assessed positive predictive value (PPV), sensitivity, negative predictive value (NPV), and specificity. The primary analysis was a narrative synthesis of review findings. Thirty studies were included, published between 1989 and 2018. Risks of bias were low, high, and unclear in 4, 14, and 12 studies, respectively. Coding systems included ICD-9, ICD-10, and Read Codes, with or without antiepileptic drugs (AEDs). PPVs included ranges of 5.2%–100% (Canada), 32.7%–96.0% (USA), 47.0%–100% (UK), and 37.0%–88.0% (Norway). Sensitivities included ranges of 22.2%–99.7% (Canada), 12.2%–97.3% (USA), and 79.0%–94.0% (UK). Nineteen studies contained at least one algorithm with a PPV >80%. Sixteen studies contained at least one algorithm with a sensitivity >80%. PPV was highest in algorithms consisting of disease codes (ICD-10 G40-41, ICD-9 345) in combination with one or more AEDs. The addition of symptom codes to this (ICD-10 R56; ICD-9 780.3, 780.39) lowered PPV. Sensitivity was highest in algorithms consisting of symptom codes with one or more AEDs. Although using AEDs alone achieved high sensitivities, the associated PPVs were low. Most NPVs and specificities were >90%. We conclude that it is reasonable to use administrative data to identify people with epilepsy (PWE) in epidemiological research. Studies prioritizing high PPVs should focus on combining disease codes with AEDs. Studies prioritizing high sensitivities should focus on combining symptom codes with AEDs. We caution against the use of AEDs alone to identify PWE.

[Correction added on June 09, 2020, after first online publication: "USA" was removed from affiliation 3.]

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.
© 2020 The Authors. *Epilepsia* published by Wiley Periodicals LLC on behalf of International League Against Epilepsy

KEYWORDS
diagnostic test accuracy, international classification of diseases, positive predictive value, routine data, seizures, sensitivity

1 | INTRODUCTION

Administrative healthcare data consist of routine diagnostic and procedural information collected about patients when they use healthcare services.[1] These national data sets are widely available and less-intrusive potential resources for medical research.[1] However, their diagnostic accuracy requires validation because the data were collected originally for routine nonscientific purposes. Poor or incomplete hospital discharge letters and clinical coding errors are examples of potential sources for administrative data inaccuracies.[2] Systematic reviews of administrative data validation studies increase confidence in case-ascertainment accuracy estimates for a particular condition and scrutinize the quality of available evidence.[3-11] To date, there has been only one systematic review of studies validating the accuracy of administrative epilepsy data sets.[12] Although this was helpful, the 11 studies reviewed were from the United States and Canada alone and published up to the year 2010 only. The International Classification of Diseases, Ninth Revision (ICD-9) system was used in all but one study, and the review lacked a risk of bias assessment. We now provide an updated systematic review of epilepsy validation studies worldwide. We include risk of bias assessments and evaluate ICD-9, ICD-10, and other coding systems routinely used.[13,14] We focus on positive predictive value (PPV) and sensitivity, as these are the most commonly reported validation outcomes in the literature.[11,12] However, we also report negative predictive value (NPV) and specificity, where available.

Key Points

- Administrative healthcare data can be used confidently to identify people with epilepsy in epidemiological research.
- Studies prioritizing high positive predictive values should focus on combining disease codes with antiepileptic drugs.
- Studies prioritizing high sensitivities should focus on combining symptom codes with antiepileptic drugs.
- Antiepileptic drugs alone are unlikely to accurately capture persons with epilepsy from administrative data.

2 | METHODS

The aims and inclusion criteria were established before conduct of the review and the a priori study protocol was registered on PROSPERO (CRD42017081212-https://bit.ly/2VOdoNj),[15] and published.[16] The review follows the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) system of reporting.[17] Ethical approval was not required.

2.1 | Eligibility criteria

We included studies according to the following criteria:

- *Study country:* No restrictions;
- *Language:* No restrictions: translations were sought for any non-English texts;
- *Study design:* No study type was excluded;
- *Participants:* People with epilepsy of all ages;
- *Validated database:* Studies evaluating the diagnostic accuracy of routinely collected (administrative) healthcare data using ICD-9, ICD-10, or any other diagnostic coding system in use since the advent of ICD-9 in 1975[18];
- *Reference standard:* There were no minimum requirements for the types of gold standard used. The risk of study bias was assessed using the Quality Assessment of Diagnostic Accuracy Studies 2 (QUADAS-2) tool[19];
- *Observations and outcomes:* Studies had to report at least a PPV, sensitivity, NPV, or specificity estimate, or provide data from which these could be calculated;
- *Timeframe:* Studies conducted from 01/01/1975 onward.

2.2 | Information sources, literature searches, and study selection

MEDLINE and Embase (including Embase gray literature)[15,20,21] were searched from 01/01/1975 to 03/07/2018 (date last searched: 04/07/2018) for potentially eligible studies (see Appendix S1 for search strategy). Reference lists of manuscripts screened were also searched to identify potentially eligible studies. Any studies made aware to us by colleagues were also screened. Two authors (GKM and KB) independently screened titles, abstracts, or full-length articles for eligibility and inclusion, with disagreements resolved by

consensus or, where necessary, third reviewer adjunction (RFMC).

2.3 | Data abstraction

GKM and KB independently abstracted data pertaining to study location, participant age, administrative data setting, coding system, sample size, algorithm(s), reference standard details, PPV, sensitivity, NPV, specificity, and confidence intervals (CIs). We used a pre-piloted data abstraction form (Appendix S2). We made attempts to collect any relevant missing or unpublished study data by contacting the corresponding study author by email. We settled any differences in abstracted data by mutual review of the relevant study article.

2.4 | Quality assessment

We assessed the risk of bias and applicability concern for included studies using the Quality Assessment of Diagnostic Accuracy Studies 2 (QUADAS-2)[19] tool adapted by Wilkinson et al[11] for administrative validation studies (Table 1). GKM and KB independently graded each study's risk of bias and applicability concern as low, unclear, or high across the four categories: patient selection, administrative database ("index test"), reference standard, and study participant flow.[19] We used the QUADAS-2 abstraction form in Appendix S3 (16 quality decisions per study), with disagreements resolved by consensus or, where necessary, third reviewer adjunction (RFMC).

2.5 | Summary measures and synthesis of results

Cohen's kappa statistic was used estimate interrater agreement (above chance) between GKM and KB for study eligibility selection and QUADAS-2 quality assessment.[22,23] We reported all of the algorithms and associated PPVs, sensitivities, NPVs, and specificities for epilepsy provided by each study, but we excluded any algorithms attempting to further identify epilepsy by seizure type or epilepsy syndrome to reduce clinical heterogeneity of the condition being assessed. The primary analysis was a narrative synthesis of results for PPV, sensitivity, NPV, and specificity.[10-12,24] This included results summarized as ranges across studies for different countries to assist in correcting for potential differences in epilepsy prevalence between different countries and also to facilitate reporting and interpreting results from a wide range of countries (each with potentially varying coding practices). Results were also summarized as ranges across studies exclusively using medical records as their diagnostic gold standard to allow interpretation of results when applied across a consistent optimal reference standard. We created forest plots consisting of the best and worst PPV, sensitivity, NPV, and specificity from each study, allowing us to discuss the study characteristics and diagnostic algorithms associated with the best cases and worst cases of PPV, sensitivity, NPV, and specificity. We did not pool PPV and NPV estimates into a formal meta-analysis, as this approach is cautioned against owing to the expected heterogeneity from variation in disease prevalence and different disease positivity thresholds.[11,12,24] We used a Reitsma random-effects bivariate meta-analysis model to estimate a summary sensitivity and specificity because these two outcomes are relatively resilient to changes in disease prevalence.[25,26] This analysis is only possible across studies providing all of the true positives, false positives, true negatives, and false negatives.[26] All estimates were reported with 95% CIs where possible. We considered participants aged ≥18 years as adults and <18 years children. Data were processed and analyzed using RSTUDIO Version 1.2.1335, Boston, MA: RStudio, Inc.

3 | RESULTS

3.1 | Search and study characteristics

Thirty studies[14,27-55] were included from 197 potentially eligible records (Figure 1). The level of study selection agreement between authors was graded as moderate,[23] detailed in Appendix S4. Twenty-seven studies were identified from searching MEDLINE and Embase, and three[36,50,55] were identified from hand-searching the reference lists of potentially eligible studies. One study had unpublished validation outcome data for epilepsy, which were kindly provided to us separately by the corresponding author upon request.[40] The included studies were all published between 1989 and 2018 (only two before 2000).[31,49] All were carried out in high-income countries. Excluded studies are listed in Appendix S5.[21] Participants were adults in eight studies,[29,33,40,42,46,48,52,53] children in two studies,[27,41] both adults and children in 14 studies,[14,28,34-39,43,49-51,54,55] and unclear age in six studies.[30-32,44,45,47] Epilepsy was the target diagnostic condition in all included studies. The majority (19 studies) used hand-searched medical records as the diagnostic gold standard.[14,30,32,37-43,45-48,50-54] In the remaining 11 studies, the other gold standards used were the general practitioner (GP: 3 studies),[35,36,49] being on antiepileptic drugs (AEDs: three studies, one of which used rufinamide only),[31,33,44] medical records and/or parent telephone interview,[27] an epilepsy center patient list,[28] a previous epilepsy prevalence study,[29] another administrative data set,[34] and a specialist epilepsy database.[55] Administrative data sets validated included inpatients, outpatients, pharmacy, accident and emergency, physician claims, and primary care. Algorithms included ICD-9, ICD-10 (including two studies

TABLE 1 Risk of bias and applicability concern judgements in QUADAS-2[19] adapted for administrative validation[11]

Domain	Patient selection	Administrative database	Reference standard	Flow and timing
Description	Describe methods of patient selection:	Describe the administrative database and how it was used and interpreted:	Describe the reference standard and how it was conducted and interpreted:	Describe any patients in the validation cohort who were not found within the reference standard or who were excluded from cross-tabulation of the administrative data diagnoses results against the results of the reference standard diagnoses:
	What is the study question?	Where available, include comment on how coding was done, by whom, and whether or not there was reimbursement for coding.	Where available, include comment on quality of the reference standard, including the level of experience of clinicians making the diagnosis, access to diagnostic tests such as electroencephalography and telemetry, and the thresholds/criteria used to make a diagnosis of epilepsy.	Describe the time interval and any interventions between administrative database diagnosis and reference standard diagnosis:
Signalling questions (yes/no/unclear)	Was a consecutive or random sample of patients enrolled?	Were the administrative database diagnosis results interpreted without knowledge of the results of the reference standard diagnosis?	Is the reference standard likely to correctly classify the epilepsy?	Was there an appropriate interval between administrative database diagnosis and reference standard diagnosis?
	Did the study avoid inappropriate exclusions?	If a diagnostic threshold was used, was it pre-specified?	Were the reference standard results interpreted without knowledge of the results of the administrative database diagnosis?	Did all patients receive a reference standard? Did all patients receive the same reference standard? Were all patients included in the analysis?
Risk of bias: High/low/unclear	Could the selection of patients have introduced bias?	Could the conduct or interpretation of the administrative database have introduced bias?	Could the reference standard, its conduct, or its interpretation have introduced bias?	Could the patient flow have introduced bias?
Concerns regarding applicability: High/low/unclear	Are there concerns that the included patients do not match the study question?	Are there concerns that the administrative database, its conduct, or interpretation differ from the study question?	Are there concerns that epilepsy, as defined by the reference standard, does not match the study question?	

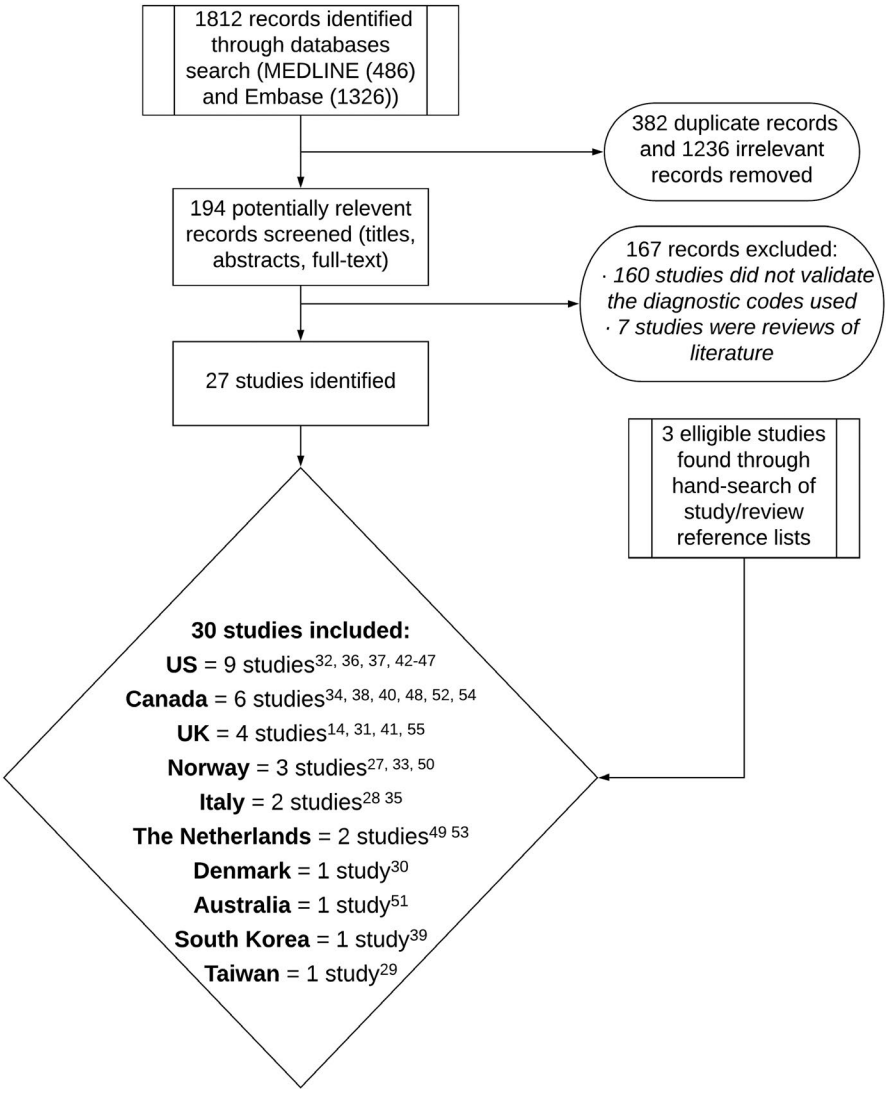


FIGURE 1 Flow diagram of study selection process. The three studies identified from hand-searching reference lists of screened literature were Pickrell 2015[55] (found in the included study Fonferko-Shadrach 2017[14]), Syvertsen 2015[50] (found in the excluded study Aaberg 2016[59]), and Frost 2000[36] (found in the systematic review Kee 2012[12])

that combined ICD-8 and ICD-10 data sets together),[30,52] Read Codes (Version 2), AEDs, hospitalizations, and procedures such as electroencephalography (EEG) or vagus nerve stimulation (VNS). Twenty-eight studies assessed PPV,[14,27-43,45-54] of which 14 also assessed sensitivity.[14,28,29,33-35,37,38,40,42,48,51,52,54] Two studies assessed sensitivity alone.[44,55] Fourteen studies assessed NPV and specificity,[14,28,29,33-35,37,38,40,42,48,51,52,54] and one study assessed specificity without NPV.[44]

3.2 | Quality assessment

The mean level of quality assessment agreement between authors was graded as fair,[23] detailed in Appendix S6A. Four studies had low risks of bias and low applicability concerns across all categories (Appendix S6B). Fourteen studies had one or more high-risk ratings. High risks of bias from reference standard conduct were common. This was largely due to inadequate blinding of the reference standard reader,[35,36,49] use of unvalidated administrative data sets as the reference standard,[34] and use of AEDs as the reference standard.[31,33,44] High risks of bias from study flow and timing were also common, largely due to all study participants not receiving the same reference standard,[27,34,36,40] or due to the diagnostic alternatives of "epilepsy" or "not epilepsy" for missing participants not being explored in a "best-case" or "worst-case" sensitivity analysis, respectively.[14,49,54] The remaining 12 studies had ratings of both low and unclear risks across different categories, resulting in overall risks being graded unclear.[28,32,37,41-43,45-47,50,52,55] The most common reason for this was lack of clarity about whether or not those reading the reference standard were blinded to the participants' administrative codes.

3.3 | Narrative analysis

3.3.1 | Overall PPV and sensitivity

Overall, 28 studies generated 172 algorithms estimating PPV and 121 of these algorithms (generated by 13 studies)

also estimated sensitivity. Figure 2A,B illustrates the range of these PPV and sensitivity results (arranged by country). There was no clear visual pattern[11] to suggest that any particular country's PPVs or sensitivities performed better than another's. Figure 2C,D illustrates the range of PPVs and sensitivities arranged by overall risks of bias or applicability concern. A larger proportion of PPVs (81% of 70 algorithms) and sensitivities (83% of 46 algorithms) from low-risk studies were of a high magnitude (>80%). By contrast, a smaller proportion of PPVs (12% of 57 algorithms) and sensitivities (52% of 42 algorithms) from high-risk studies were of a high magnitude (>80%). For unclear-risk studies, a small proportion of PPVs (33% of 45 algorithms) and sensitivities (29% of 35 algorithms) were of a high magnitude (>80%). Figure 3A,B illustrates the range of these PPV and sensitivity results grouped by the 19 studies exclusively using medical records as the diagnostic gold standard against the 11 studies that used the other methods. PPV ranged 5.2%–100% and sensitivity ranged 61.0%–100% across the studies using medical records as the diagnostic gold standard. PPV ranged 5.0%–93.0% and sensitivity ranged 36.6%–97.3% across the studies using other methods as the diagnostic gold standard. The lookup table in Appendix S7 ranks all the algorithm results from highest to lowest PPV alongside their sensitivity results (where available), other associated study characteristics, and overall risks of study bias or applicability concern.

Figure 4A,B demonstrates the optimal diagnostic PPV and sensitivity algorithms, respectively, taken from each study and arranged in a forest plot. At their best, PPV or sensitivity were >80% in the majority of studies (19 of 28 studies for PPV and 12 of 16 studies for sensitivity). From visual inspection[11] of these figures, it was difficult to appreciate any potential effect of varying diagnostic gold standards used on optimal PPV and sensitivity estimates because hand-searched medical records predominated the gold standards used across studies. Inspecting the figures[11] otherwise revealed that the high PPV and sensitivity estimates >80% appeared to cross multiple different coding systems and algorithms used (including ICD-8, ICD-9, ICD-10, or Read Codes with and without AEDs or procedures), across various population ages and settings including accident and emergency, inpatients, outpatients, and primary care. Nearly all of the studies with an optimal PPV ≤80% had a high risk of bias or applicability concern (six of nine studies, with unclear risks in the remaining three studies, Figure 4A). The quality ratings were more variable for studies generating optimal sensitivity algorithms (Figure 4B). Studies providing multiple within-study PPVs and sensitivities were used to generate forest plots illustrating the worst-performing diagnostic PPV and sensitivity algorithms, respectively, in Figure 5A,B. These figures illustrate that PPV was low, <80%, in 19 of the 20 worst-performing PPV algorithms, and sensitivity was <80% for nine of the 10 worst-performing sensitivity algorithms. Once again, hand-searched medical records predominated the diagnostic gold standard across this comparison.

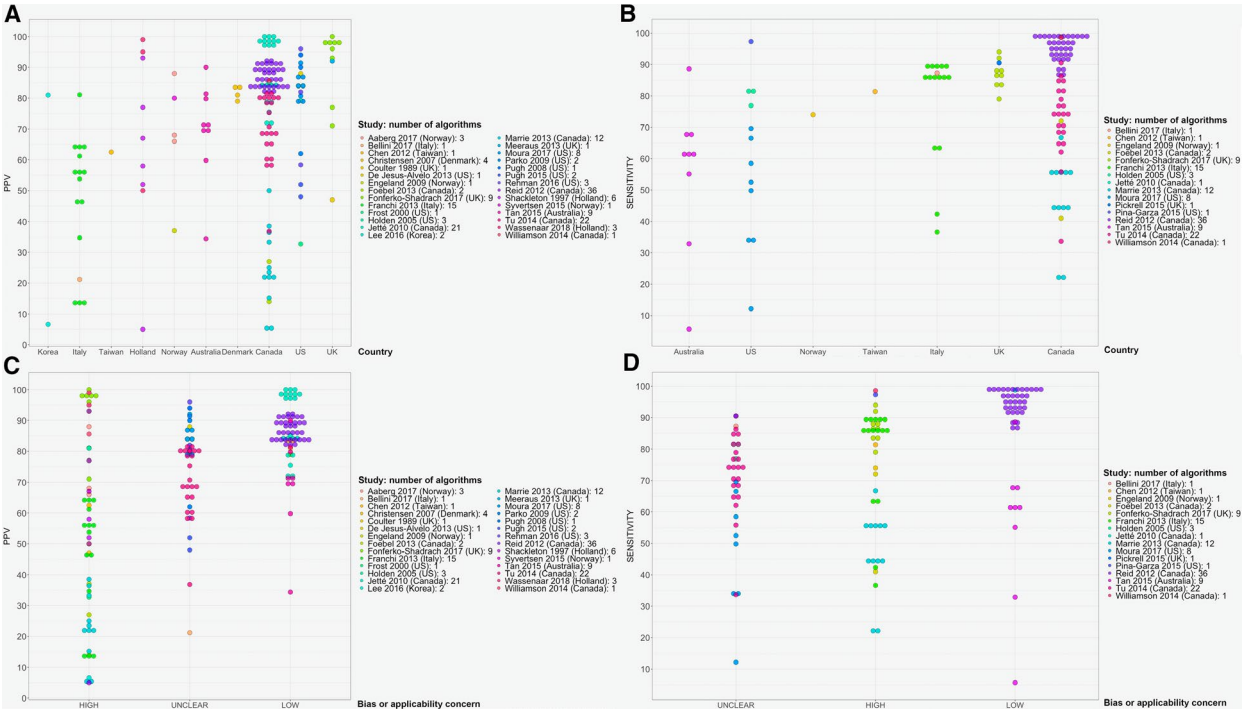


FIGURE 2 Positive predictive value (PPV, %) and sensitivity (%) dot plots: Dots represent the PPV or sensitivity for each algorithm provided by studies, organized by country (2A = PPV, 2B = sensitivity), and by Quality Assessment of Diagnostic Accuracy Studies 2 (QUADAS-2)[19] score for overall risk of study bias or applicability concern (2C = PPV, 2D = sensitivity)

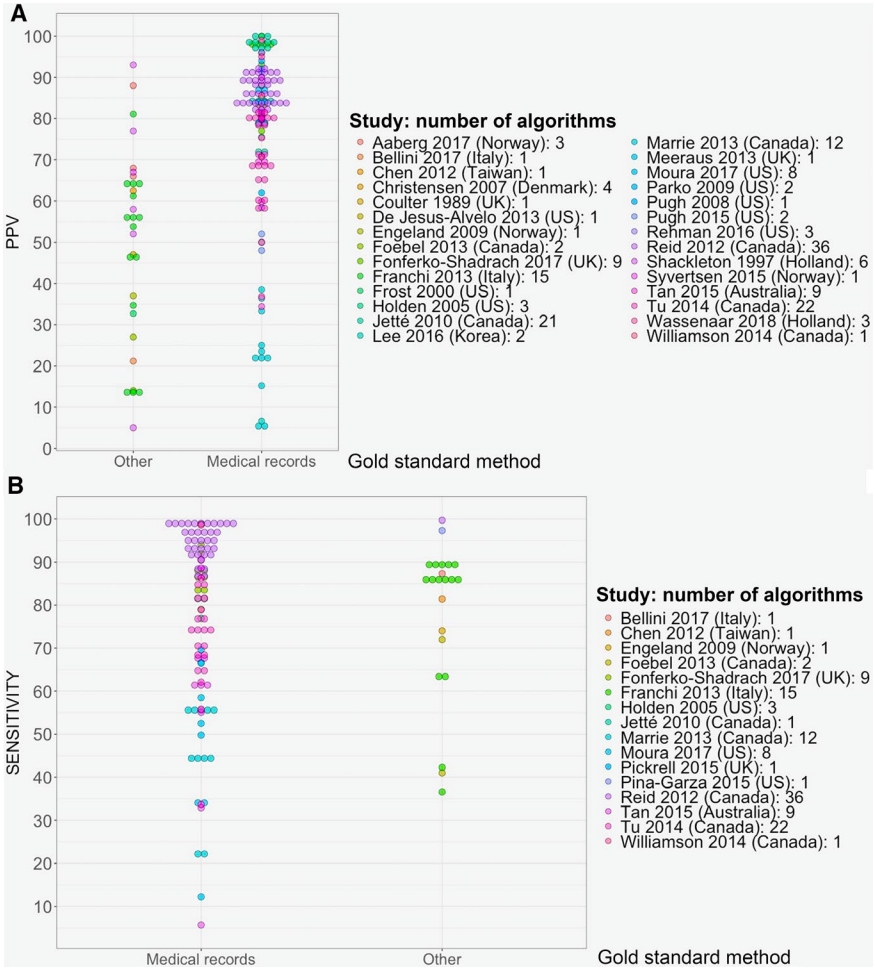


FIGURE 3 Gold standard method dot plots: Dots represent the PPV (3A, %) or sensitivity (3B, %) for each algorithm, grouped by whether the reference standard used was exclusively medical records or other methods (ie, the general practitioner,[35,36,49] antiepileptic drugs,[31,33,44] medical records and/or parent telephone interview,[27] an epilepsy center patient list,[28] a previous epilepsy prevalence study,[29] another administrative data set,[34] or a specialist epilepsy database)[55]

3.4 | PPVs and sensitivities by country

3.4.1 | Canada

Positive predictive value ranged from 5.2%–100% and sensitivity ranged from 22.2%–99.7% across six Canadian studies (Appendix S7).[34,38,40,48,52,54] Jetté 2010[38] and Reid 2012[48] had low risks of bias and applicability concerns (Appendix S6B). All but one study (Foebel 2013[34]) used medical records as their diagnostic gold standard, and PPV and sensitivity range remained 5.2%–100% and 22.2%–99.7%, respectively, across these five studies. Jetté 2010[38] estimated an optimal PPV for epilepsy of 100% (95% CI 93.2%–100%) for ICD-10 codes G40 (epilepsy) and G41 (status epilepticus) combined (Figure 4A). The addition of ICD-10 code R56 (convulsions) lowered this PPV to a worst estimate of 71.6% (CI 60.5%–80.6%, Figure 5A). The sensitivity for G40-41 was 98.8% (CI 93.3%–99.8%) and was not tested with R56. The study also tested the PPVs of ICD-9 coding and found that performance between ICD-9 and ICD-10 was similar, and that there was also little difference between results for children and adults. Reid 2012[48] tested algorithms consisting of ICD-9 code 345 (epilepsy) or ICD-10 G40-41 when used as various combinations of physician claims, hospitalizations, or accident and emergency visits over 1-2 years. Spreading these codes over one physician claim, one hospitalization, or one accident and emergency visit over 2 years had the highest sensitivity (99.7%, CI 99.3%–100%, Figure 4B), demonstrating an ability to capture almost all persons with epilepsy in the population, albeit with the trade-off of also giving the worst PPV of 81.6% (CI 79.0%–84.2%, Figure 5A),[56] indicating that many false positives were also captured by this use of multiple sites. By limiting to those algorithms with two physician claims over 1 year, PPV was raised to an optimal of 92.1% (CI 90.1%–94.1%, Figure 4A) at a trade-off of lowering sensitivity to, at worst, 86.2% (CI 83.3%–90.2%, Figure 5B); suggesting that a good proportion of the epilepsy population was still captured. Williamson 2014[54] validated primary care using ICD-9 codes and had high risks of bias or applicability concerns (Appendix S6B). The algorithm was epilepsy ICD-9 345 or convulsions 780.3 (with a requirement for AEDs in "petit mal" 345.2, "grand mal" 345.3, or in 780.3). This generated a very high sensitivity of 98.6% (CI 96.6%–100%, Figure 4B) with a trade-off of PPV 85.6% (CI 80.2%–91.1%, Figure 4A), perhaps suggesting

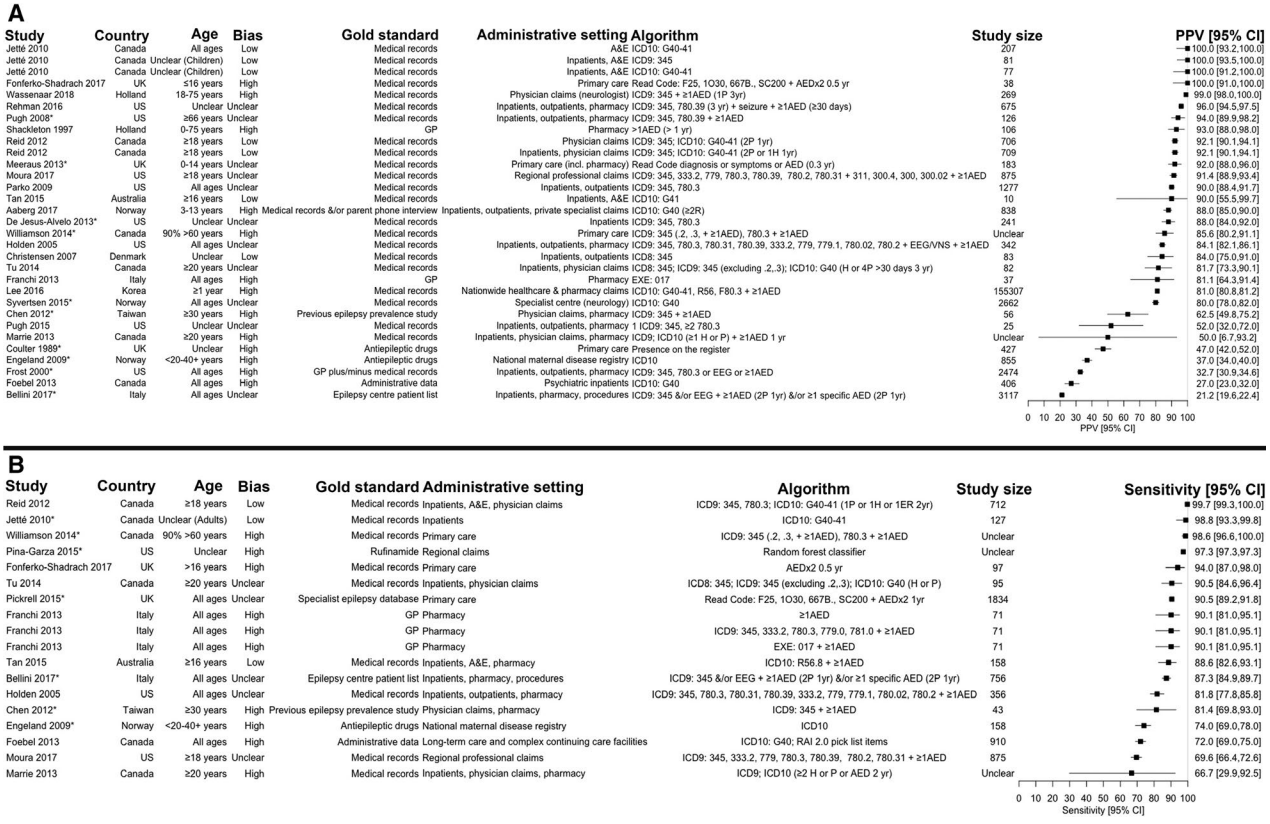


FIGURE 4 Forest plot showing the optimal diagnostic algorithms (4A = highest positive predictive value [PPV, %], 4B = highest sensitivity [%]) in each included study that estimated PPV or sensitivity, alongside various study characteristics. * = Studies that provided one PPV or sensitivity algorithm only. Bias = Quality Assessment of Diagnostic Accuracy Studies 2 (QUADAS-2)[19] score for overall risk of study bias or applicability concern for that study. A and E: Accident and emergency; AED: Antiepileptic drug; CI: Confidence interval (%); EEG: Electroencephalography; EXE: Exception code; F25, 1O30, 667B., SC200: Epilepsy Read Codes; H: Hospitalization; ICD: International Classification of Diseases; ICD10 F80.3: Acquired aphasia with epilepsy; ICD10 G40: Epilepsy; ICD10 G41; Status epilepticus; ICD8 345: Epilepsy; ICD9 311, 300.4, 300, 300.02: Anxiety and depression codes; ICD9 333.2: Myoclonus; ICD9 345: Epilepsy; ICD9 779: Perinatal illness; ICD9 780.2: Syncope and collapse; ICD9 780.3: Convulsions; ICD9 780.31: Febrile convulsions; ICD9 780.39: convulsions; P: Physician claim; RAI: Resident Assessment Instrument; VNS: Vagus nerve stimulation; Yr: year

that the convulsions code in ICD-9 lowers PPV similarly to the ICD-10 convulsions code R56, as was seen in Jetté 2010.[38] Other studies included the Tu 2014[52] study, which had an unclear risk of bias and tested an algorithm consisting of ICD-8 345 (epilepsy), ICD-9 345, or ICD-10 G40 as hospitalization or physician claims over varying times. The optimal PPV of 81.7% (CI 73.3%–90.1%, Figure 4A) was generated by one hospitalization or four physician claims separated by ≥30 days within a 3-year period and greatly contrasted with the worst PPV of 36.8% (CI 30.6%–42.9%, Figure 5A) generated by only one hospitalization or physician claim. However, this poor PPV was associated with a trade-up in sensitivity to 90.5% (CI 84.6%–96.4%, Figure 4B). The worst sensitivity consisted of hospitalization alone (33.7%, CI 24.2%–43.2%) indicating that hospitalization alone is not a reliable way of capturing all persons with epilepsy. Another study, by Marrie 2013,[40] validated epilepsy codes as a comorbidity of multiple sclerosis and had high risks of bias and low PPV and sensitivity estimates (with wide CIs, Figures 4 and 5, Appendix S7). These results were difficult to interpret in the context of the different algorithms provided. Foebel 2013[34] had high risks of bias and validated ICD-10 G40 or Resident Assessment Instrument (RAI) 2.0 pick list items within psychiatric inpatients, long-term care, and complex continuing care, generating low PPV and sensitivity estimates, which suggests that these environments and/or pick list items brought inaccuracy to estimates (Figures 4 and 5). It is possible that the study's use of an administrative data set as the diagnostic gold standard compounded the poorer accuracy estimates.

3.4.2 | United States

Positive predictive value and sensitivity ranged from 32.7%–96.0% and 12.2%–97.3%, respectively, across nine US studies (Appendix S7).[32,36,37,42-47] Risks of bias and applicability concern were unclear in all but two studies,[36,44] where risks were

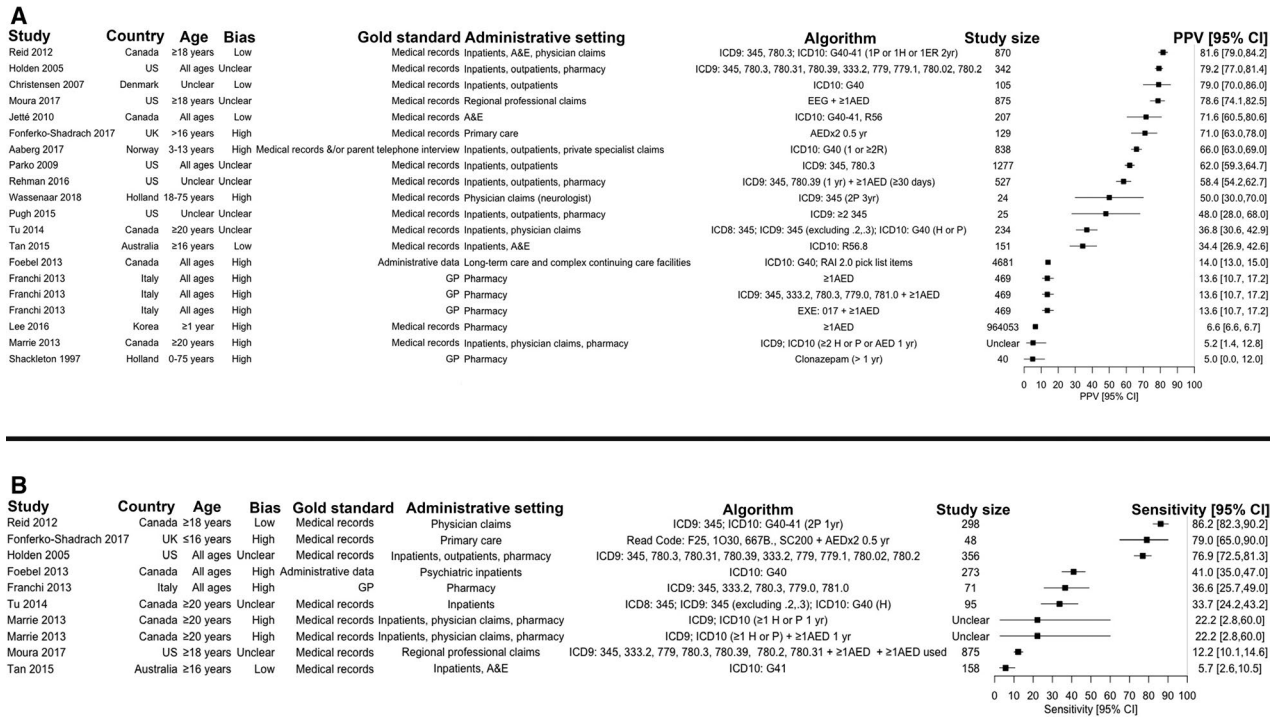


FIGURE 5 Forest plot showing the worst-performing diagnostic algorithms (5A = lowest positive predictive value [PPV, %], 5B = lowest sensitivity [%]) in each included study that estimated PPV or sensitivity, alongside various study characteristics. Bias = Quality Assessment of Diagnostic Accuracy Studies 2 (QUADAS-2)[19] score for overall risk of study bias or applicability concern for that study. A and E: Accident and emergency; AED: Antiepileptic drug; CI: Confidence interval (%); EEG: Electroencephalography; EXE: Exception code; F25, 1O30, 667B., SC200: Epilepsy Read Codes; H: Hospitalization; ICD: International Classification of Diseases; ICD10 F80.3: Acquired aphasia with epilepsy; ICD10 G40: Epilepsy; ICD10 G41; Status epilepticus; ICD8 345: Epilepsy; ICD9 311, 300.4, 300, 300.02: Anxiety and depression codes; ICD9 333.2: Myoclonus; ICD9 345: Epilepsy; ICD9 779: Perinatal illness; ICD9 780.2: Syncope and collapse; ICD9 780.3: Convulsions; ICD9 780.31: Febrile convulsions; ICD9 780.39: Convulsions; P: Physician claim; RAI: Resident Assessment Instrument; VNS: vagus nerve stimulation; Yr: year

high (Appendix S6B). Six US studies used medical records as their diagnostic gold standard (PPV range 48.0%–96.0%, sensitivity range 12.2%–81.8%).[37,42,43,45-47] The US studies all tested ICD-9 codes, including at least ICD-9 345 (epilepsy) and either 780.3 or 780.39 (seizures). When these diagnostic codes were looked at without a requirement for AEDs, PPV ranged between 48.0% and 90.0% (seven algorithms from five studies),[32,37,42,43,45] with PPV <80% in four of the seven algorithms. When diagnostic codes were looked at with a requirement for AEDs, PPV ranged between 58.4% and 94.0% (nine algorithms from four studies).[37,42,46,47] PPV was <80% in only two of nine algorithms, indicating that an AED requirement increased the likelihood of higher PPV estimates. When diagnostic codes were looked at with a requirement for AEDs and procedures (EEG or VNS), PPV ranged between 32.7% and 86.9% (four algorithms from three studies,[36,37,42] with two algorithm PPVs <80%). The best and worst algorithm combinations from the United States are shown in Figures 4 and 5. In the Parko 2009[43] study, the algorithm ICD-9 345 or 780.3 had a PPV of 90.0% (CI 88.4%–91.7%) when capturing epilepsy or seizures (Figure 4A), and fell to 62.0% (CI 59.3-64.7) when capturing epilepsy. ICD-9 345 without 780.3 was not tested.

3.4.3 | United Kingdom

Positive predictive value and sensitivity ranged from 47.0%–100% and 79.0%–94.0%, respectively, across four UK studies (Appendix S7).[14,31,41,55] Two studies (Fonferko-Shadrach 2017[14] and Meeraus 2013[41]) used medical records as their diagnostic gold standard (PPV range 71.0%–100%, sensitivity range 79.0%–94.0%). The Fonferko-Shadrach 2017[14] study had a high risk of bias or applicability concern (Appendix S6B). This study demonstrated that within Read Codes, two prescriptions of the same AED within 6 months provided the optimal sensitivity (94.0%, CI 87.0%–98%, Figure 4B), capturing nearly all persons with epilepsy. However, many false positives were also captured, generating the lowest PPV (71.0%, CI 63.0%–78.0%, Figure 5A). This was in participants >16 years of age and rose to 98% (CI 94%–100%) in participants ≤16 years of age. AEDs are seldom prescribed for indications other than epilepsy in children in the UK.[14] This was the only included study to allow direct comparison of AEDs alone between adults and children. Diagnostic epilepsy Read Codes combined with two AED prescriptions in 6 months generated

the highest PPV (100%, CI 91.0%–100%, Figure 4A), at a trade-off of slightly poorer sensitivity (79.0%, CI 65.0%–90.0%, Figure 5B). PPVs and sensitivities for diagnostic Read Codes alone ranged between 93.0% and 98.0% and 83.0% and 88.0%, respectively. Pickrell 2015[55] had unclear risks of bias (Appendix S6B) and demonstrated a similar sensitivity result (90.5%, CI 89.2%–91.8%) to Fonferko-Shadrach 2017[14] by also using diagnostic epilepsy Read Codes combined with two AED prescriptions, although over 12 months. Pickrell 2015[55] used a specialist epilepsy database as the diagnostic gold standard. Meeraus 2013[41] had unclear risks of bias and demonstrated high PPVs for Read Code epilepsy diagnosis, symptoms, or AED prescription repeated within 4 months (92.0%, CI 88.0%–96.0%). Coulter 1989,[31] the oldest included study, had high risks of bias. The chronic primary care disease register validated was found to have a poor PPV (47.0%, CI 42.0%–52.0%). It is possible that this was compounded by the study's use of AEDs as a diagnostic gold standard.

3.4.4 | Norway

Positive predictive value ranged from 37.0%–88.0% across three Norwegian studies (Appendix S7).[27,33,50] Risks of bias were high in all but one study, which had unclear risks (Appendix S6B).[50] This same study (Syvertsen 2015[50]) was the only one to exclusively use medical records as the diagnostic gold standard (PPV 80%). Both Syvertsen 2015[50] and Aaberg 2017[27] validated ICD-10 code G40 alone (two or more registrations in Aaberg 2017)[27] and achieved good optimal PPVs ranging from 80%–88% (Figure 4A). Reducing to one or more registrations lowered PPV to 66.0% (CI 63.0%–69.0%, Figure 5A).[27] Aaberg 2017[27] used medical records and/or parent telephone interview as the diagnostic gold standard. Engeland 2009[33] validated a maternal disease registry to capture epilepsy in pregnant women using ICD-10 (codes were not further specified), identifying globally poor estimates of PPV 37.0% (CI 34.0%–40.0%, Figure 4A) and sensitivity 74.0% (CI 69.0%–78.0%, Figure 4B). It is possible that this was compounded by the study's use of AEDs as the diagnostic gold standard.

3.4.5 | Italy

Positive predictive value ranged from 13.6%–81.1% and sensitivity ranged from 36.6%–90.1% across two Italian studies (Appendix S7), one using the GP[35] and the other an epilepsy center patient list[28] as the diagnostic gold standard. Risks of bias were high in Franchi 2013[35] and unclear in Bellini 2017[28] (Appendix S6B). The highest PPV came from the former study's use of lone exemption codes (EXE) for epilepsy 017 (based on ICD-9): 81.1% (CI 64.3%–91.4%, Figure 4A).[35] These are codes that qualify people with epilepsy for free-of-charge healthcare. However, sensitivity was low (42.3%, CI 30.8%–54.5%) indicating that these were not able to capture all persons with epilepsy. The addition of AEDs to exemption codes or to diagnostic ICD-9 codes (or use of AEDs alone) dropped PPV to its worst (13.6%, CI 10.7%–17.2%, Figure 5A). Correspondingly, sensitivity was highest from use of the diagnostic ICD-9 or exemption codes with an additional requirement for AEDs or from AEDs alone (90.1%, CI 81.0%–95.1%, Figure 4B).[35] Bellini 2017[28] used a complex case-ascertainment algorithm arranged as follows: [≥1 ICD-9 345 code] and/or [≥1 EEG recorded in a person prescribed ≥2 AEDs over 12 months] and/or [≥2 prescriptions of one epilepsy-specific AED over 12 months]. This algorithm had a good sensitivity (87.3%, CI 84.9%–89.7%, Figure 4B) at a trade-off of generating a particularly low PPV (21.2%, CI 19.6%–22.4%, Figure 4A), suggesting that despite capturing most epilepsy cases there were many additional false positives. The PPV and sensitivity ranges across nine algorithms that included codes for procedures (EEG) were 21.2%–64.2% and 85.9%–88.7%, respectively.[28,35] PPV and sensitivity ranges across seven algorithms without procedures were 13.6%–81.1% and 36.6%–90.1%, respectively, suggesting that there was a more stable tendency for good sensitivity when procedures were included but PPVs were variable (Appendix S7).

3.4.6 | The Netherlands

Positive predictive value ranged from 5.0%–99.0% across two Dutch studies (Appendix S7), one using the GP[49] and the other medical records[53] as the diagnostic gold standard. Risks of bias or applicability concern were high for both studies (Appendix S6B). AED monotherapy or ICD-9 coding without AEDs tended to perform poorly, with PPV range 5.0%–77.0% for the former notwithstanding requirement for continuous prescription over more than 1 year, and PPV 50.0 (CI 30.0%–70.0%) for the latter notwithstanding requirement for two attendances over 3 years. Polytherapy (>1 AED) improved PPV to 93.0% (CI 88.0%–98.0%), and adding in ICD-9 codes to the AED algorithm improved PPV to 95.0%–99.0%.

3.4.7 | Other countries

Little additional information was gained from the findings of countries within which only one study was conducted. These studies are summarized in Appendix S8 (Denmark,[30] Australia,[51] South Korea,[39] Taiwan).[29] In the South Korean study (gold standard medical records), one or more AEDs

alone generated a worst PPV of 6.6% (CI 6.6%–6.7%, Figure 5A), reinforcing the potential problems of using AEDs alone.[39]

3.5 | Overall NPV and specificity

Less than half of the studies provided data on NPV and specificity: 14 studies, generating 131 algorithms estimating NPV (range 13.2%–100%), with 121 of these algorithms also estimating specificity (range 61.0%–100%).[14,28,29,33-35,37,38,40,42,48,51,52,54] One study provided specificity (95.6%) without NPV.[44] Figures 6 and 7 illustrate the range of these NPV and specificity results, arranged by country (Figure 6A,B), study quality score (Figure 6C,D), and gold standard method (Figure 7A,B). Nearly all of the NPV estimates (93% of 131 algorithms) and specificity estimates (95% of 122 algorithms) were high, >80%, with little visible influence seen from grouping by country or study quality (Figure 6). Grouping by gold standard method suggested that the highest NPV and specificity estimates were found in the studies using methods other than medical records as the gold standard: All 20 algorithms from five studies generated NPVs >97% and all 21 algorithms from six studies generated specificities >94% (Figure 7). Figure 8A,B demonstrates the optimal NPV and specificity algorithms taken from each study and arranged in a forest plot. In each figure, all but one algorithm generated a very high NPV or specificity (>90%). This holds true despite multiple different coding systems and algorithms used, across various population ages and settings, and between different countries. Most studies used medical records as their diagnostic gold standard. The worst-performing NPV and specificity algorithms (Figure 9A,B) demonstrated similar trends, with nearly all achieving stable estimates >80%–90% despite varying study characteristics. NPV and specificity estimates are also included in the Appendix S7 look-up table.

3.6 | Meta-analysis of sensitivity and specificity

Only 10 of the 30 included studies provided true positives, false positives, true negatives, and false negatives for the algorithms being tested (91 algorithms, Appendix S9A), allowing limited opportunity for comprehensive meta-analysis. The bivariate diagnostic random-effects meta-analysis revealed an overall sensitivity estimate of 88.6% (CI 85.7%–91.0%), with an overall specificity estimate of 97.9% (CI 98.5%–96.9%) false positive rate 0.021 (CI 0.015–0.031, Appendix S9B). This matches closely with our narrative assessment that in general, sensitivity, and specificity estimates were high across studies. We were unable to proceed to a

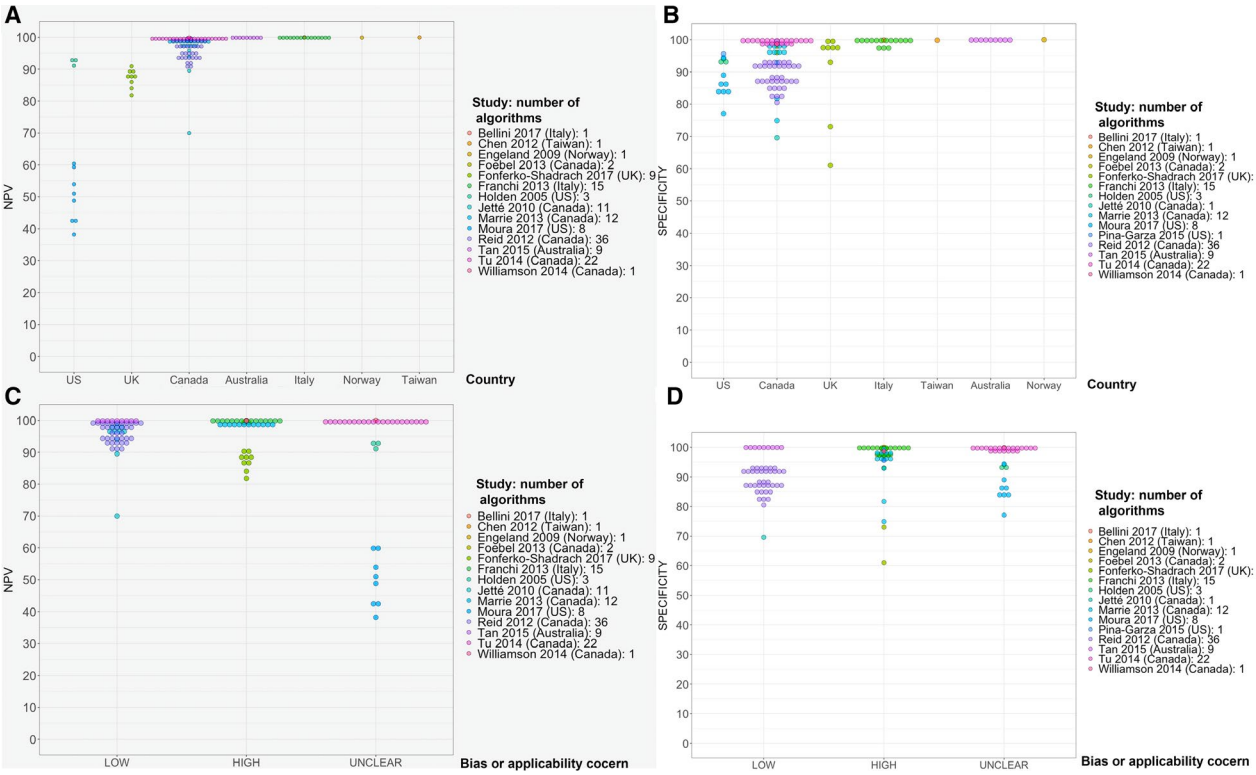


FIGURE 6 Negative predictive value (NPV, %) and specificity (%) dot plots: Dots represent the NPV or specificity for each algorithm provided by studies, organized by country (6A = NPV, 6B = specificity), and by Quality Assessment of Diagnostic Accuracy Studies 2 (QUADAS-2)[19] score for overall risk of study bias or applicability concern (6C = NPV, 6D = specificity)

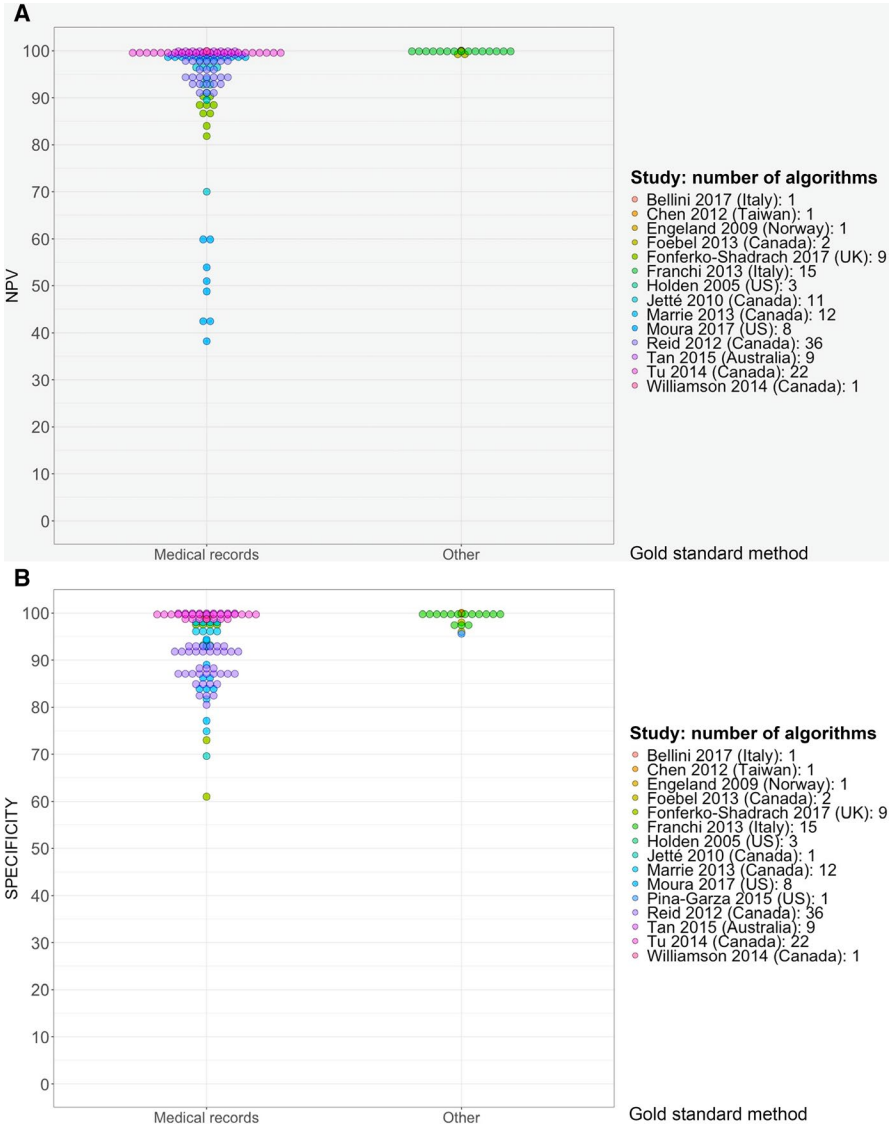


FIGURE 7 Gold standard method dot plots: Dots represent the NPV (7A, %) or specificity (7B, %) for each algorithm, grouped by whether the reference standard used was exclusively medical records or other methods

valid bivariate meta-regression using study characteristics as covariates because the full spectrum of available study characteristics from the 30 included studies were not represented in a small sample of only 10 studies. Furthermore, such a meta-regression would assume each of the within-study algorithms providing estimates were independent, but they are not.

4 | DISCUSSION

4.1 | Summary of findings

Overall, this systematic review illustrates that it is reasonable to use administrative healthcare data to identify people with epilepsy in epidemiological research, with studies tending to achieve high estimates (>80%) of PPV, sensitivity, NPV, and specificity. Because the large majority of validation studies in epilepsy have used hand-searched medical records as the diagnostic gold standard, limiting analysis to only such studies makes little difference to the overall results. Little difference is also seen in optimal epilepsy case-identification accuracy between the different ICD coding versions and between the ICD system and others. However, for a particular disease-coding system to perform well in and of itself, the most important elements seem to be algorithm structure and composition. There is a known trade-off relationship between PPV and sensitivity.[56] In situations where researchers wish to prioritize achieving a high PPV, our findings suggest that the algorithm should consist of disease codes (ICD-10 G40-41, ICD-9 345) without symptom codes (ICD-10 R56, ICD-9 780.3, 780.39), and one or more AEDs should be included. Most optimal case-identification algorithms capturing PPVs >80% (and even above 90%) used this basic arrangement, with or without the need to have the disease code registered more than once (Figure 4A). Where a balance needs to be struck between PPV and sensitivity, researchers may wish to choose disease codes alone (ie, without AEDs). This managed

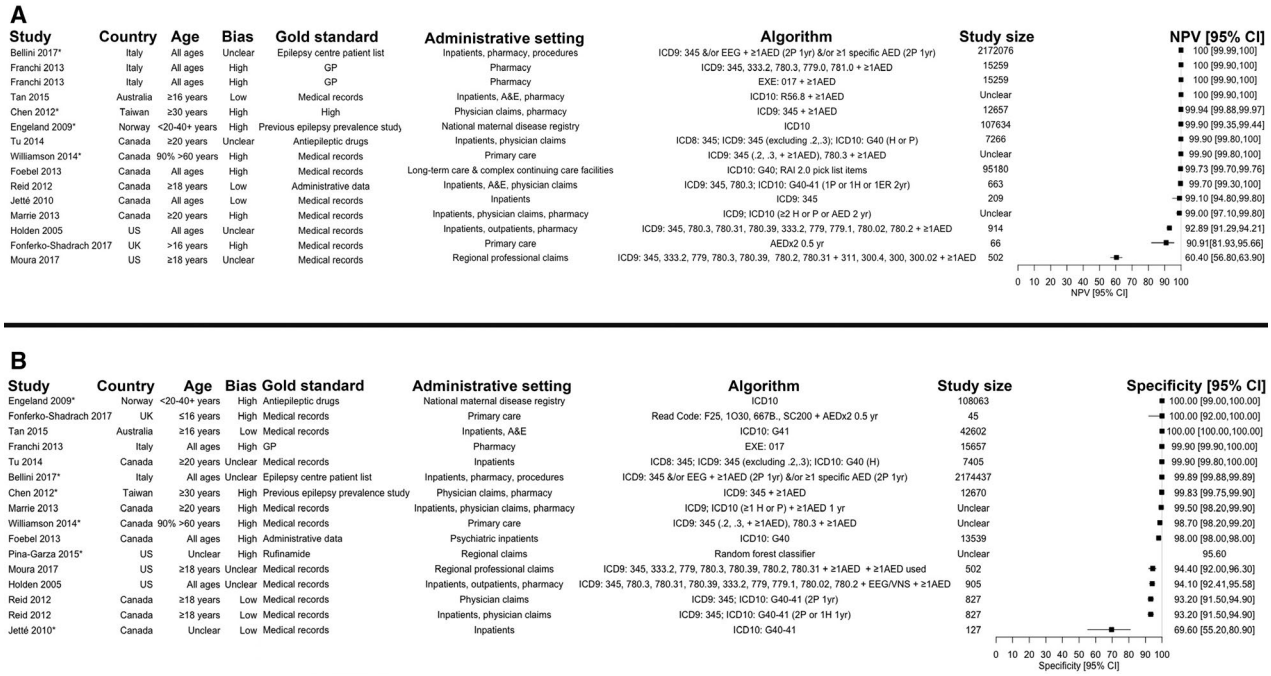


FIGURE 8 Forest plot showing the optimal diagnostic algorithms (8A = highest negative predictive value [NPV, %], 8B = highest specificity [%]) in each included study that estimated NPV or specificity, alongside various study characteristics. * = Studies that provided one NPV or specificity algorithm only. Bias = Quality Assessment of Diagnostic Accuracy Studies 2 (QUADAS-2)[19] score for overall risk of study bias or applicability concern for that study. A and E: Accident and emergency; AED: Antiepileptic drug; CI: Confidence interval (%); EEG: Electroencephalography; EXE: Exception code; F25, 1O30, 667B., SC200: Epilepsy Read Codes; H: Hospitalization; ICD: International Classification of Diseases; ICD10 F80.3: Acquired aphasia with epilepsy; ICD10 G40: Epilepsy; ICD10 G41; Status epilepticus; ICD8 345: Epilepsy; ICD9 311, 300.4, 300, 300.02: Anxiety and depression codes; ICD9 333.2: Myoclonus; ICD9 345: Epilepsy; ICD9 779: Perinatal illness; ICD9 780.2: Syncope and collapse; ICD9 780.3: Convulsions; ICD9 780.31: Febrile convulsions; ICD9 780.39: Convulsions; P: Physician claim; RAI: Resident Assessment Instrument; VNS: Vagus nerve stimulation; Yr: year

to retain both PPV and sensitivity >80% in several studies (Figure 4A,B). When sensitivity is the priority, our findings suggest that the algorithm should consist of symptom codes with one or more AEDs. We consistently show that an algorithm consisting of one or more AEDs alone, while often able to provide a high sensitivity, is associated with very poor PPVs and should, therefore, not be used. High PPV may be preserved when one or more AEDs alone are used to try and identify children,[14] although this was reported in one study and requires confirmation. For both PPV and sensitivity, there seems to be no added value in adding peripheral epilepsy codes including: ICD-9 333.2 (myoclonus), 779 (convulsions in new-borns), 780 (alteration of consciousness) except 780.3 or 0.39, or ICD-10 F80.3 (acquired aphasia with epilepsy). The algorithms consisting of long combinations of core and peripheral epilepsy codes were no more likely to achieve a higher PPV or sensitivity for epilepsy than those using just core epilepsy disease codes (ICD-10 G40-41, ICD-9 345, Figures 4 and 5, Appendix S7). Although adding in a procedure code (such as EEG) to core epilepsy disease codes optimized sensitivity, it was associated with an unnecessary drop in PPV, and therefore it is unlikely to be helpful. The ICD system is largely used in administrative hospital settings worldwide, and one of its main purposes is to help summarize the incidence and prevalence of diseases on a national or worldwide basis.[57] On the other hand, Read Codes are used primarily to create electronic patient records in the UK primary care system (although they can also be used to summarize disease incidence and prevalence).[57] Taking these differences in function into account, our finding that there was little difference in the diagnostic accuracy of ICD codes and Read Codes for identification of epilepsy cases in administrative healthcare records should be interpreted with caution, particularly as only three studies used Read Codes. Overall, epilepsy coding was particularly good at correctly identifying those without the disease, with estimates of NPV and specificity mostly remaining >90%, despite varying study characteristics.

Our other findings are that no countries appeared to outperform others in terms of overall case-identification accuracy using administrative data. Second, the factors seeming to be most consistently associated with lower optimal PPV or sensitivity estimates globally were modifiable, namely, study design elements that bring about high risk of bias or applicability concern. To lower these risks, future validation studies should aim to clarify/adequately blind the reference standard

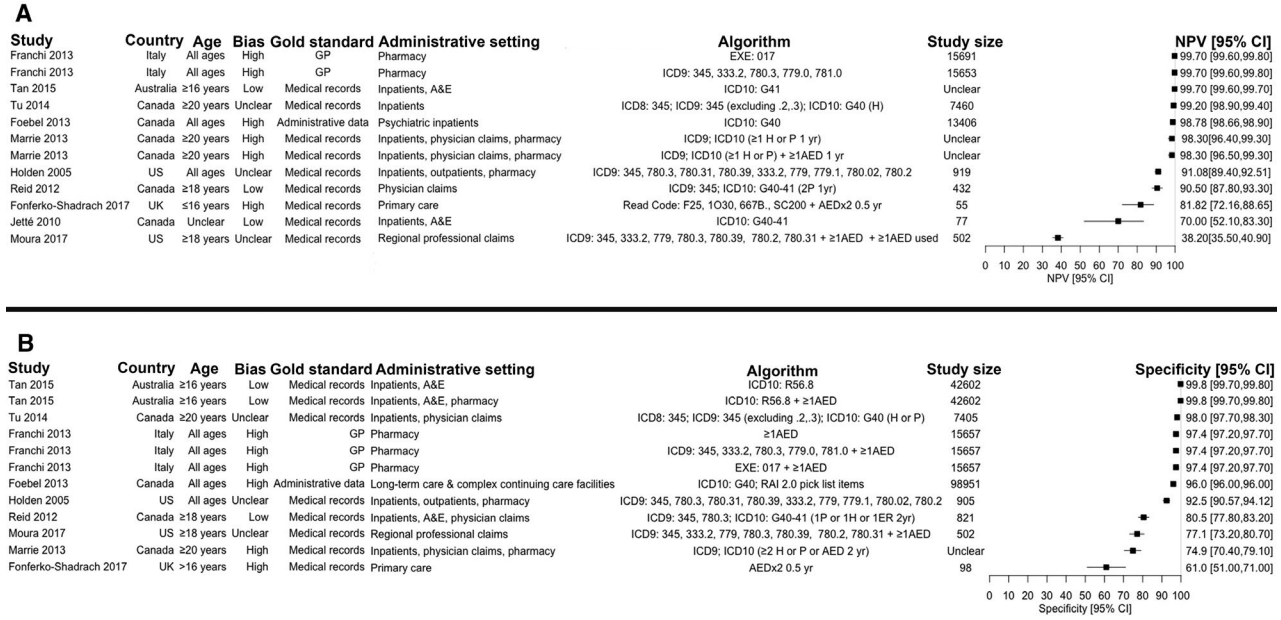


FIGURE 9 Forest plot showing the worst-performing diagnostic algorithms (9A = lowest negative predictive value [NPV, %], 9B = lowest specificity [%]) in each included study that estimated NPV or specificity, alongside various study characteristics. Bias = Quality Assessment of Diagnostic Accuracy Studies 2 (QUADAS-2)[19] score for overall risk of study bias or applicability concern for that study. A and E: Accident and emergency; AED: Antiepileptic drug; CI: Confidence interval (%); EEG: Electroencephalography; EXE: Exception code; F25, 1O30, 667B., SC200: Epilepsy Read Codes; H: Hospitalization; ICD: International Classification of Diseases; ICD10 F80.3: Acquired aphasia with epilepsy; ICD10 G40: Epilepsy; ICD10 G41; Status epilepticus; ICD8 345: Epilepsy; ICD9 311, 300.4, 300, 300.02: Anxiety and depression codes; ICD9 333.2: Myoclonus; ICD9 345: Epilepsy; ICD9 779: Perinatal illness; ICD9 780.2: Syncope and collapse; ICD9 780.3: Convulsions; ICD9 780.31: Febrile convulsions; ICD9 780.39: Convulsions; P: Physician claim; RAI: Resident Assessment Instrument; VNS: Vagus nerve stimulation; Yr: year

reader, avoid using unvalidated administrative data sets or AEDs as reference standards, administer the same reference standard to all participants, and account for any missing data by performing best- and worst-case sensitivity analyses. A third observation is that administrative settings that specialize in non-neurology and non-general medicine or accident and emergency patients appeared to perform worst in terms of epilepsy PPVs and sensitivities. These were psychiatric inpatients, maternal disease registries, and long-term and complex continuing care facilities. Optimal PPVs were no more than 37% in these settings, and sensitivities were no more than 74% (Figure 4A,B). However, as these studies were also part of the small handful of studies to not use medical records as the diagnostic gold standard (instead using AEDs or other administrative data sets as gold standard), the predominant cause of the lower accuracy estimates becomes less clear and will require further investigation in future studies using gold standard medical records. Finally, the data indicate that administrative setting and age of participants appeared to play little influence on optimal epilepsy case-identification accuracy.

Placing our results in context, the accuracy of administrative healthcare data in identifying epilepsy appears to be similar to that of other neurological diseases with rapidly progressive physical symptoms, such as multiple sclerosis (PPV 86%, sensitivity 84%, NPV 99%, specificity 100%)[58] or motor neuron disease (PPV 55%–92%, sensitivity 75%–93%, NPV/specificity not reported).[10]

4.2 | Strengths and limitations

This review expands on findings of the 2012 US/Canada systematic review of this topic[12] by now taking advantage of algorithms tested after that review within the United States or Canada (10 studies),[32,34,40,42,44,45,47,48,52,54] and algorithms tested outside of the United States and Canada (15 studies).[14,27-31,33,35,39,41,49-51,53,55] We also capture 10 more than the single study validating ICD-10,[27,30,33,34,39,40,48,50-52] and we review other non-ICD coding systems for the first time.[14,31,41,49,55] It was concluded previously that a corrective factor would be necessary when relying on ICD-9 to identify epilepsy because the US Parko 2009[43] study had shown a PPV of 90% for a diagnosis of epilepsy or seizures and yet only a PPV of 62% for a diagnosis of epilepsy. However, we now demonstrate several other studies relying on ICD-9 in which no corrective factor was necessary to accurately identify epilepsy (see in Figure 4A, for example, the US study of De Jesus-Alvelo 2013).[32] We also demonstrate this for ICD-10 (eg, Tan 2015[51] in Australia). It was also concluded

previously that multiple epilepsy diagnoses over time are required to identify epilepsy using PPV. Although this is intuitive and was a condition for many of the optimal algorithms with PPVs >80%, our much larger sample of studies demonstrates that it was not a requirement for all (Figure 4A, Appendix S7). This is the first systematic review of epilepsy validation studies to assess included studies for bias and applicability concern. We have been able to demonstrate that studies with a high risk of bias or applicability concern have a tendency toward lower PPV estimates, and we make recommendations on how to reduce these risks. This is also the first systematic review of epilepsy validation studies to incorporate NPV and specificity estimates, demonstrating that these estimates are generally favorable for this condition. Our algorithm look-up table (Appendix S7) is a novel resource that should help future researchers screen epilepsy patients for inclusion into studies recruiting from administrative data sets.

This review demonstrates that there is currently limited scope for detailed meta-analysis of epilepsy validation data, indicating that this an area that will require more investigation in the future as more data become available. The forest plots (Figures 4, 5, 8, 9, Appendix S9) suggest that heterogeneity is low, but this would benefit from further investigation in a meta-regression, which was not possible in the current review because of limitations in the available data. Furthermore, we were unable to assess for publication bias due to the absence of a well-established technique to do so in validation reviews.[10,11]

5 | CONCLUSIONS

This review helps researchers in deciding on optimal case-identification methods when using administrative healthcare data to capture persons with epilepsy. It also improves our understanding of the likely accuracy of global estimates for the incidence and prevalence of epilepsy, which have largely been made using administrative data. Our findings are consistent with the hypothesis that administrative epilepsy data accurately identify epilepsy cases, with optimal estimates of PPV, sensitivity, NPV, and specificity >80% in the majority of studies. This is likely to be sufficient for most epidemiological studies, although we show that careful consideration is needed when choosing algorithm structure and code composition for PPV and sensitivity. Although the published algorithms are a useful approximation for accuracy and there are several epilepsy validation studies available around the world, future investigators should still aim to validate putatively selected algorithms within their own data sets first in order to help maintain transparency in the likely accuracy of subsequent study findings.

ACKNOWLEDGMENTS

This work was charitably supported by Epilepsy Research UK (R44007) and the Juliet Bergqvist Memorial Fund. The funders played no role in the design or conduct of this review. We are also grateful to Kathryn Bush and Tim Wilkinson for support with RStudio, and Sarah Nevitt, Cat Graham, and Jane Hutton for statistical support. We are grateful to Ruth Ann Marrie for providing unpublished study data. We are also grateful to Siddharthan Chandran and Cathie Sudlow for their strategic support.

CONFLICTS OF INTEREST

None of the authors has any conflict of interest to disclose.

ETHICAL PUBLICATION STATEMENT

We confirm that we have read the Journal's position on issues involved in ethical publication and affirm that this report is consistent with those guidelines.

ORCID

Gashirai K. Mbizvo https://orcid.org/0000-0002-9588-2944
Kyle H. Bennett https://orcid.org/0000-0002-3922-7056
Colin R. Simpson https://orcid.org/0000-0002-5194-8083
Susan E. Duncan https://orcid.org/0000-0003-0454-6564
Richard F.M. Chin https://orcid.org/0000-0002-7256-3027

REFERENCES

1. Garratt E, Barnes H, Dibben C. Health administrative data: exploring the potential for academic research. St Andrews, UK: Administrative Data Liaison Service; 2010.
2. Loke YK. Use of databases for clinical research. Arch Dis Child. 2014;99(6):587–9.
3. St Germaine-Smith C, Metcalfe A, Pringsheim T, Roberts JI, Beck CA, Hemmelgarn BR, et al. Recommendations for optimal ICD codes to study neurologic conditions: a systematic review. Neurology. 2012;79(10):1049–55.
4. Abraha I, Montedori A, Eusebi P, Orso M, Grilli P, Cozzolino F, et al. The current state of validation of administrative healthcare databases in Italy: a systematic review. Pharmacoepidemiol Drug Saf. 2012;21:400.
5. Traversa G, Bianchi C, Da Cas R, Abraha I, Menniti-Ippolito F, Venegoni M. Cohort study of hepatotoxicity associated with nimesulide and other non-steroidal anti-inflammatory drugs. BMJ. 2003;327(7405):18–22.
6. Hottes TS, Skowronski DM, Hiebert B, Janjua NZ, Roos LL, Van Caeseele P, et al. Influenza vaccine effectiveness in the elderly based on administrative databases: change in immunization habit as a marker for bias. PLoS ONE. 2011;6(7):e22618.
7. Lalmohamed A, Vestergaard P, Cooper C, de Boer A, Leufkens HGM, van Staa TP, et al. Timing of stroke in patients undergoing total hip replacement and matched controls. Stroke. 2012;43(12):3225–9.

8. Schneeweiss S, Avorn J. A review of uses of health care utilization databases for epidemiologic research on therapeutics. J Clin Epidemiol. 2005;58(4):323–37.
9. Krarup LH, Boysen G, Janjua H, Prescott E, Truelsen T. Validity of stroke diagnoses in a national register of patients. Neuroepidemiology. 2007;28(3):150–4.
10. Horrocks S, Wilkinson T, Schnier C, Ly A, Woodfield R, Rannikmäe K, et al. Accuracy of routinely-collected healthcare data for identifying motor neurone disease cases: a systematic review. PLoS ONE. 2017;12(2):e0172639.
11. Wilkinson T, Ly A, Schnier C, Rannikmäe K, Bush K, Brayne C, et al. Identifying dementia cases with routinely collected health data: a systematic review. Alzheimers Dement. 2018;14(8):1038–51.
12. Kee VR, Gilchrist B, Granner MA, Sarrazin NR, Carnahan RM, et al. A systematic review of validated methods for identifying seizures, convulsions, or epilepsy using administrative and claims data. Pharmacoepidemiol Drug Saf. 2012;21:183–93.
13. Benson T. The history of the Read Codes: the inaugural James Read Memorial Lecture 2011. Inform Prim Care. 2011;19(3):173–82.
14. Fonferko-Shadrach B, Lacey AS, White CP, Powell HWR, Sawhney IMS, Lyons RA, et al. Validating epilepsy diagnoses in routinely collected data. Seizure. 2017;52:195–8.
15. Shea BJ, Grimshaw JM, Wells GA, Boers M, Andersson N, Hamel C, et al. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. BMC Med Res Methodol. 2007;15(7):10.
16. Mbizvo GK, Bennett K, Simpson CR, Duncan SE, Chin RFM. Accuracy and utility of using administrative healthcare databases to identify people with epilepsy: a protocol for a systematic review and meta-analysis. BMJ Open. 2018;8(6):e020824.
17. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. BMJ. 2009;21(339):b2535.
18. World Health Organization. Classification of Diseases (ICD). 2016. Available from www.who.int/classifications/icd/en/
19. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med. 2011;155(8):529–36.
20. Shea BJ, Bouter LM, Peterson J, Boers M, Andersson N, Ortiz Z, et al. External validation of a measurement tool to assess systematic reviews (AMSTAR). PLoS ONE. 2007;2(12):e1350.
21. Shea BJ, Hamel C, Wells GA, Bouter LM, Kristjansson E, Grimshaw J, et al. AMSTAR is a reliable and valid measurement tool to assess the methodological quality of systematic reviews. J Clin Epidemiol. 2009;62(10):1013–20.
22. Viera AJ, Garrett JM. Understanding interobserver agreement: the kappa statistic. Fam Med. 2005;37(5):360–3.
23. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977;33(1):159–74.
24. Deeks JJ, Bossuyt PM, Gatsonis C. Handbook for Systematic Reviews of Diagnostic Test Accuracy Version 1.02010. [updated 2019 May 19]. Available from https://methods.cochrane.org/sdt/handbook-dta-reviews
25. Reitsma JB, Glas AS, Rutjes AW, Scholten RJPM, Bossuyt PM, Zwinderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. J Clin Epidemiol. 2005;58(10):982–90.
26. Doebler P, Holling H. Meta-Analysis of Diagnostic Accuracy with mada 2012. [Accessed 2019 Jan 6]. Available from https://cran.r-project.org/web/packages/mada/vignettes/mada.pdf
27. Aaberg KM, Gunnes N, Bakken IJ, Lund Søraas C, Berntsen A, Magnus P, et al. Incidence and prevalence of childhood epilepsy: a nationwide cohort study. Pediatrics. 2017;139:5.
28. Bellini I, Policardo L, Zaccara G, Palumbo P, Rosati E, Torre E, et al. Identification of prevalent patients with epilepsy using administrative data: the Tuscany experience. Neurol Sci. 2017;38(4):571–7.
29. Chen CC, Chen LS, Yen MF, Chen H-H, Liou H-H. Geographic variation in the age- and gender-specific prevalence and incidence of epilepsy: analysis of Taiwanese National Health Insurance-based data. Epilepsia. 2012;53(2):283–90.
30. Christensen J, Vestergaard M, Olsen J, Sidenius P. Validation of epilepsy diagnoses in the Danish National Hospital Register. Epilepsy Res. 2007;75(2–3):162–70.
31. Coulter A, Brown S, Daniels A. Computer held chronic disease registers in general practice: a validation study. J Epidemiol Community Health. 1989;43(1):25–8.
32. De Jesus-Alvelo I, Labovitz D. How reliable are the ICD9-CM billing codes in the administrative data to estimate the risk of seizures and epilepsy after stroke? Neurology. 2013;12:80.
33. Engeland A, Bjorge T, Daltveit AK, Vollset SE, Furu K. Validation of disease registration in pregnant women in the Medical Birth Registry of Norway Acta Obstet Gynecol Scand. 2009;88(10):1083–9.
34. Foebel AD, Hirdes JP, Heckman GA, Kergoat M-J, Patten S, Marrie RA. Diagnostic data for neurological conditions in interRAI assessments in home care, nursing home and mental health care settings: a validity study. BMC Health Serv Res. 2013;1(13):457.
35. Franchi C, Giussani G, Messina P, Montesano M, Romi S, Nobili A. Validation of healthcare administrative data for the diagnosis of epilepsy. J Epidemiol Community Health. 2013;67(12):1019–24.
36. Frost FJ, Hurley JS, Petersen HV, Gunter MJ, Gause D. A comparison of two methods for estimating the health care costs of epilepsy. Epilepsia. 2000;41(8):1020–6.
37. Holden EW, Grossman E, Nguyen HT, Gunter MJ, Grebosky B, Worley AV, et al. Developing a computer algorithm to identify epilepsy cases in managed care organizations. Dis Manag. 2005;8(1):1–14.
38. Jette N, Reid AY, Quan H, Hill MD, Wiebe S. How accurate is ICD coding for epilepsy? Epilepsia. 2010;51(1):62–9.
39. Lee SY, Chung SE, Kim DW, Eun SH, Kang HChu, Cho YW, et al. Estimating the prevalence of treated epilepsy using administrative health data and its validity: ESSENCE study. J Clin Neurol. 2016;12(4):434–40.
40. Marrie RA, Yu BN, Leung S, Elliott L, Caetano P, Warren S, et al. The utility of administrative data for surveillance of comorbidity in multiple sclerosis: a validation study. Neuroepidemiology. 2013;40(2):85–92.
41. Meeraus WH, Petersen I, Chin RF, Knott F, Gilbert R. Childhood epilepsy recorded in primary care in the UK. Arch Dis Child. 2013;98(3):195–202.
42. Moura LM, Price M, Cole AJ, Hoch DB, Hsu J. Accuracy of claims-based algorithms for epilepsy research: revealing the unseen performance of claims-based studies. Epilepsia. 2017;58(4):683–91.
43. Parko K, Thurman DJ. Prevalence of epilepsy and seizures in the Navajo Nation 1998–2002. Epilepsia. 2009;50(10):2180–5.
44. Pina-Garza JE, Vekeman F, Cheng W, Tuttle E, Giguère-Duval P, Oganisian A, et al. Development of a claims-based classifier to identify lennox-gastaut syndrome. Neurology Conference: 67th American Academy of Neurology Annual Meeting, AAN. 2015;84 (Suppl 14). Conference Abstract.

45. Pugh MJ, Parko K. Research using archival health care data: let the buyer beware. Epilepsia. 2015;56(2):321–2.
46. Pugh MJ, Van Cott AC, Cramer JA, Knoefel JE, Amuan ME, Tabares J, et al. Trends in antiepileptic drug prescribing for older patients with new-onset epilepsy: 2000-2004. Neurology. 2008;70(Issue 22, Part 2):2171–8.
47. Rehman R, Everhart A, Frontera AT, Kelly PR, Lopez M, Riley D, et al. Implementation of an established algorithm and modifications for the identification of epilepsy patients in the veterans health administration. Epilepsy Res. 2016;127:284–90.
48. Reid AY, St Germaine-Smith C, Liu M, Sadiq S, Quan H, Wiebe S, et al. Development and validation of a case definition for epilepsy for use with administrative health data. Epilepsy Res. 2012;102(3):173–9.
49. Shackleton DP, Westendorp RG, Kasteleijn-Nolst Trenite DG, de Boer A, Herings RMC. Dispensing epilepsy medication: a method of determining the frequency of symptomatic individuals with seizures. J Clin Epidemiol. 1997;50(9):1061–8.
50. Syvertsen M, Nakken KO, Edland A, Hansen G, Hellum MK, Koht J. Prevalence and etiology of epilepsy in a Norwegian county-a population based study. Epilepsia. 2015;56(5):699–706.
51. Tan M, Wilson I, Braganza V, Ignatiadis S, Boston R, Sundararajan V, et al. Development and validation of an epidemiologic case definition of epilepsy for use with routinely collected Australian health data. Epilepsy Behav. 2015;51:65–72.
52. Tu K, Wang M, Jaakkimainen RL, Butt D, Ivers NM, Young J, et al. Assessing the validity of using administrative data to identify patients with epilepsy. Epilepsia. 2014;55(2):335–43.
53. Wassenaar M, Carpay JA, Sander JW, Thijs RD. Validity of health insurance data to identify people with epilepsy. Epilepsy Res. 2018;139:102–6.
54. Williamson T, Green ME, Birtwhistle R, Khan S, Garies S, Wong ST, et al. Validating the 8 CPCSSN case definitions for chronic disease surveillance in a primary care database of electronic health records. Ann Fam Med. 2014; 12(4):367–72.
55. Pickrell WO, Lacey AS, Bodger OG, Demmler JC, Thomas RH, Lyons RA, et al. Epilepsy and deprivation, a data linkage study. Epilepsia. 2015;56(4):585–91.
56. Chubak J, Pocobelli G, Weiss NS. Tradeoffs between accuracy measures for electronic health care data algorithms. J Clin Epidemiol. 2012;65(3):343–9.e2.
57. Watson N. Using clinical coding systems to best effect in electronic records. Guidelines in practice [Internet]. 2001. [Accessed 2020 Jan 9]. Available from https://www.guidelinesinpractice.co.uk/using-clinical-coding-systems-to-best-effect-in-electronic-records/305085.article
58. Widdifield J, Ivers NM, Young J, Green D, Jaakkimainen L, Butt DA, et al. Development and validation of an administrative data algorithm to estimate the disease burden and epidemiology of multiple sclerosis in Ontario. Canada Mult Scler. 2015;21(8):1045–54.
59. Aaberg KM, Bakken IJ, Lossius MI, Lund Soraas C, Haberg SE, Stoltenberg C, et al. Comorbidity and childhood epilepsy: a nationwide registry study. Pediatrics. 2016;138:3.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Mbizvo GK, Bennett K, Schnier C, Simpson CR, Duncan SE, Chin RFM. The accuracy of using administrative healthcare data to identify epilepsy cases: A systematic review of validation studies. *Epilepsia*. 2020;61:1319–1335. https://doi.org/10.1111/epi.16547

3.5 Conclusion

This chapter addresses our aim to estimate the overall accuracy of using administrative healthcare data to identify epilepsy cases. The results indicate that it is likely to be reasonable to use such datasets for this purpose, with PPV, sensitivity, NPV, and specificity estimates frequently >80–90%. The datasets appear to have somewhat preserved resilience for accurate case ascertainment even across different countries, coding systems, reference standards, and study designs. We learn that higher PPVs are likely to be promoted by combining disease codes (ICD-10 G40–41 for epilepsy and status epilepticus) with AEDs, and that replacing disease codes with symptom codes (ICD-10 R56 for seizures) within this equation helps promote higher sensitivities. AEDs on their own are unlikely to be sufficiently accurate for use. We also learn in this chapter that some studies validating administrative epilepsy datasets have been somewhat poor in reporting certain aspects of study design, such as if the reference standard reader was blinded from the administrative codes, alongside using inconsistent or poor reference standards. This has led to unclear or high risks of bias, often associated with poorer accuracy estimates. Such an awareness should help inform and improve the design of any subsequent diagnostic accuracy work within the current thesis.

This systematic review searched the literature extensively and yet did not identify any studies validating administrative epilepsy datasets in Scotland. Furthermore, no studies assessed the accuracy of administrative data identifying epilepsy in deceased adults. Specifically, it remains unclear how

accurately causes of death featuring disease or symptom codes indicative of epilepsy (ICD-10 G40–41, R56) reflect a correct reference standard epilepsy diagnosis, yet coded causes of death remain administrative and therefore liable to the same sources of inaccuracy as in any administrative dataset. It may be that they also need to be combined with AEDs, for example, to promote greater case-ascertainment accuracy. The diagnostic accuracy of coded causes of death has been validated in other conditions outside of epilepsy literature.[179] In the next chapter, we validate the accuracy of administrative data identifying epilepsy in deceased adults in Scotland, and seek to identify the optimal case-ascertainment algorithm to allow investigation of epilepsy-related deaths in subsequent chapters.

The systematic review protocol and manuscript methodologies differed mainly in that we primarily favoured a narrative synthesis of review findings in the latter, over a meta-analysis initially proposed in the former. A narrative synthesis was preferred because almost all studies validating administrative epilepsy datasets estimated PPVs, with or without sensitivities, and very few estimated NPV or specificity. PPVs cannot be reasonably pooled into a meta-analysis due to the expected heterogeneity from variation in disease prevalence and different disease positivity thresholds.[130,155,177] However, as measures, PPV and sensitivity focus on identifying individuals with disease correctly, whilst NPV and specificity focus on identifying individuals without disease correctly. Therefore, PPV and sensitivity remain the most important outcomes to validate in our study of epilepsy-related deaths.

Chapter 4. Validating the accuracy of administrative healthcare data identifying epilepsy in deceased adults: A Scottish data linkage study

4.1 Introduction

The aim of this chapter is to identify the optimal algorithms for identifying epilepsy cases in deceased adults using administrative data in Scotland. We undertake a nationwide administrative data linkage study in which we link mortality records, hospital admissions, drug prescribing, and primary care data on deceased adults (aged ≥16 years) in Scotland. The diagnostic accuracy of these datasets is assessed both alone and in algorithmic combinations of epilepsy disease and symptom codes, as well as AEDs. The focus is to validate whether ICD-10 codes G40–41 (epilepsy and status epilepticus) and R56.8 (seizures) coded as causes of death correctly reflect a reference standard epilepsy diagnosis, and if this can be optimised by linking in the additional datasets (AEDs, hospital admissions, or primary care attendances related to epilepsy). Medical records are used as the reference standard. We estimate PPVs and sensitivities as the outcomes of interest. This validation study was published in *Epilepsy Research* in 2020. The published manuscript has been used as the body of this thesis chapter, with permission to do so granted from publishers. References listed in the published manuscript are integrated into the thesis references.[1,3,5,12,13,15,16,31,42,43,55,57,59,61-63,65,66,68,85,102,109,120,126-128,130,143,145,162,166,173,174,179-203]

Student contributions: My supervisors had always been keen for the first step in this work to involve validating the diagnostic accuracy of epilepsy coding in Scotland and I was tasked with coming up with a sensible way to do

so. I proposed various methods and we agreed on the one described in this chapter, and I subsequently analysed data gathered in the linked datasets against data abstracted from medical records. I drafted and successfully submitted the work for publication.

4.2 Manuscript

Epilepsy Research 167 (2020) 106462

Contents lists available at ScienceDirect

Epilepsy Research

journal homepage: www.elsevier.com/locate/epilepsyres

Validating the accuracy of administrative healthcare data identifying epilepsy in deceased adults: A Scottish data linkage study

Gashirai K. Mbizvo [a,*], Christian Schnier [b], Colin R. Simpson [b,c], Susan E. Duncan [a,d], Richard F. M. Chin [a,e]

[a] The University of Edinburgh Centre for Clinical Brain Sciences, Muir Maxwell Epilepsy Centre, Edinburgh, UK
[b] The University of Edinburgh, Usher Institute, Edinburgh, UK
[c] Victoria University of Wellington, School of Health, Wellington Faculty of Health, New Zealand
[d] Western General Hospital, Department of Clinical Neurosciences, Edinburgh, UK
[e] Royal Hospital for Sick Children, Edinburgh, UK

ARTICLE INFO

Keywords:
Seizures
Diagnostic accuracy study
Routine data
Cause of death
ICD-10
Mortality

ABSTRACT

Background: We investigate the case-ascertainment accuracy for potentially active epilepsy of four administrative healthcare datasets used to identify deceased adults in Scotland.

Methods: In this diagnostic accuracy study, unique patient identifiers were used to link administrative healthcare data for adults (aged 16 years and over) who died in Scotland between 01/01/09–01/01/16. Cases were ascertained from linking mortality records, hospital admissions, antiepileptic drug (AED) prescriptions, and primary care attendances. We assessed ICD-10 codes G40 (epilepsy), G41 (status epilepticus), and R56.8 (seizures) listed as causes of death and as hospital admission reasons, various AEDs, and F25 primary care epilepsy Read codes. These epilepsy indicators were searched through 01/01/09–01/01/16, suggesting active epilepsy during a maximal period of seven years before death. They were compared to epilepsy diagnoses made from medical records reviewed by a senior epileptologist, with a second senior epileptologist independently reviewing the medical records in a 10 % sample to check for specialist interrater agreement in epilepsy diagnoses. We validated how accurately epilepsy was identified by each dataset alone and when combined, calculating positive predictive value (PPV) and sensitivity (with 95 % confidence intervals (CIs)).

Results: 159,032 deceased potential epilepsy cases were captured across the four datasets. Medical records reviewed in a random sample of 936 confirmed that epilepsy was present in 614 and absent in 322. Specialist interrater diagnostic agreement was substantial (100 medical records reviewed in duplicate, kappa = 0.72, CI 0.58–0.86). G40–41 cause of death codes had a PPV of 86 % (CI 84–89 %) and sensitivity of 73 % (CI 69–76 %). Adding R56.8 lowered PPV to 69 % (CI 65–72 %) and raised sensitivity to 87 % (CI 84–90 %). The optimal algorithm combining two datasets consisted of F25 Read codes paired with AEDs (PPV 86 % (CI 80–92 %), sensitivity 93 % (CI 88–97 %)). Also effective was pairing G40–41 and/or R56.8 cause of death codes with AEDs (PPV 91 % (CI 89–94 %), sensitivity 81 % (CI 77–84 %)). Whilst algorithms combining three datasets raised PPV to as high as 93–95 %, the associated sensitivities were low (71 % at most).

Conclusions: Routinely-collected Scottish data can accurately identify epilepsy in deceased adults. It may be necessary to combine the diagnostic coding used with AEDs to ensure optimal case-ascertainment. The results help inform the design of future Scottish epilepsy mortality studies recruiting from administrative data sources.

1. Introduction

Epilepsy is common and contributes 0.7 % of the global burden of disease according to various administrative healthcare data sources (Murray et al., 2012; Thijs et al., 2019). Although administrative data are widely available resources for epilepsy research (Garratt et al.,

* Corresponding author at: University of Edinburgh Centre for Clinical Brain Sciences, Muir Maxwell Epilepsy Centre, Child Life and Health, 20 Sylvan Place, EH9 1UW, Edinburgh, UK.
E-mail addresses: gashiraimbizvo@hotmail.com (G.K. Mbizvo), christian.schnier@ed.ac.uk (C. Schnier), c.simpson@ed.ac.uk (C.R. Simpson), susanxduncan@gmail.com (S.E. Duncan), rchin@exseed.ed.ac.uk (R.F.M. Chin).

https://doi.org/10.1016/j.eplepsyres.2020.106462
Received 25 June 2020; Received in revised form 29 July 2020; Accepted 4 September 2020
Available online 13 September 2020
0920-1211/© 2020 Elsevier B.V. All rights reserved.

2010), their accuracy in correctly identifying epilepsy cases requires validation because the potential for epilepsy misdiagnosis remains large (Garratt et al., 2010; ILAE, 1993; Kee et al., 2012; Mbizvo et al., 2020). This is because such data were originally collected for routine non-scientific purposes, such as to help organise health insurance claims (England et al., 2013).

People with epilepsy (PWE) are at increased risk of premature death (Escalaya et al., 2015; Levira et al., 2016; Nevalainen et al., 2014; Shackleton et al., 2002; Thurman et al., 2017; Watila et al., 2018). In a recent systematic review of mortality in PWE, we identified a paucity of studies assessing the validity of using administrative data to identify epilepsy cases (Mbizvo et al., 2019a, 2020). Most studies relied on identifying cases using unvalidated epilepsy coding within hospital admissions data or International Classification of Disease (ICD) codes derived from causes of death listed within death certificates (Mbizvo et al., 2019a, 2020; Pickrell and Kerr, 2020). The presence of an epilepsy code within the listed causes of death of a patient does not necessarily mean they had epilepsy. Various potential sources of error remain possible, such as an incorrect diagnosis of epilepsy being listed by the death certifying doctor in a patient who did not actually have epilepsy (Death Certifcation Review Service, 2016), or human error occurring during the subsequent coding process (O'Malley et al., 2005). For example, the G40–41 ICD-10 codes for epilepsy and status epilepticus are often incorrectly coded as G43.1 (migraine) or G45.x (transient cerebral ischemia) (Jette et al., 2010; Tan et al., 2015). Furthermore, a deceased PWE might not be identified in an epilepsy mortality study if death certificates are the only source for case-ascertainment. This is because the diagnosis of epilepsy would be excluded from their death certificate if epilepsy was not felt to have contributed to causing death (Chief Medical Officer, 2018; National Records of Scotland, 2020).

In an ongoing, retrospective, population-based, Scottish Epilepsy Deaths Study (SEDS) (Mbizvo et al., 2019b), we identified people who died who may have had active epilepsy before death by searching four linked administrative healthcare databases for indicators of epilepsy occurring within a maximum of seven years before death occurred, suggesting that epilepsy was still active (Fisher et al., 2014). In the current paper, we perform a diagnostic accuracy study to investigate how accurately the coding strategies we used identified epilepsy within these four databases. The outcomes of interest were positive predictive value (PPV) and sensitivity, the most commonly used outcomes in studies validating the accuracy of administrative data in epilepsy (Kee et al., 2012; Mbizvo et al., 2020).

2. Methods

2.1. Environment

Healthcare to the five million population of Scotland is delivered by a single state-provided National Health Services (NHS), administered by 14 regional Health Boards (Pavis and Morris, 2015). To access NHS care, patients first register with a general practitioner (GP), who provides them with a Community Health Index (CHI) number. This is a unique patient identifier used whenever they subsequently access any NHS service. As 96.5–99.9 % of the Scottish population has a CHI number (Pavis and Morris, 2015), it can be used to create nationally linked research datasets derived from almost the entire Scottish population (Katikireddi et al., 2018).

2.2. Study design

We undertook a diagnostic accuracy study to assess how accurately administrative data identified epilepsy cases in adults who had died between 01/01/09–01/01/16 and had been identified by various diagnostic coding and/or antiepileptic drug (AED) strategies used within the following administrative healthcare databases: A) National Records of Scotland (NRS) death records (Information Services Division, 2016c); B) Scottish Morbidity Record 01 (SMR01) hospital admissions (Information Services Division, 2016b); C) Prescribing Information System (PIS) Scottish prescribing data (Information Services Division, 2016a); and GP primary care data. Each database was searched for cases retrospectively between 01/01/09–01/01/16. We used hand-searched medical records as the diagnostic gold standard against which to compare the administrative data for accuracy (Mbizvo et al., 2020). Given the expectation of a large number of deceased participants to be captured by the four datasets, we aimed to review a randomly generated convenience sample focusing on eight of 14 Scottish NHS Health Boards (Ayrshire and Arran, Fife, Forth Valley, Grampian, Glasgow, Lanarkshire, Lothian, and Tayside) (Bossuyt et al., 2015). This study was designed to conform to the Standards for Reporting of Diagnostic Accuracy Studies (STARD 2015) (Bossuyt et al., 2015).

2.3. Study population

Eligible participants were adults (aged ≥16 years) who had died between 01/01/09–01/01/16, were captured within the administrative databases, and had medical records available for review. The focus was on adults as mortality in children with epilepsy has been studied in detail elsewhere (RCPCH, 2013) and was therefore not the focus of our epilepsy mortality study (SEDS) – which was the data substrate for the current diagnostic accuracy study.

2.4. Linkage of datasets

The Scottish Information Services Division (ISD) is the legislated producer of NHS healthcare statistics to inform quality assessment and Government resource allocation (Public Health Scotland, 2020). Therefore, it is mandatory for NHS care providers to submit their administrative data to ISD. ISD then collate and process these data to produce several key national datasets (Information Services Division, 2016b, 2016c, 2016a). These national datasets are also routinely extracted and made available to approved researchers in a secure analytical environment (Safe Haven) as linked and deidentified research datasets (Pavis and Morris, 2015). This process is managed by the electronic Data Research and Innovation Service (eDRIS) of ISD, based at Farr Institute Scotland (Pavis and Morris, 2015). Primary care data from GPs can also be extracted for research purposes by an approved third party (Albasoft) (Albasoft, 2020) and transferred to eDRIS (Douglas et al., 2015). The availability of CHI numbers as unique patient identifiers makes it possible to link primary care data to NRS, SMR01, and PIS as these datasets also hold the CHI variable (Pavis and Morris, 2015).

We utilised these existing linkage infrastructures to establish an epilepsy mortality research dataset containing linked national data from ISD and regional primary care data (Fig. 1) (eDRIS Team, 2020). To facilitate this, eDRIS created a linked dataset in our Safe Haven containing deceased participants recruited according to the coding and/or AED prescription criteria listed in 2.5.1–2.5.4 below (see *Datasets*) having occurred between 01/01/09–01/01/16. The codes chosen were based on previous epilepsy validation literature (Kee et al., 2012; Mbizvo et al., 2020). We validated accuracy of the coding and/or prescription conditions listed in 2.5.1–2.5.4 both alone and in multiple algorithm combinations (61 in total, see Tables 1–3).

2.5. Datasets (each searched through 01/01/09–01/01/16)

2.5.1. NRS death records

(Information Services Division, 2016c): those who had died with ≥1 G40–41 (epilepsy and status epilepticus) and/or R56.8 (seizure) code listed within their causes of death in the NRS death records. All Scottish deaths are registered at NRS within eight days of occurrence using death certification (Chief Medical Officer, 2018). The causes of death are listed in Parts 1 and 2 of a death certificate as free-text. An automated method

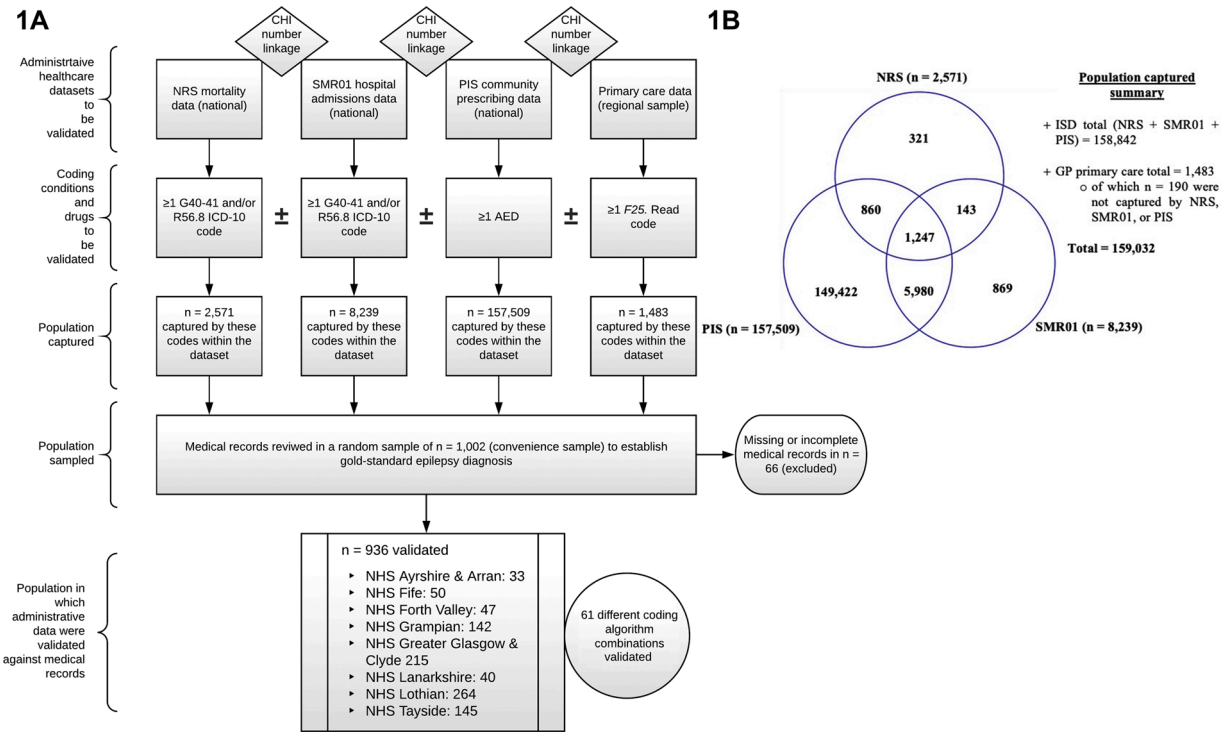


Fig. 1. Study design – **1A Flow diagram** summarising flow of participants to create a linked research dataset; **1B – Venn diagram** summarising distribution of the deceased population captured across the databases. The primary care database was unable to be included in the Venn diagram due to risk of patient deidentification as some of the overlapping categories with other databases contained five or fewer participants.
Abbreviations: AED – antiepileptic drug; *F25* – primary care diagnostic Read codes for epilepsy; G40–41 – International Classification of Disease 10 (ICD-10) codes for epilepsy and status epilepticus; R56.8 – ICD-10 code for seizures; CHI – Scottish Community Health Index NHS patient identification number; GP – General practitioner

Table 1
Results of Level 1 validation of a single database coding or AED strategy.

	Database	Coding algorithm	Participants captured	Sample validated	Epilepsy confirmed	TP	FP	FN	PPV (95 % CI)	Sn (95 % CI)
1*	**PIS**	**≥1 AED (NL)**	**22,460**	**936**	**614**	**560**	**64**	**54**	**90 %(87–92 %)**	**91 %(89–93 %)**
2	NRS	≥1 G40	2001	936	614	422	47	192	90 % (87–93 %)	69 % (65–72 %)
3	NRS	≥1 G40–41	2143	936	614	446	70	168	86 % (84–89 %)	73 % (69–76 %)
4	SMR01	≥1 G40–41, R56.8	8239	936	614	450	110	164	80 % (77–84 %)	73 % (70–77 %)
5	PIS	≥1 AED (BL)	157,509	936	614	576	149	38	79 % (77–82 %)	94 % (92–96 %)
6	Primary care	≥1 F25	1483	155	120	120	35	NA	77 % (70–83 %)	NA
7	NRS	≥1 G40–41, R56.8	2571	936	614	533	243	81	69 % (65–72 %)	87 % (84–90 %)
8	NRS	≥1 G41	179	936	614	39	25	575	61 % (49–73 %)	6% (4–8%)
9	NRS	≥1 R56.8	609	936	614	102	160	512	39 % (33–45 %)	17 % (14–20 %)

Key: * = Optimal case-ascertainment strategy(s).
Abbreviations: SMR – Scottish Morbidity Record; PIS – Prescription Information Service; NRS – National Records of Scotland; AED – antiepileptic drug; F25 – primary care diagnostic Read Codes for epilepsy; G40–41 – International Classification of Disease 10 (ICD-10) codes for epilepsy and status epilepticus; R56.8 – ICD-10 code for seizures; NL – AEDs on the narrow list (appendix S2b); BL – AEDs on the broad list (appendix S2a); TP – true positive; FP – false positives; FN – false negatives, PPV – positive predictive value; Sn – sensitivity; CI – confidence intervals; NA – not applicable (negative cases unavailable as primary care data were taken from a 10 % sample of Scottish GP practices).

of coding, based on a complex set of modification and selection rules, is then used alongside experienced coding staff at NRS to convert the free-text causes of death in Parts 1–2 into their corresponding ICD-10 codes (National Records Scotland, 2016). NRS arrange these ICD-10 codes into a single primary cause of death field and up to nine secondary cause of death fields. Unique to Scotland, the causes of death are reviewed for accuracy by the NRS staff and, where necessary, amended using information from various sources (see appendix A for further details) (Death Certifcation Review Service, 2016; Fernie, 2019; National Records Scotland, 2016). We extracted all of the ICD-10-coded causes of death and the corresponding Part 1–2 death certificate free-texts held by NRS.

AND/OR

2.5.2. SMR01 hospital admissions

(Information Services Division, 2016b): those who had experienced ≥1 hospital admission diagnosis coded as G40–41 and/or R56.8, regardless of their cause of death. SMR01 is a mandatory national dataset capturing all of the inpatient hospital admissions in Scotland. 1.4 million SMR01 records are generated yearly (Information Services Division, 2016b). Data collected include a mandatory diagnostic field containing up to six ICD-10-coded hospital admission diagnoses (Information Services Division, 2016b).

AND/OR

2.5.3. AEDs on PIS

(Information Services Division, 2016a): those who were prescribed

Table 2
Results of Level 2 validation of algorithms combining two database coding or AED strategies together.

	Database	Coding algorithm	Participants captured	Sample validated	Epilepsy confirmed	TP	FP	FN	PPV (95 % CI)	Sn (95 % CI)
1	Primary care + NRS	≥1 F25 + ≥1 R56.8	13	155	120	8	0	112	100 % (100–100 %)	7% (2–11 %)
2	NRS + PIS	≥1 G40−41 + ≥1 AED (NL)	1781	936	614	419	28	195	94 % (92–96 %)	68 % (65–72 %)
3	NRS + PIS	≥1 G40 + ≥1 AED (NL)	1732	936	614	402	26	212	94 % (92–96 %)	66 % (62–69 %)
4	SMR01 + PIS	≥1 G40−41, R56.8 + ≥1 AED (NL)	6501	936	614	418	34	196	93 % (90–95 %)	68 % (64 %–72 %)
5	NRS + PIS	≥1 G40 + ≥1 AED (BL)	1798	936	614	407	31	207	93 % (91–95 %)	66 % (63–70 %)
6	SMR01 + NRS	SMR01 ≥ 1 G40−41, R56.8 + NRS ≥ 1 G40	1167	936	614	325	23	289	93 % (91–96 %)	53 % (49–57 %)
7*	**NRS + PIS**	**≥1 G40−41, R56.8 + ≥1 AED (NL)**	**1921**	**936**	**614**	**494**	**47**	**120**	**91 %(89–94 %)**	**81 %(77–84 %)**
8	NRS + PIS	≥1 G40−41 + ≥1 AED (BL)	1883	936	614	426	42	188	91 % (88–94 %)	69 % (66–73 %)
9	SMR01 + PIS	≥1 G40−41, R56.8 + ≥1 AED (BL)	7227	936	614	424	48	190	90 % (87–93 %)	69 % (65–73 %)
10	Primary care + NRS	≥1 F25 + ≥1 G40−41, R56.8	199	155	120	68	8	52	90 % (83–96 %)	57 % (48–66 %)
11	SMR01 + NRS	SMR01 ≥ 1 G40−41, R56.8 + NRS ≥ 1 G40−41	1273	936	614	347	42	267	89 % (86–92 %)	57 % (53–60 %)
12	Primary care + NRS	≥1 F25 + ≥1 G40−41	192	155	120	63	8	57	89 % (81–96 %)	53 % (44–61 %)
13	Primary care + NRS	≥1 F25 + ≥1 G41	8	155	120	X	X	X	89 % (81–96 %)	53 % (44–61 %)
14	Primary care + NRS	≥1 F25 + ≥1 G40	188	155	120	62	8	58	89 % (81–96 %)	52 % (43–61 %)
15	NRS + PIS	≥1 G41 + ≥1 AED (NL)	81	936	614	X	X	X	88 % (77–99 %)	5% (3–7%)
16*	**Primary care + PIS**	**≥1 F25 + ≥1 AED (NL)**	**1089**	**155**	**120**	**111**	**18**	**9**	**86 %(80–92 %)**	**93 %(88–97 %)**
17	Primary care + SMR01	≥1 F25 + ≥1 G40−41, R56.8	613	155	120	85	15	35	85 % (78–92 %)	71 % (63–79 %)
18	NRS + PIS	≥1 G40−41, R56.8 + ≥1 AED (BL)	2107	936	614	507	98	107	84 % (81–87 %)	83 % (80–86 %)
19	SMR01 + NRS	≥1 G40−41, R56.8	1390	936	614	386	76	228	84 % (80–87 %)	63 % (59–67 %)
20	NRS + PIS	≥1 R56.8 + ≥1 AED (NL)	281	936	614	89	17	525	84 % (77–91 %)	15 % (12–17 %)
21	Primary care + PIS	≥1 F25 + ≥1 AED (BL)	1255	155	120	112	27	8	81 % (74–87 %)	93 % (89–98 %)
22	NRS + PIS	≥1 G41 + ≥1 AED (BL)	117	936	614	32	13	582	71 % (58–84 %)	5% (4–7%)
23	SMR01 + NRS	SMR01 ≥ 1 G40−41, R56.8 + NRS ≥ 1 R56.8	220	936	614	58	30	556	66 % (56–76 %)	9% (7–12 %)
24	NRS + PIS	≥1 R56.8 + ≥1 AED (BL)	403	936	614	95	52	519	65% (57–72 %)	16 % (13–18 %)
25	SMR01 + NRS	SMR01 ≥ 1 G40−41, R56.8 + NRS ≥ 1 G41	139	936	614	36	21	578	63 % (51–76 %)	6% (4–8%)

Key: * = Optimal case-ascertainment strategy(s); X – categories with five or less events hidden to protect patient identity.
Abbreviations: SMR – Scottish Morbidity Record; PIS – Prescription Information Service; NRS – National Records of Scotland; AED – antiepileptic drug; F25 – primary care diagnostic Read Codes for epilepsy; G40–41 – International Classification of Disease 10 (ICD-10) codes for epilepsy and status epilepticus; R56.8 – ICD-10 code for seizures; NL – AEDs on the narrow list (appendix S2b); BL – AEDs on the broad list (appendix S2a); TP – true positive; FP – false positives; FN – false negatives, PPV – positive predictive value; Sn – sensitivity; CI – confidence intervals.

≥1 AEDs (regardless of their cause of death). PIS is a mandatory national prescribing dataset used to track the community drug history of the entire Scottish population in order to facilitate public drug reimbursement (Information Services Division, 2016a). Approximately 100 million prescription data items are uploaded to PIS yearly (Information Services Division, 2016a). We used PIS to screen for whether or not any deceased adults in Scotland were prescribed one or more AEDs from a broad list of 36 AEDs provided in appendix B1 during the study period, recruiting any who were. Such a broad list was aimed at increasing sensitivity, creating a file with the potential to have captured nearly all PWE in the country who died (assuming that most were prescribed an AED) (Hamer et al., 2012). However, as this broad list also included several AEDs often prescribed for indications other than epilepsy (e.g. gabapentin and diazepam), we created an additional narrow AED filter in which such agents were excluded (leaving 21 AEDs, see appendix B2) to try and explore whether this would recruit a sample with fewer false positives.

AND/OR

2.5.4. *Primary care*

those who had ≥1 epilepsy-related (Read code *F25*) primary care attendance (regardless of cause of death). Read codes are a coded thesaurus of clinical terms used in primary care to create electronic health records (Fonferko-Shadrach et al., 2017). Currently, there is no routine primary care epilepsy dataset held at ISD (Information Services Division, 2020). Therefore, we established one consisting of 100 GP practices (a convenience sample of 10 % of 900 Scottish GP practices) using an extension of methods detailed elsewhere (see appendix C for Read codes used) (Douglas et al., 2015; Fonferko-Shadrach et al., 2017; Pickrell et al., 2015). To our knowledge, this is a larger sample of Scottish GP practices providing epilepsy data than other established primary care datasets, including The Health Improvement Network and

Table 3
Results of Level 3 validation of algorithms combining three database coding or AED strategies together.

	Database	Coding algorithm	Participants captured	Sample validated	Epilepsy confirmed	TP	FP	FN	PPV (95 % CI)	Sn (95 % CI)
1	Primary care + NRS + PIS	≥1 F25 + ≥1 R56.8 + ≥1 AED (BL)	11	155	120	7	0	113	100 % (100–100 %)	6% (2–10 %)
2	Primary care + NRS + PIS	≥1 F25 + ≥1 R56.8 + ≥1 AED (NL)	11	155	120	7	0	113	100 % (100–100 %)	6% (2–10 %)
3	Primary care + SMR01 + NRS	≥1 F25 + SMR01 ≥ 1 G40–41, R56.8 + NRS ≥ 1 R56.8	11	155	120	7	0	113	100 % (100–100 %)	6% (2–10 %)
4	Primary care + NRS + PIS	≥1 F25 + ≥1 G41 + ≥1 AED (BL)	8	155	120	X	X	X	100 % (100–100 %)	3% (0–5%)
5	Primary care + NRS + PIS	≥1 F25 + ≥1 G41 + ≥1 AED (NL)	8	155	120	X	X	X	100 % (100–100 %)	3% (0–5%)
6	Primary care + SMR01 + NRS	≥1 F25 + SMR01 ≥ 1 G40–41, R56.8 + NRS ≥ 1 G41	7	155	120	X	X	X	100 % (100–100 %)	2% (0–4%)
7	SMR01 + NRS + PIS	SMR01 ≥ 1 G40–41, R56.8 + NRS ≥ 1 G40–41 + ≥1 AED (NL)	1119	936	614	328	19	286	95 % (92–97 %)	53 % (50–57 %)
8	SMR01 + NRS + PIS	SMR01 ≥ 1 G40–41, R56.8 + NRS ≥ 1 G40 + ≥1 AED (BL)	1101	936	614	314	17	300	95 % (93–97 %)	51 % (47–55 %)
9	SMR01 + NRS + PIS	SMR01 ≥ 1 G40–41, R56.8 + NRS ≥ 1 G40 + ≥1 AED (NL)	1075	936	614	314	17	300	95 % (93–97 %)	51 % (47–55 %)
10	SMR01 + NRS + PIS	≥1 G40–41, R56.8 + ≥1 AED (NL)	1185	936	614	366	26	248	93 % (91–96 %)	60 % (56–64 %)
11	SMR01 + NRS + PIS	SMR01 ≥ 1 G40–41, R56.8 + NRS ≥ 1 G40–41 + ≥1 AED (BL)	1168	936	614	331	26	283	93 % (90–95 %)	54 % (50–58 %)
12	SMR01 + NRS + PIS	≥1 G40–41, R56.8 + ≥1 AED (BL)	1247	936	614	370	38	244	91 % (88–94 %)	60 % (56–64 %)
13	Primary care + SMR01 + PIS	≥1 F25 + ≥1 G40–41, R56.8 + ≥1 AED (NL)	555	155	120	79	9	41	90 % (83–96 %)	66 % (57–74 %)
14	SMR01 + NRS + PIS	SMR01 ≥ 1 G40–41, R56.8 + NRS ≥ 1 R56.8 + ≥1 AED (NL)	156	936	614	56	6	558	90 % (83–98 %)	9% (7–11 %)
15	Primary care + NRS + PIS	≥1 F25 + ≥1 G40–41, R56.8 + ≥1 AED (BL)	195	155	120	66	8	54	89 % (82–96 %)	55 % (46–64 %)
16	Primary care + NRS + PIS	≥1 F25 + ≥1 G40–41, R56.8 + ≥1 AED (NL)	194	155	120	66	8	54	89 % (82–96 %)	55 % (46–64 %)
17	Primary care + SMR01 + NRS	≥1 F25 + ≥1 G40–41, R56.8	115	155	120	48	6	72	89 % (81–97 %)	40 % (31–49 %)
18	Primary care + NRS + PIS	≥1 F25 + ≥1 G40–41 + ≥1 AED (BL)	188	155	120	61	8	59	88 % (81–96 %)	51 % (42–60 %)
19	Primary care + NRS + PIS	≥1 F25 + ≥1 G40–41 + ≥1 AED (NL)	187	155	120	61	8	59	88 % (81–96 %)	51 % (42–60 %)
20	Primary care + NRS + PIS	≥1 F25 + ≥1 G40 + ≥1 AED (BL)	184	155	120	60	8	60	88 % (81–96 %)	50 % (41–59 %)
21	Primary care + NRS + PIS	≥1 F25 + ≥1 G40 + ≥1 AED (NL)	183	155	120	60	8	60	88 % (81–96 %)	50 % (41–59 %)
22	Primary care + SMR01 + NRS	≥1 F25 + SMR01 ≥ 1 G40–41, R56.8 + NRS G40–41	111	155	120	45	6	75	88 % (79–97 %)	38 % (29–46 %)
23	Primary care + SMR01 + NRS	≥1 F25 + SMR01 ≥ 1 G40–41, R56.8 + NRS ≥ 1 G40	107	155	120	44	6	76	88 % (79–97 %)	37 % (28–45 %
24	SMR01 + NRS + PIS	SMR01 ≥ 1 G40–41, R56.8 + NRS ≥ 1 G41 + ≥1 AED (NL)	73	936	614	X	X	X	88 % (76–99 %)	5% (3–6%)
25	Primary care + SMR01 + PIS	≥1 F25 + ≥1 G40–41, R56.8 + ≥1 AED (BL)	577	155	120	80	11	40	85 % (78–92 %)	71 % (63–79 %)
26	SMR01 + NRS + PIS	SMR01 ≥ 1 G40–41, R56.8 + NRS ≥ 1 R56.8 + ≥1 AED (BL)	175	936	614	57	11	557	84 % (75–93 %)	9% (7–12 %)
27	SMR01 + NRS + PIS	SMR01 ≥ 1 G40–41, R56.8 + NRS ≥ 1 G41 + ≥1 AED (BL)	96	936	614	29	11	585	73 % (59–86 %)	5% (3–6%)

Key: X – categories with five or less events hidden to protect patient identity.
Abbreviations: SMR – Scottish Morbidity Record; PIS – Prescription Information Service; NRS – National Records of Scotland; AED – antiepileptic drug; F25 – primary care diagnostic Read Codes for epilepsy; G40–41 – International Classification of Disease 10 (ICD-10) codes for epilepsy and status epilepticus; R56.8 – ICD-10 code for seizures; NL – AEDs on the narrow list (appendix S2b); BL – AEDs on the broad list (appendix S2a); TP – true positive; FP – false positives; FN – false negatives, PPV – positive predictive value; Sn – sensitivity; CI – confidence intervals.

the Clinical Practice Research Datalink (Herrett et al., 2015; Meeraus et al., 2013).

2.6. Establishing gold standard epilepsy diagnosis from medical records

Gold standard epilepsy diagnosis was established from reviewing all available medical records. Access to medical records was made possible through the support of a national network of consultant neurology colleagues (one in each of the 14 Scottish Health Boards), who were provided with CHI numbers of the validation sample by eDRIS. The medical records were then reviewed by an experienced consultant epileptologist (S.E.D.), who was blinded to the participants' administrative data codes and used the medical records to confirm the presence or absence of epilepsy (Kee et al., 2012; Mbizvo et al., 2020). This was based on corroborative evidence such as the presence of two or more unprovoked seizures or clear documentation of a diagnosis of epilepsy from a neurologist (Fisher et al., 2005, 2014). The medical records were extensive, and included electronic and paper records from general and/or specialist inpatients, emergency care, outpatients (including neurology clinics), hospital discharge summaries, referral letters from GPs, radiology scans, neurophysiology reports, and medication lists. This meant that although our network of neurology colleagues could

provide access to medical records, it was not necessary for the patients themselves to have been followed by a neurologist or under specialist care (i.e. they could have had only non-specialist general inpatient/emergency care attendances in their medical records). A second consultant epileptologist (R.F.M.C), also blinded to the participants' administrative data codes, independently reviewed a random sample of 10 % of the validated medical records in NHS Lothian to confirm the presence or absence of epilepsy using the same diagnostic inclusion methods (Fisher et al., 2005, 2014). This was done to assess for interrater agreement (Landis and Koch, 1977; Viera and Garrett, 2005). Participants were sampled using electronically-generated random numbers.

2.7. Statistical analysis

The primary focus was to validate cause of death ICD-10 coding, secondarily supported by validating epilepsy hospital admissions coding in SMR01, AEDs prescribed in PIS, and primary care epilepsy coding. Therefore, we ensured randomisation was weighted toward oversampling from the NRS dataset such that 80 % of the validation sample were captured by at least one G40–41 and/or R56.8 code within their listed causes of death, using electronically generated random numbers to sample participants. Cohen's kappa statistic was used estimate interrater agreement (above chance) between S.E.D and R.F.M.C for medical record epilepsy diagnoses (as yes, no or unclear due to missing or incomplete medical records). The following cut-offs were used: kappa < 0: no agreement; 0.00–0.20: slight agreement; 0.21–0.40: fair agreement; 0.41–0.60: moderate agreement; 0.61–0.80: substantial agreement; 0.81–1.00: almost perfect agreement (Landis and Koch, 1977; Viera and Garrett, 2005).

We validated each of the 61 study algorithms (Tables 1–3), reporting a PPV and sensitivity for each (with 95 % confidence intervals (CIs) generated using an exact method) (Du and Hao, 2019). For clarity, we grouped results into Levels 1, 2 and 3 – presenting the results of a single, two combined, and three combined database coding or AED strategies, respectively. We arranged results in order of highest to lowest PPV. We identified the optimal case-ascertainment algorithm(s) as those with the highest PPV and sensitivity in combination. Categories with five or less events/participants were combined to minimise risk of patient identification (eDRIS Team, 2018). Data were analysed using the "reportROC" and "cutpointr" packages in RStudio Version 1.2.1335 (Du and Hao, 2019; RStudio Team, 2015; Thiele, 2020).

2.8. Approvals

The study was approved by South East Scotland Research Ethics Committee 2 (IRAS 181131, 15/SS/0165), and Scottish Public Benefit and Privacy Panel for Health and Social Care.

3. Results

3.1. Population capture and validation sample characteristics

eDRIS removed 11 people (<1%) and 420 people (5%) from the NRS and SMR01 data files, respectively, as they had no CHI number; precluding linkage. Linkage was otherwise complete and a large dataset containing information from 159,032 deceased adults across Scotland was created for this study (see Fig. 1). From these, a sample of 1002 participants had their medical records reviewed, confirming the presence of epilepsy in 614, and absence in and 322. The remaining 66 participants were excluded from the validation analysis due to missing or incomplete medical records. Specialist interrater diagnostic agreement was classed as substantial (100 medical records reviewed in duplicate, 87 observed agreements, 54 agreements expected by chance, kappa = 0.72, 95 % CI 0.58–0.86) (Landis and Koch, 1977). The 936 validated participants had a mean age of 62 years (±19 years SD). 491 were male (52 %). Participants were mainly from Lothian (n = 264), Glasgow (n = 215), Tayside (n = 145), and Grampian (n = 142), with the remaining 170 spread between Ayrshire and Arran, Fife, Forth Valley, and Lanarkshire. The denominator for validating the primary care dataset was 155 persons for whom primary care data were available.

3.2. Level 1 – validation of a single database coding or AED strategy

Table 1 summarises Level 1 validation results. Within causes of death in the NRS death records, the optimal coding strategy in terms of PPV was using either G40 (epilepsy) alone or G40–41 (epilepsy and/or status epilepticus): PPV range 86–90 %, with overlapping CIs. This indicates that when present, these causes of death normally reflected a correct diagnosis of epilepsy. However, the associated sensitivities were relatively low (ranging 69–73 %, with overlapping CIs), indicating that many positive cases were also missed by each of these coding strategies (high burden of 168–192 false negatives). This is intuitive, as epilepsy and/or status epilepticus might not have appeared on the death certificate of a PWE if they did not die of these (National Records of Scotland, 2020), making them appear falsely negative within NRS cause of death coding. The addition of seizures (R56.8) to the G40–41 epilepsy and/or status epilepticus cause of death coding strategy pushed sensitivity up to a much higher 87 % (CI 84–90 %), lowering the burden of false negatives to 81. However, the PPV for this coding strategy was poor (69 %), owing to many false positives (243) also being associated with the addition of seizures (R56.8). Within causes of death in the NRS death records, the poorest PPV and sensitivity results were from lone R56.8 codes for seizures (PPV 39 %, sensitivity 17 %) and lone G41 codes for status epilepticus (PPV 61 %, sensitivity 6%). These results likely reflect appropriate use of the death certificate to indicate instances when people died of acute symptomatic seizures or *de novo* status epilepticus without an underlying diagnosis of epilepsy (Chief Medical Officer, 2018; Middleton et al., 2018).

A coding strategy combining G40–41 and/or R56.8 within hospital admissions data from SMR01 had a PPV of 80 % and sensitivity of 73 %. This indicates that when present, a hospital admission with these codes often reflected a correct diagnosis of epilepsy but many positive cases were also missed (high number of 164 false negatives). This is intuitive, as not all deceased PWE would have had a prior hospital admission due to epilepsy and they would therefore appear falsely negative on SMR01.

Within the PIS dataset, both our broad list and narrow list of AEDs (appendix B) performed similarly well in terms of a very high sensitivity (91–94 %, with overlapping CIs), indicating that they both had the ability to capture almost all cases of epilepsy in the sample validated (capturing 560–576 of the 614 cases). However, PPV was much higher for the narrow list AEDs (90 %, CI 87–92 %) compared to the broad list AEDs (79 %, CI 77–82 %), indicating that there were many more additional false positive cases captured by the broad list of AEDs (149 false positives) than the narrow list AEDs (64 false positives). The narrow list AEDs were, therefore, more reliable as a case-ascertainment method. The potential impact of this may be illustrated by the large difference in numbers of participants captured by these two AED strategies within the wider PIS dataset (see Fig. 1). There were 157,509 deceased persons captured by the broad list of AEDs within the wider PIS dataset. Restricting this to only those taking AEDs in the narrow list removed a substantial proportion, 135,049 participants, leaving a remaining 22,460 captured by the narrow list AEDs alone. The PPV and sensitivity figures in the sample validated are highly suggestive most of the 22,460 captured by this narrow list had epilepsy whilst a large number of the additional 135,049 captured by the broad AED list were probably false positives. This is consistent with our clinical knowledge that the AEDs that differentiated our narrow list from our broad list of AEDs (e.g. diazepam and gabapentin) are frequently used for conditions other than epilepsy and therefore are likely to have contributed many false positives.

F25 epilepsy Read codes within GP primary care data had a low PPV of 77 %, suggesting that these were poor at correctly identifying epilepsy.

3.3. Level 2 – validation of algorithms combining two database coding or AED strategies together

Table 2 summarises the Level 2 validation results, assessing 25 algorithms. The highest PPV of 100 % was conferred by combining *F25* epilepsy Read codes in primary care with R56.8 seizures codes within the NRS causes of death, but this is unlikely to have been reliable as there were very few people actually captured by this approach and it missed most of the positive cases (7% sensitivity). PPV was otherwise similarly high between 85–94 % for the next 16 algorithms (all with overlapping CIs) regardless of the database assessed or coding strategy used within these 16 algorithms. This suggests that the combination of two databases together normally captured a correct diagnosis of epilepsy and was more important than the specific algorithm used within each of those databases. However, there was some variation in the associated sensitivities across these 16 algorithms, with the highest sensitivity (93 %, CI 88–97 %) conferred by combining *F25* epilepsy Read codes from primary care with AEDs in the narrow list (appendix B2). This strategy had the ability to capture nearly all of the epilepsy cases within the validation sample for whom primary care data were available (111 of 120 epilepsy cases captured). The next most effective coding strategy was combining G40–41 and/or R56.8 causes of death codes with AEDs in the narrow list (sensitivity 81 %, CI 77–84 %). This combination of both a high PPV and sensitivity strongly suggests the 1921 cases this algorithm captured within the wider dataset, all of whom died with epilepsy (G40), status epilepticus (G41), or seizure (R56.8) codes listed as their causes of death, were likely to have had epilepsy. Removing R56.8 from this algorithm in NRS did not change PPV as it was 94 % (CI 92–96) in G40–41 without R56.8 and 91 % (CI 89–94 %) with R56.8 included. However, the loss of R56.8 lowered sensitivity greatly to 68 % (CI 65–72 %) from having been 81 % (CI 77–84 %) with R56.8 included. This means that in these circumstances, the R56.8 cause of death code was helpful in allowing the algorithm to capture a larger proportion of the deceased cases (494 of 614) without creating many additional false positives. The narrow list of AEDs also helped to increase the PPVs of SMR01 hospital admissions captured by G40–41 and/or R56.8 codes to 93 % (CI 90–95 %) compared to when hospital admissions were coded alone in Level 1 (PPV 80 %, CI 77–84 %), although sensitivities remained similarly low: 68 % (CI 64 %–72 %) for the former and 73 % (CI 70–77 %) for the latter.

3.4. Level 3 – validation of algorithms combining three database coding or AED strategies together

Table 3 summarises the Level 3 validation results, assessing 27 algorithms. The first six achieved PPVs of 100 % and included data from primary care. However, these are unlikely to have been reliable as there were very few people actually captured by these algorithms and they missed most of the positive cases (2–6% sensitivity). Although the next 20 algorithms demonstrated a very high PPV between 84–95 % (with overlapping CIs), the overall limitation to accuracy came from generally poor sensitivities, ranging between 5–71 % across these 20 algorithms (most between 50–60 %). This is intuitive in that whilst having epilepsy indicators in three different databases may streamline a patient well into a correct administrative epilepsy diagnosis (high PPV and low false positive rate), not all PWE will have had an opportunity to be captured in three different databases (lowering sensitivity and increasing false negatives).

4. Discussion

We have undertaken the first data-linkage and validation study to focus on establishing how accurately administrative healthcare data identify epilepsy cases in deceased adults (Kerr, 2012; Mbizvo et al., 2020). The findings are likely to be relevant mostly to active cases of epilepsy prior to death, as we required the coded seizure and/or AED prescription indicators of epilepsy be present within a maximum of seven years before death. The International League Against Epilepsy defines epilepsy remission as 10 years seizure-free, with the last 5 years off AEDs (Fisher et al., 2014). We have made three key findings that may help future researchers aiming to identify deceased epilepsy cases using administrative data.

First, we show that ICD-10-coded causes of death alone can be used to identify epilepsy in a fairly accurate manner, although the codes need to be selected appropriately to include either G40 (epilepsy) alone or G40–41 (epilepsy and status epilepticus) *without* R56.8 (seizures) to allow PPVs of >90 % to be achieved. However, as the associated sensitivity is likely to be ~70 %, we caution that such an approach is likely to also miss many positive cases.

Second, if the resources are available to do so, our study indicates that there is additional value in linking a second administrative healthcare dataset to cause of death coding in order to help improve the overall diagnostic accuracy. Where the option is present, the second dataset should consist of a carefully selected group of AEDs more commonly used in epilepsy. When these are used in combination with cause of death codes G40–41 and/or R56.8, PPVs >90 % and sensitivities >80 % can be achieved. Whilst primary care data alone (*F25* epilepsy Read codes captured by GPs) are potentially unhelpful in identifying epilepsy cases in a deceased population, combining these with a selected group of AEDs more commonly used in epilepsy may be effective, achieving PPVs >85 % and sensitivities >90 %. There have been few studies validating the accuracy of primary care data in capturing epilepsy cases worldwide (Fonferko-Shadrach et al., 2017; Mbizvo et al., 2020; Meeraus et al., 2013; Pickrell et al., 2015; Tu et al., 2014), perhaps because of difficulties obtaining such data. Our findings help promote increased use of primary care data to help identify epilepsy cases.

Finally, we caution against using an algorithm linking in a third dataset to two already linked datasets when trying to identify deceased epilepsy cases because although this tips the balance towards a very high PPV (often 93–95 %), there are likely to be many missed positive cases and therefore unhelpfully low sensitivity. The maximum sensitivity we were able to achieve across algorithms containing three linked datasets was 71 %.

In a previous systematic review, we showed that ICD-10 codes are generally good at identifying epilepsy cases within administrative healthcare datasets of living populations, with PPVs and sensitivities frequently >80 % (Mbizvo et al., 2020). The results of our current study are supportive of this, although we caution against making direct comparisons between living and deceased epilepsy populations. More research will be needed to help further ascertain how accurately administrative healthcare data can identify epilepsy cases in deceased populations (Pickrell and Kerr, 2020).

Our study's strengths include that the sample validated were taken from a linked dataset containing virtually the entire deceased Scottish population of interest using the three national ISD datasets, supplemented by potentially the largest deceased primary care epilepsy dataset in Scotland (Herrett et al., 2015; Meeraus et al., 2013). This is also the first dedicated diagnostic test accuracy study for cause of death coding in epilepsy (Mbizvo et al., 2020; Pickrell and Kerr, 2020). Although the diagnostic accuracy of mortality coding has been extensively validated in other areas, such as in sudden cardiac death (Singh et al., 2019), there is a currently a paucity in epilepsy literature perhaps due to lack of a robust infrastructure to allow access to individual patient medical records as a diagnostic reference standard (Mbizvo et al., 2019a; Pickrell and Kerr, 2020). Our study was strengthened by national access to in-depth medical records through a supportive network of neurology colleagues. The sample was not limited to patients under specialist

neurology care as all of the available medical records were reviewed, including in cases were only non-specialist attendances to emergency care/inpatients had been made by a patient. This helped to allow the sample to be more representative of the general population, where some patients are not known to a specialist. Our study focuses on validating coding strategies used within potentially active epilepsy cases as future research into epilepsy-related mortality is likely to benefit more from ascertaining active than resolved epilepsy cases.

Our study's main limitation was being unable to provide valid negative predictive value (NPV) and specificity estimates, as all study participants were included based on having at least one epilepsy indicator across one or more databases. Therefore, any NPV and specificity estimates could not be generalised to a population without epilepsy indicators. This limitation is unlikely to have changed the external validity of our study as in a recent systematic review, we identified that NPVs and specificities in administrative epilepsy data accuracy studies are typically >90 % (Mbizvo et al., 2020), meaning our study is perhaps unlikely to have contributed any new information about these outcomes. However, our systematic review also noted that across the 30 available studies validating administrative epilepsy data globally, less than half were able to provide data on NPV and specificity due to similar problems gaining permissions or access to a disease-free cohort (Mbizvo et al., 2020). This is an obstacle that will require further consideration in future studies validating administrative epilepsy data. Another limitation of our study was failure to have all of the medical record cases reviewed in duplicate by two specialists. This was due to study resource limitation. We ameliorated this limitation by checking diagnostic interrater agreement in a 10 % sample (n = 100) of duplicate-reviewed medical records, ensuring experienced epileptologists reviewed the medical records and were also blinded to the administrative coding of participants. It remains possible that there may have been bias towards a positive gold-standard epilepsy diagnosis in cases previously managed by a neurologist versus those which were not although this is unlikely to have been significant in this study because the medical records available were extensive and included GP referral letters from primary care, secondary care records, and emergency care records, increasing the chances that the clinical information available was sufficient to allow our experienced team of consultant epileptologists to make an independent diagnosis of epilepsy in the absence of any additional neurology support. A potential limitation was failure to establish a gold-standard diagnosis in 66 participants with incomplete/missing medical records. Overall, this is unlikely to have changed our conclusions given these individuals would have formed only seven percent of the sample validated. Finally, not all countries will have the advantage of administrative datasets with national coverage as in Scotland, perhaps limiting the overall generalisability of our study. However, this is something many countries will be working towards, e.g. in Canada, Holland, and Italy, where there are already vast networks of administrative data (Franchi et al., 2013; Tu et al., 2014; Wassenaar et al., 2018), and so our study helps provide a framework for the methodological avenues researchers could use to validate such datasets.

In conclusion, our study increases confidence in the validity of using administrative data to identify epilepsy cases in mortality studies, particularly when using ICD-10-coded causes of death linked to appropriately chosen AEDs. This helps provide a context within which further work can be done to investigate the rate of epilepsy-related deaths, proportion that were potentially avoidable, and modifiable risk factors (Mbizvo et al., 2019b). Our findings also support the utility of establishing disease-based mortality registries for surveillance, such those available in the UK, US, and Canada (Donner and Devinsky, 2018; Thomas and Osland, 2020; Verducci et al., 2019). The study focuses on validating data from deceased adults and future work will be required to also assess how accurately administrative healthcare data identify epilepsy in deceased children.

Funding

This work was charitably supported by Epilepsy Research UK (R44007) and the Juliet Bergqvist Memorial Fund. The funders played no role in the design or conduct of this review.

Ethics approval, informed consent, trial registration

The study was approved by South East Scotland Research Ethics Committee 2 (IRAS 181131, 15/SS/0165), and Scottish Public Benefit and Privacy Panel for Health and Social Care.

Declaration of Competing Interest

The authors report no declarations of interest.

Acknowledgements

We are grateful to Saif Razvi (NHS Ayrshire and Arran), Myles Connor (NHS Borders), Ondrej Dolezal (NHS Dumfries and Galloway), Russell Hewett (NHS Greater Glasgow and Clyde), Linda Gerrie and Graham Mackay (NHS Grampian), Martin Zeidler (NHS Fife), Katy Murray (NHS Forth Valley), Kate Taylor (NHS Highland), John Paul Leach (NHS Lanarkshire), Kathleen White and Ian Morrison (NHS Tayside) for arranging access to medical records in the respective areas. We thank Dave Kelly and Albasoft Ltd for facilitating access to primary care data. We would like to acknowledge the support of the eDRIS Team (National Services Scotland) for their involvement in obtaining approvals, provisioning and linking data and the use of the secure analytical platform within the National Safe Haven. We are also grateful to Siddharthan Chandran and Catherine Sudlow for tactical support and advice on the project. We are grateful to Jane Andrews for administerial support in relation to this project.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.eplepsyres.2020.106462.

References

Albasoft, 2020. Albasoft Ltd.. Accessed 01/05/2020. http://www.albasoft.co.uk/.
Bossuyt, P.M., et al., 2015. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. Clin. Chem. 61 (12), 1446–1452.
Chief Medical Officer, 2018. Guidance for Doctors Completing Medical Certificates of the Cause of Death (MCCD) and Its Quality Assurance. Accessed 03/03/2020. https://bit.ly/2yyMzlC.
Death Certifcation Review Service, 2016. Annual Report 2015–2016. HTTPS://goo.gl/sBDSoU, Accessed 25/06/2020.
Donner, E.J., Devinsky, O., et al., 2018. Registries for SUDEP research. In: Panelli, R. (Ed.), Continuing the Global Conversation [Online]. SUDEP Action & SUDEP Aware.
Douglas, A., et al., 2015. Pilot study linking primary care records to Census, cardiovascular hospitalization and mortality data in Scotland: feasibility, utility and potential. J. Public Health Oxf. (Oxf).
Du, Z., Hao, Y., 2019. Package 'reportROC': An Easy Way to Report ROC Analysis. CRAN. Accessed 08/04/2020. https://bit.ly/2V3X6OA.
eDRIS Team, 2018. Researcher Guide: Requesting Outputs From Safe Haven and Disclosure Control. Accessed 16/11/2019. https://bit.ly/2XzD1Bk.
eDRIS Team, 2020. How Do We Link the Data? https://bit.ly/39DJ4ZB.
England, M.J., et al., 2013. IOM Report 2012: epilepsy across the spectrum: promoting health and understanding abstracts. Epilepsy Curr. 13 (1), 2–492.
Escalaya, A.L., et al., 2015. Epilepsy and mortality in Latin America. Seizure 25, 99–103.
Fernie, C.G.M., 2019. Scotland already has world leading mortality review system. BMJ-Br. Med. J. 364.
Fisher, R.S., et al., 2005. Epileptic seizures and epilepsy: definitions proposed by the International League Against Epilepsy (ILAE) and the International Bureau for Epilepsy (IBE). Epilepsia 46 (4), 470–472.
Fisher, R.S., et al., 2014. ILAE official report: a practical clinical definition of epilepsy. Epilepsia 55 (4), 475–482.
Fonferko-Shadrach, B., et al., 2017. Validating epilepsy diagnoses in routinely collected data. Seizure 52, 195–198.

Franchi, C., et al., 2013. Validation of healthcare administrative data for the diagnosis of epilepsy. J. Epidemiol. Community Health 67 (12), 1019–1024.
Garratt, E., Barnes, H., Dibben, C., 2010. Health Administrative Data: Exploring the Potential for Academic Research. St Andrews: Administrative Data Liaison Service.
Hamer, H.M., et al., 2012. Prevalence, utilization, and costs of antiepileptic drugs for epilepsy in Germany–a nationwide population-based study in children and adults. J. Neurol. 259 (11), 2376–2384.
Herrett, E., et al., 2015. Data Resource Profile: Clinical Practice Research Datalink (CPRD). Int. J. Epidemiol. 44 (3), 827–836.
ILAE, 1993. Guidelines for Epidemiologic Studies on Epilepsy. Epilepsia 34 (4), 592–596.
Information Services Division, 2016a. Prescribing Information System (PIS). National Data Catalogue, Scotland: ISD Scotland.
Information Services Division, 2016b. General Acute Inpatient and Day Case - Scottish Morbidity Record (SMR01). National Data Catalogue, Scotland: ISD Scotland.
Information Services Division, 2016c. National Records of Scotland (NRS) - Deaths Data, (ISD Scotland).
Information Services Division, 2020. General Practice - SPIRE accessed 03/03/2020. https://bit.ly/3eldK3B.
Jette, N., et al., 2010. How accurate is ICD coding for epilepsy? Epilepsia 51 (1), 62–69.
Katikireddi, S.V., et al., 2018. Assessment of health care, hospital admissions, and mortality by ethnicity: population-based cohort study of health-system performance in Scotland. Lancet Public Health 3 (5), 226–236.
Kee, V.R., et al., 2012. A systematic review of validated methods for identifying seizures, convulsions, or epilepsy using administrative and claims data. Pharmacoepidemiol. Drug Saf. 21, 183–193.
Kerr, M.P., 2012. 'The impact of epilepsy on patients' lives'. Acta Neurol. Scand., Suppl. 194, 1–9.
Landis, J.R., Koch, G.G., 1977. 'The measurement of observer agreement for categorical data'. Biometrics 33 (1), 159–174.
Levira, F., et al., 2016. 'Premature mortality of epilepsy in low- and middle-income countries: a systematic review from the Mortality Task Force of the International League Against Epilepsy'. Epilepsia 1–11.
Mbizvo, G.K., et al., 2019a. 'Epilepsy-related and other causes of mortality in people with epilepsy: a systematic review of systematic reviews'. Epilepsy Res. 157, 106192.
Mbizvo, G.K., et al., 2019b. The Scottish Epilepsy Deaths Study (Seds): Identifying Avoidable Epilepsy-Related Deaths. J. Neurol. Neurosurg. Psychiatr. 90 (12), E16–E17.
Mbizvo, G.K., et al., 2020. The accuracy of using administrative healthcare data to identify epilepsy cases: A systematic review of validation studies. Epilepsia 61 (7), 1319–1335.
Meeraus, W.H., et al., 2013. Childhood epilepsy recorded in primary care in the UK. Arch. Dis. Child. 98 (3), 195–202.
Middleton, O., et al., 2018. National Association of Medical Examiners position paper: Recommendations for the investigation and certification of deaths in people with epilepsy. Epilepsia 59 (3), 530–543.
Murray, C.J., et al., 2012. Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet 380 (9859), 2197–2223.
National Records of Scotland, 2020. The Medical Certificate of the Cause of Death. https://bit.ly/3aPkqnZ.
National Records Scoltand, 2016. Coding the Causes of Death. Death Certificates and Coding the Causes of Death. https://goo.gl/oWTxQU.
Nevalainen, O., et al., 2014. 'Epilepsy-related clinical characteristics and mortality: a systematic review and meta-analysis'. Neurology 83 (21), 1968–1977.
O'Malley, K.J., et al., 2005. 'MEasuring diagnoses: ICD code accuracy'. Health Serv. Res. 40 (5), 1620–1639.
Pavis, S., Morris, A.D., 2015. Unleashing the power of administrative health data: the Scottish model. Public Health Res. Pract. 25 (4), e2541541.
Pickrell, W.O., Kerr, M.P., 2020. SUDEP and mortality in epilepsy: The role of routinely collected healthcare data, registries, and health inequalities. Epilepsy Behav. 103 (Pt B), 106453.
Pickrell, W.O., et al., 2015. Epilepsy and deprivation, a data linkage study. Epilepsia 56 (4), 585–591.
Public Health Scotland, 2020. Data and Intelligence - Previously ISD Scotland. Accessed 29/04/2020. https://www.isdscotland.org/.
RCPCH, 2013. Coordinating Epilepsy Care: a UK-wide Review of Healthcare in Cases of Mortality and Prolonged Seizures in Children and Young People With Epilepsies. Child Health Reviews - UK (London: RCPCH).
RStudio Team, 2015. RStudio: Integrated Development for R, 1.2.1335. RStudio, Inc., Boston, MA.
Shackleton, D.P., et al., 2002. Survival of patients with epilepsy: an estimate of the mortality risk. Epilepsia 43 (4), 445–450.
Singh, S., et al., 2019. Diagnostic algorithms for cardiovascular death in administrative claims databases: a systematic review. Drug Saf. 42 (4), 515–527.
Tan, M., et al., 2015. Development and validation of an epidemiologic case definition of epilepsy for use with routinely collected Australian health data. Epilepsy Behav. 51, 65–72.
Thiele, C., 2020. cutpointr: Determine and Evaluate Optimal Cutpoints in Binary Classification Tasks. https://rdrr.io/cran/cutpointr/man/roc.html.
Thijs, R.D., et al., 2019. 'Epilepsy in adults'. Lancet 393 (10172), 689–701.
Thomas, R.H., Osland, K., 2020. Learnings from deaths - the Epilepsy Deaths Register. Epilepsy Behav. 103 (Pt B), 106454.
Thurman, D.J., et al., 2017. The burden of premature mortality of epilepsy in high-income countries: a systematic review from the Mortality Task Force of the International League Against Epilepsy. Epilepsia 58 (1), 17–26.
Tu, K., et al., 2014. Assessing the validity of using administrative data to identify patients with epilepsy. Epilepsia 55 (2), 335–343.
Verducci, C., et al., 2019. SUDEP in the North American SUDEP Registry: The full spectrum of epilepsies. Neurology 93 (3), 227–236.
Viera, A.J., Garrett, J.M., 2005. Understanding interobserver agreement: the kappa statistic. Fam. Med. 37 (5), 360–363.
Wassenaar, M., et al., 2018. Validity of health insurance data to identify people with epilepsy. Epilepsy Res. 139, 102–106.
Watila, M.M., et al., 2018. Overall and cause-specific premature mortality in epilepsy: a systematic review. Epilepsy Behav.

4.3 Conclusion

This chapter set out to address our aim to identify the optimal algorithms for identifying epilepsy cases in deceased adults in Scotland using administrative data. Specifically, we show that cause-of-death coding can capture epilepsy with a high degree of accuracy. The optimal algorithm involving cause-of-death coding combined ICD-10 codes G40–41 (epilepsy and status epilepticus) and/or R56.8 (seizures) with a requirement for ≥1 community-prescribed AED. This achieved PPVs >90% and sensitivities >80%, and identified 1,921 deceased individuals within the dataset. We will investigate the causes of death in these individuals in greater depth in the next chapter.

Chapter 5. A national study of epilepsy-related deaths in Scotland: trends, mechanisms, and avoidable deaths

The following content was published in *Epilepsia* in November of 2021 and has been presented here in a monograph.[204]

Student contributions: The bulk of my work in this chapter involved reviewing, cleaning and analysing the large amounts of linked data we had available for the study, and figuring out how best to interpret and present it, noting any unique trends based on my knowledge of the literature. All of these steps required statistical support. Finally, I drafted and successfully submitted the work for publication.

5.1 Introduction

Epilepsy is common and contributes to 0.7% of the global burden of diseases.[1] Of the 600,000 PWE in the UK, 60,000 live in Scotland.[52] Although it is not often discussed with patients,[205,206] epilepsy is associated with significantly increased risk of premature death.[12] Some of those deaths may be unrelated to epilepsy. However, a substantial proportion relate to the epilepsy itself, its treatment, or comorbidities worsened by epilepsy.[12,23] Together, these are termed 'epilepsy-related deaths', and they are likely to contribute a significant burden of total years of potential life lost.[12,23] A nationwide study quantifying epilepsy-related deaths is yet to be undertaken in Scotland.[12] From a public health perspective, this is a pressing issue as Scotland has a lower life expectancy than the rest of the UK and Western Europe.[53] It also has a larger proportion of its population living in rural and remote communities compared

with the rest of the UK.[53] This inevitably presents challenges in epilepsy healthcare provision.[43]

More generally, the exact mechanisms by which epilepsy-related deaths occur remain unclear,[12,23] often making it difficult for clinicians to have a broad discussion about these with patients.[205,206] A recently proposed classification system for epilepsy-related deaths suggests these should include deaths occurring directly due to epilepsy (including SUDEP and seizure-related accidents), deaths occurring due to acute symptomatic seizures that would likely have gone on to become epilepsy (including after stroke or traumatic brain injury), deaths indirectly occurring due to epilepsy (including aspiration pneumonia or suicide), and deaths occurring due to the underlying aetiology of epilepsy (including brain tumours and genetic conditions).[23] A way to test for the presence all of these mechanisms in real-world data is to conduct an MCOD analysis on deaths involving epilepsy.[30,31] To our knowledge, this is yet to be undertaken.[12] Traditionally, cause of death statistics are derived from the underlying cause of death on a death certificate.[29-31] This is defined as the disease or condition which initiated the chain of morbidities and mechanisms that lead directly to death.[31] MCOD analysis differs by including not only the underlying cause of death, but also all of the other causes of death listed as contributory morbidities and mechanisms in the death certificate.[29,30] Reporting MCOD data, advocated by the World Health Organisation, would allow a more complete clinical picture to be drawn of all of the mechanisms implicated in epilepsy-related deaths.[29-31]

It is thought many epilepsy-related deaths may be avoidable,[15,39,40] although this is yet to be quantified.[12] The ONS has recently produced a list of underlying causes of death which should not, or should only infrequently, have given rise to death in the presence of timely and effective healthcare or public health policies.[45] This Revised Definition of Avoidable Mortality list (RDAM 2016) is based on research from Nolte (2004)[207] and Page (2006),[208] and has been refined through public consultation with multiple organisations including UK government, academics and charities.[45] The list identifies conditions for which there are effective public health and primary prevention strategies before the onset of disease (e.g. immunisation) to reduce the incidence of death, or conditions for which, after the onset of disease, effective treatments are available to reduce case fatality.[209] For most causes, there is an upper age limit of 74 years.[45] The RDAM 2016 is used by the UK government and The Department of Health and Social Care to determine progress in reducing the prevalence of preventable ill-health, premature death, and the gap between local authorities.[209] In Scotland, a seminal paper in The Lancet evaluated nationwide health system performance through 2001–2013 using RDAM 2016 to illustrate that during the 50.5 million person-years of study, there had been 166,245 potentially avoidable deaths.[43] While a specific condition can be considered avoidable, this does not mean that every death from that condition could be averted. Instead, this measure is designed to highlight areas of potential weaknesses in healthcare that could benefit from further in-depth investigation.[209]

We report the first national study of epilepsy-related deaths in Scotland. We aimed to first establish the total number of epilepsy-related deaths in the country, whether these have changed over time, and how they compare to risk of mortality in the general population by age. We then aimed to describe all of the mechanisms by which epilepsy-related deaths had occurred in the age groups at increased risk (using MCOD data), and to estimate what proportion of these deaths were potentially avoidable (using the RDAM 2016 list).

5.2 Methods

5.2.1 Setting

All individuals accessing healthcare in Scotland are provided with a unique identification number (CHI). This can, therefore, be used to create nationally linked research datasets derived from the entire Scottish population.[43] All deaths are registered centrally at NRS within eight days of occurrence (circa. 55,000 annually).[59] The free-text causes of death listed on death certificates are coded by NRS into their corresponding ICD-10 codes.[59] There is a world-leading mortality review system in place in Scotland in which these causes of death are routinely scrutinised for accuracy and amended where necessary (summarised in Figure 5-1).[60,61] Deaths in which a cause cannot be immediately identified because they are sudden, suspicious, accidental or unexplained are reported to a statutory legal official (the Procurator Fiscal).[210] The Procurator Fiscal reviews the history and circumstances of death and either accepts the cause if satisfied it was natural, or arranges an autopsy. Suspected SUDEP would normally proceed to autopsy in Scotland.[211]

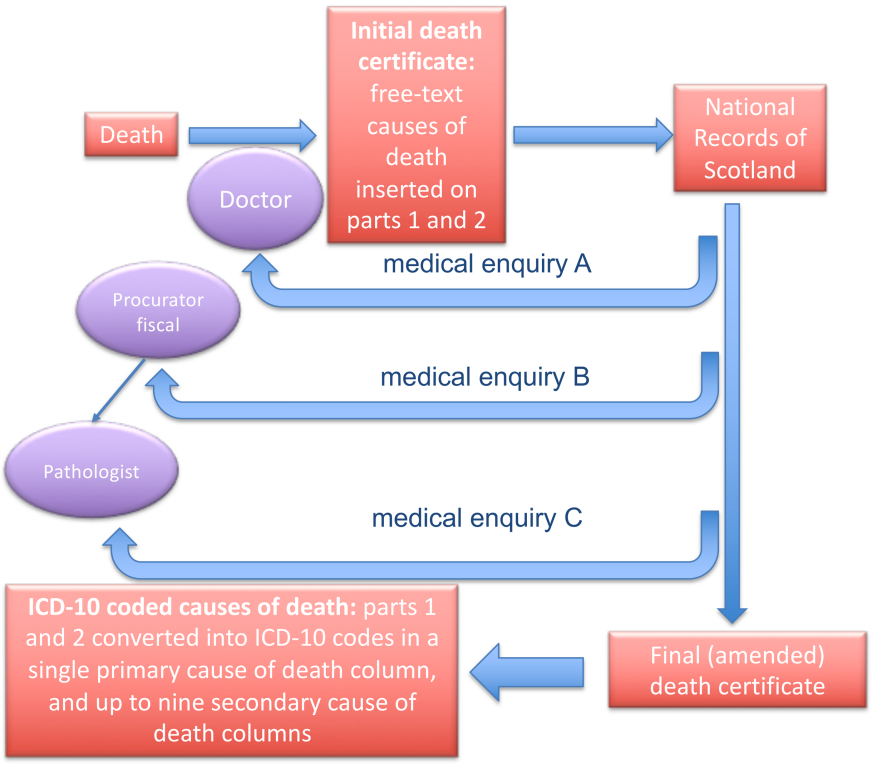


Figure 5-1: Flow diagram summarising the National Records of Scotland (NRS) death certification and cause of death coding process

Figure legend: The initial death certificate is completed by the treating physician. Part 1 lists the immediate cause of death and works back in time by listing the diseases and conditions that lead to this (Parts 1A–D). Part 2 lists other significant diseases and conditions that contributed to the occurrence of the death (i.e. this is not simply a list of comorbidities). The underlying cause of death is taken from either part of the death certificate as appropriate. Codes for each death are checked by an NRS Vital Events Medical Coder to ensure their electronic software has allocated the correct codes and taken proper account of the details of the cause of death. Subsequently, additional information about a particular death may then be sought from e.g. general practitioners, hospital physicians, and the Procurators Fiscal (a statutory legal official that investigates deaths that are accidental, unexpected, unexplained, sudden or suspicious (including ordering a post mortem, where necessary)).[60] This may result in changes to the cause of death codes. The software is used to allocate revised codes, based on the amended details of the cause of death, which are then checked and, if necessary, amended by a Medical Coder. Double-coding is a process where an experienced Medical Coder reviews all the information received in relation to a particular death, and checks that the correct cause-of-death codes have been allocated. Where there is a difference of opinion, the person double-coding the death record discusses the case with the original Coder and they come to an agreement on the final cause of death codes.[60] All Scottish deaths registered since 2009 have been double coded.[60]

Abbreviations: ICD-10 – International Classification of Diseases 10

5.2.2 Study design, data sources and participants

This was a national administrative data linkage study. Using CHI numbers, Scotland-wide person-level data were retrospectively linked over seven consecutive years (01/01/09–01/01/16) in order to identify all of the epilepsy-related deaths in adults (aged ≥16 years) during this period. We first linked NRS death records[59] to Prescribing Information System (PIS) Scottish prescribing data.[190] PIS is a mandatory national prescribing dataset used to track the community drug history of the entire Scottish population in order to facilitate public drug reimbursement.[190] Epilepsy-related deaths were defined as those in which death had occurred with at least one ICD-10-coded cause listed as G40 (epilepsy), G41 (status epilepticus), or R56.8 (seizures) in any position within the NRS cause of death record, combined with a requirement for such individuals to have also had at least one community-prescribed AED in their PIS dataset. This combined NRS and PIS algorithm demonstrated a high PPV of 91% (95% confidence interval (CI) 89–94%), and a high sensitivity of 81% (95% CI 77–84%) for correctly identifying a diagnosis of epilepsy in Chapter Four.[62,212] It was also consistent with national guidance that the presence of an epilepsy code anywhere in the cause of death record indicates that epilepsy was thought by the certifying doctor to be either part of the sequence of events leading to death or else a factor contributing to the death.[31,213] Data on autopsy status (performed or not), place of death, sex, and Scottish Index of Multiple Deprivation (SIMD) quintiles were automatically uploaded from NRS.[57,59,214]

5.2.3 Enrichment linking

We enriched the NRS and PIS data by linking it to nationwide hospital admissions[63] and outpatient attendance data,[64] as well as to regional primary care data (see Panel A).[57]

Panel A

Identification of epilepsy-related deaths = A + B

A. **NRS death records**:[59] those who had died with ≥1 G40–41 ICD-10 code (epilepsy and status epilepticus) and/or R56.8 (seizures) listed within their causes of death.
B. **AEDs on the Scottish Prescribing Information System (PIS)**:[190] those who were prescribed ≥1 AED (regardless of their cause of death).

***Enrichment linking* = A + B ± *C and/or D* ± E**

C. **Scottish Morbidity Record 01 (SMR01) hospital admissions**:[63] those who had experienced ≥1 hospital admission diagnosis coded as G40–41 and/or R56.8, regardless of their cause of death. SMR01 is a mandatory national dataset capturing all of the inpatient hospital admissions in Scotland. 1.4 million SMR01 records are generated annually.[63] Data collected include a mandatory diagnostic field containing up to six ICD-10-coded hospital admission diagnoses.[63]
D. **Primary care:** those who had ≥1 epilepsy-related (Read code *F25*) primary care attendance (regardless of cause of death). Read codes are a coded thesaurus of clinical terms used in primary care to create electronic health records. General practices (GPs) use Vision medical software to build electronic health records (EHRs) for patients in primary care. Within this system, GPs record symptoms, diagnoses and prescriptions using Read codes.[126] This makes it possible to track diagnostic primary care NHS data within a data linkage study environment. As there is no routine primary care epilepsy dataset available in Scotland currently, we established one by putting together a convenience sample of 10% of all 900 Scottish GP practices using methods published in detail previously (Appendix S1).[212] We extracted all other coded diseases and activities occurring for the *F25* patients in order to gain insight into the burden of comorbidity from mental health problems (including depression, psychiatric illness and suicidality), learning disabilities and congenital abnormalities, and substance misuse (including alcohol and drug abuse). Additionally, for primary care data, we were able to estimate a duration of epilepsy as the time between the first ever *F25* code and the date of death (see Appendix S2–3 for full list of Read codes used). The baseline characteristics of those with primary care data were compared to the rest of the group to assess their representativeness. For the categorical baseline characteristics, comparison was made using a chi-square independence test alongside post-hoc *z*-tests for independent proportions, applied with Bonferroni correction. For continuous variables, an independent samples t-test was used. A 2-sided 5% significance cut-off was used throughout.
E. **SMR00 outpatient attendance:** We linked participants to their SMR00 dataset in order to flag which had attended neurology clinic during the study period. SMR00 is a national outpatient activity dataset in which it is mandatory to record whether or not patients attended clinic and under which speciality/discipline, allowing hospitals to be reimbursed for providing clinic services.[64] Neurology clinic attendances are coded as AH (AH1 = patient was seen, AH5 = patient attended but was not seen (could not wait),

AH8 = patient did not attend (DNA)).[64] This makes it possible for us to track whether or not deceased patients had previously attended neurology clinics anywhere in the country during the study period. We excluded the outpatient diagnosis field as this is poorly recorded due to being optional.[64]

5.2.4 Variables and data sources

5.2.4.1 Identifying the mechanisms of epilepsy-related death using MCOD data

We created a MCOD dataset by pooling together all of the ICD-10-coded causes of death listed alongside G40–41/R56.8 in the NRS dataset.[30,31] As there is no specific ICD-10 code for SUDEP,[119] we captured SUDEP from the free-text causes of death section of the NRS records, where SUDEP deaths were identified as those with "sudden unexplained/unexpected death in epilepsy" listed in the free-text cause of death section. We identified *definite* SUDEP as those that underwent autopsy, and *probable* SUDEP as those without autopsy.[24] Given our access to MCOD data, we were able to identify SUDEP *Plus* as those with an additional condition(s) listed as a cause of death alongside SUDEP.[24] These definitions were consistent with the Nashef 2012 classification of *definite*, *probable*, and SUDEP *Plus*.[24] Given that coding all free-text causes of death is mandatory, but there is no specific ICD-10 code for SUDEP,[119] we took the opportunity to extract all of the ICD-10 codes that were listed alongside free-text causes of death listed as SUDEP in order to help explore which codes in the ICD-10 system are being used in cases of SUDEP.

5.2.4.2 Identifying potentially avoidable epilepsy-related deaths

Table 5-5 lists the underlying causes of death used by the ONS to indicate a potentially avoidable death in the RDAM 2016.[43,45] Consequently, we identified

the potentially avoidable epilepsy-related deaths as any in which the underlying cause of death was in this RDAM 2016 list.[43,44]

5.2.5 Bias and study size

Selection bias was reduced by examining data from the entire Scottish population,[43,59,63,190] also maximising sample size for observation. Misclassification bias was reduced by using the optimal case-ascertainment codes identified during prior validation,[62,212] cause of death records with a central quality assurance process in place,[60,61] and an established list of ICD-10 codes produced by the ONS for identifying potentially avoidable mortality.[43-45]

5.2.6 Statistical analysis

5.2.6.1 Age-standardised mortality rate (MR)

We reported the total number of epilepsy-related deaths in the study period. We applied a direct method of determining age-standardised MR using the 2013 European Standard Population (ESP)[215,216] to allow for differences in the age structure of populations and allow valid comparisons to be made between geographical areas and through time.[216] We calculated age-standardised MR in each year of the study period using $\frac{\sum(P_\kappa \mathrm{m}_\kappa)}{\sum P_\kappa}$,[216] where P_κ = Standard Population in age group κ, m_κ = observed MR (observed epilepsy-related deaths per 100,000 Scottish persons) in age group, and κ = age group 16–24 years, 25–34 years, 35–44 years, 45–54 years, 55–64 years, 65–74 years, and 75+ years. The yearly Scottish population size was extracted from NRS annual mid-year population estimates.[217] We applied a linear regression model of epilepsy-related MR plotted against time (in years) to assess for significant

change over the study period (indicated by change in MR (ΔMR) with a 95% CI not crossing 0). Primary analysis looked at age-standardised MR overall. A subgroup analysis was reported by sex.

5.2.6.2 Age-specific mortality ratios

To compare epilepsy-related mortality to mortality in the general Scottish population, we used an indirect method to estimate age-specific mortality ratios by age group κ.[218] Annual mortality figures for the general Scottish population were extracted from NRS Death Time Series Data.[219] Annual Scottish epilepsy population figures were extracted from national ISD primary care releases where available (2009–13),[220] and estimated using Scottish NHS Board prevalence figures (0.8%) in the remaining years.[52] Annual change in mortality ratios was assessed using linear regression. Primary analysis looked at the overall age-specific mortality. A subgroup analysis was reported by sex.

5.2.6.3 MCOD data

MCOD analysis was conducted for the age groups in which SMR was significantly increased. The deaths were split into SUDEP and non-SUDEP epilepsy-related deaths. We reported the total number of SUDEP cases, identifying what proportion were *definite*, *probable*, and SUDEP *Plus*.[24] For the non-SUDEP deaths, we counted the number of times each individual ICD-10 code and chapter group was listed as a cause of death, listing each chapter frequency as a percentage of the total code frequency.[221]

5.2.6.4 Avoidable deaths

We reported the total number of avoidable epilepsy-related deaths and grouped these according to the RDAM 2016 disease groups to which they

belonged (Table 5-5),[43,45] reporting these as percentages of the total number affected.

Data were processed using RStudio Version 1.2.1335, Boston, MA. 95% CIs were listed, where relevant, including using an exact binomial test for proportional estimates. Categories with less than five events/participants were combined or reported as '<5' to avoid patient identification.

5.2.7 Approvals

This study was approved by South East Scotland Research Ethics Committee 2 (15/SS/0165, IRAS 181131), and the Scottish PBPP.

5.3 Results

5.3.1 Frequency and demographics of epilepsy-related deaths

1,921 epilepsy-related deaths were identified across Scotland through 2009–16. This was from nationwide NRS and PIS data combined using the validated algorithm ≥1 G40–41 and/or R56.8 cause of death coded in NRS + ≥1 community-prescribed AED in PIS.[212] Baseline characteristics of the 1,921 people are summarised Table 5-1. The deaths mostly occurred in the largest four Scottish Health Board regions, at a median age of 67 years (range 16–101 years), with 53% of the 1,921 deaths occurring in men. Most deaths occurred in an NHS hospital (47% of the 1,921 deaths) or at home (29% of the 1,921 deaths), and the majority had lived in more deprived areas of Scotland (Figure 5-2).[214] Whilst 1,185/1,921 people who went on to suffer an epilepsy-related death had experienced at least one seizure- or epilepsy-related hospital admission during the study period (i.e. 62%), only 518/1,921 had been seen in a neurology clinic during that same period (i.e. 27%). 360/518 seen in

neurology clinic were from the 1,185 with at least one seizure- or epilepsy-related hospital admission (i.e. 30% of the 1,185 people).

Table 5-1: Characteristics of epilepsy-related deaths population

Scottish NHS Health Board *(NRS data)*	***Total number of epilepsy-related deaths 2009–16***	***Average yearly general Scottish population size 2009–16***
NHS Greater Glasgow & Clyde	451	956,927
NHS Lanarkshire	259	536,332
NHS Lothian	254	721,051
NHS Grampian	198	488,611
NHS Ayrshire & Arran	155	308,405
NHS Tayside	144	346,970
NHS Fife	137	303,999
NHS Forth Valley	116	249,793
NHS Highland	112	267,752
NHS Dumfries & Galloway	48	125,994
NHS Borders	33	95,055
NHS Orkney, NHS Shetland, & NHS Western Isles	14	40,903
Total	1,921 (1,010 male)	
Median age at death (years) *– NRS data*	67 (range 16–101)	
Place of death *– NRS data*	**Total**	**Proportion (95% CI)**
NHS hospital	898	46.7% (44.5–49.0%)
Non-Institutional, e.g. at home	562	29.3% (27.2–31.3%)
Private nursing home/private hospital	377	19.6% (17.9–21.5%)
Home for the elderly	57	3.0% (2.3–3.8%)
Miscellaneous (GP surgery, clinic, prison, contractual hospital, other home)	27	1.4% (0.9–2.0%)
Cause of death enquiries *– NRS data*		
Medical enquiry into cause of death	189	9.8% (8.5–11.3%)
Procurator Fiscal input to cause of death	466	24.3% (22.4–26.2%)
Post mortem input to cause of death	356	18.5% (16.8–20.3%)
Clinical care indicators *– linked SMR01, SMR00, PIS data*		
Seen in neurology clinic ≥ once during 2009–16 (≥1 AH1 code in SMR00)	518	27.0% (25.0–29.0%)
≥1 seizure-/epilepsy-related hospital admission during 2009–16 (≥ 1 G40–41 and/or R56.8 code in SMR01)	1,185	61.7% (59.5–63.9%)
≥1 seizure-/epilepsy-related hospital admission during 2009–16 (≥ 1 G40–41 and/or R56.8 code in SMR01) and also seen in neurology clinic ≥ once during 2009–16 (≥1 AH1 code in SMR00)	360	

Abbreviations: NRS – National Records of Scotland; SMR01 – Scottish Morbidity Record (01 = hospital admissions, 00 = outpatient clinic attendance); PIS –Prescribing Information System; AH1 – patient was seen neurology clinic attendance; G40–41 – ICD-10 codes for epilepsy; R56.8 – ICD-10 code for seizure; NHS – National Health Service; AED – antiepileptic drug.

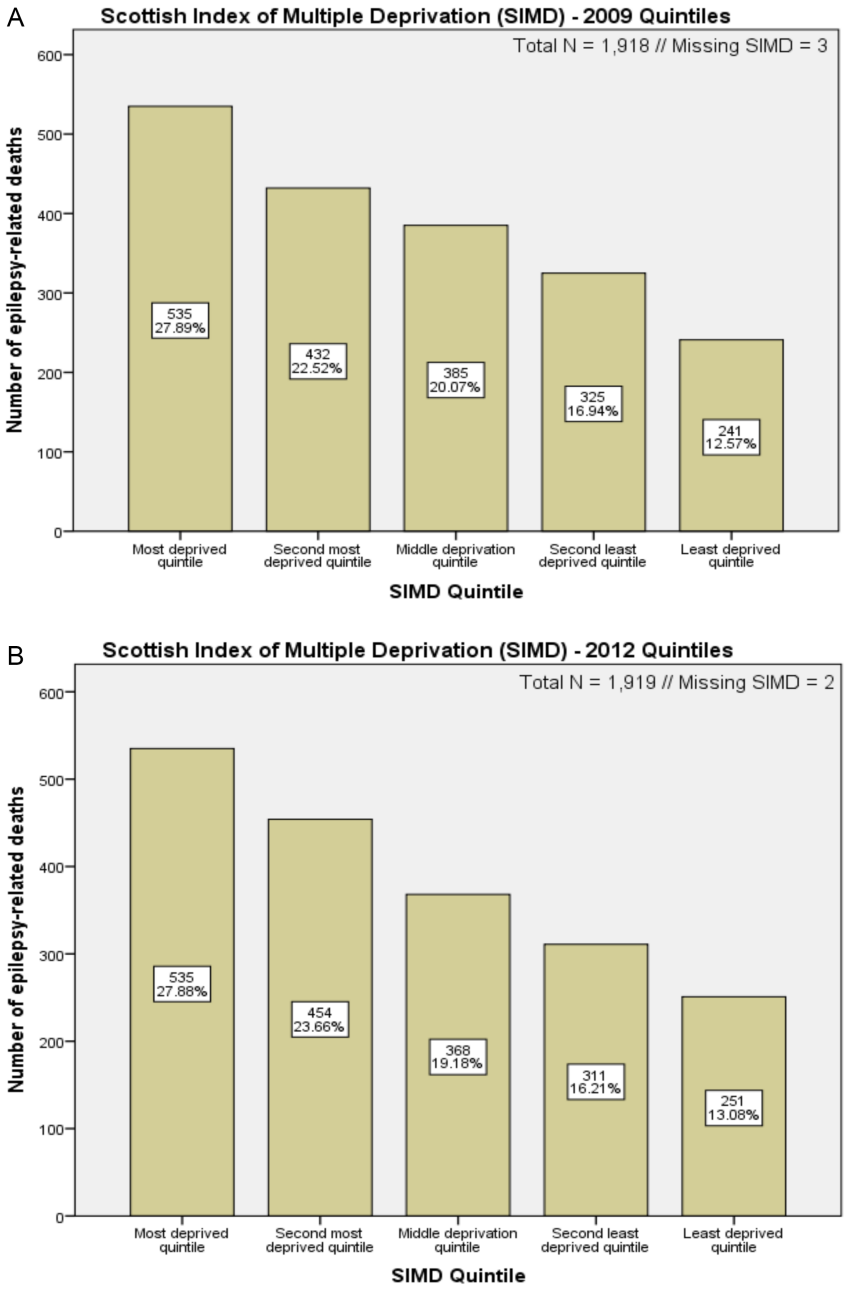


Figure 5-2: Bar chart of Scottish Index of Multiple Deprivation (SIMD) Quintiles

Figure legend: The SIMD quintiles split the Scottish population by socioeconomic deprivation, ranked from 1 = most deprived to 5 = least deprived areas of residence.[214] The quintiles were updated in 2009 (2A) and 2012 (2B).

5.3.2 Annual trend in epilepsy-related death MR

Figure 5-3A is a linear regression plot of annual age-standardised epilepsy-related MR per 100,000 Scottish population.[216] The overall MR ranged between 5.9–8.7/100,000 during the study period, demonstrating no significant change over time (95% CI -0.05–0.66/100,000 for annual ΔMR – calculated in Appendix S4a).

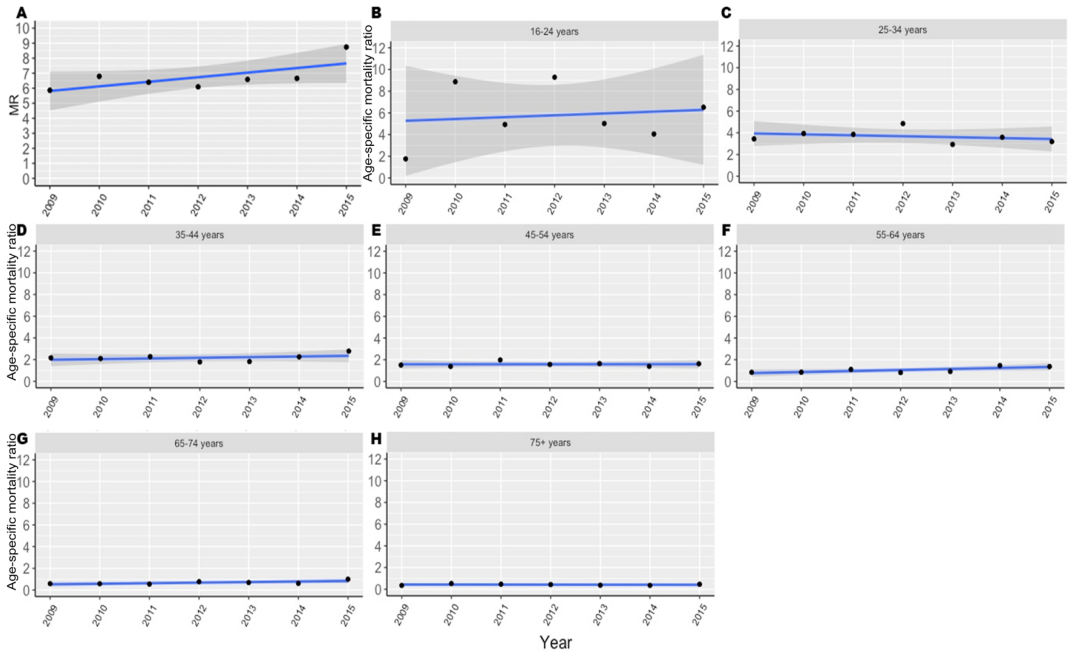


Figure 5-3: Linear regression plots of mortality rate and age-specific mortality ratios overall

Figure legend: Each dot plot is fitted with linear regression line and confidence interval (grey shadow). 5-**3A** – Direct age-standardised mortality rate (MR) per 100,000 Scottish population each year. Figure illustrates little change in epilepsy-related mortality rate over time. 5-**3B–H** – Annual age-specific mortality ratios for epilepsy-related death in each age group of study. This figure illustrates little change over time, with persistently increased estimates in those aged 16–54 years.

In men, MR ranged between 6.6–10.1/100,000 men, demonstrating no significant change over time (95% CI -0.19–0.80/100,000 for annual ΔMR, Appendix S4b). In women, it ranged between 5.2–7.3/100,0000 women, also demonstrating no significant change over time (95% CI -0.009–0.544/100,000

for annual ΔMR, Appendix S4c). Although there was a tendency for MRs to be slightly higher in men than women, the differences were largely abolished by overlapping CIs (Figure 5-4).

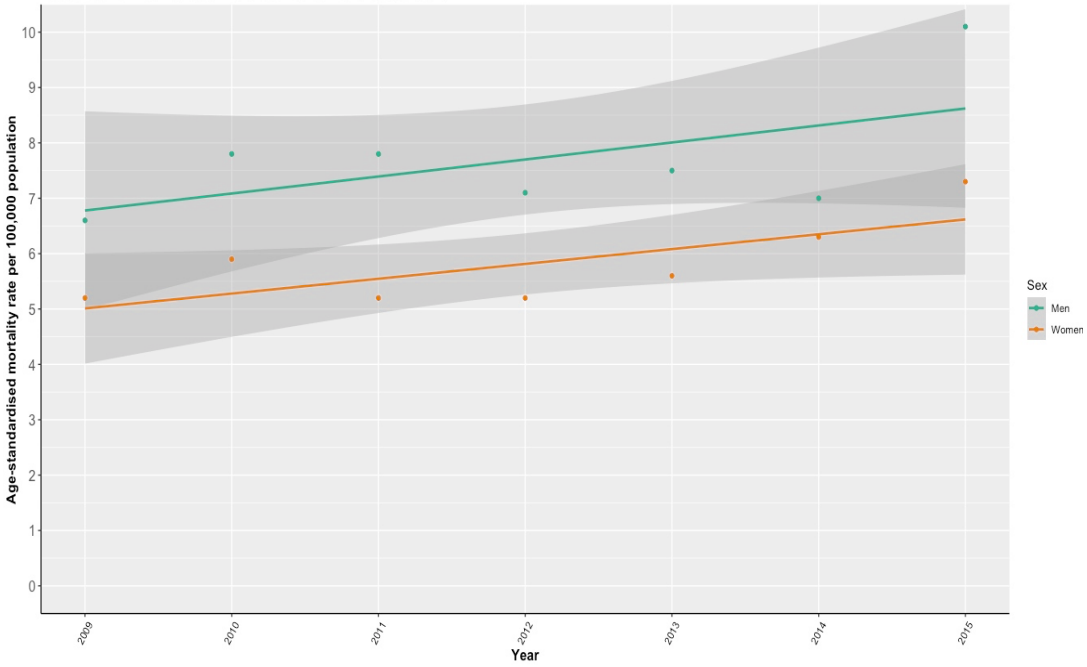


Figure 5-4: Linear regression plots of direct age-standardised mortality rate per 100,000 population each year according to sex

Figure legend: The dot plot is fitted with linear regression line and confidence interval (grey shadow). The figure illustrates little change in epilepsy-related mortality rate over time in both groups.

5.3.3 Trend in age-specific mortality ratios

Figure 5-3B–H is a panel of linear regression plots for age-specific mortality ratios for epilepsy-related deaths each year. These increased significantly over the study period for those aged 55–64 years (95% CI 0.00065–0.19 for annual change in mortality ratio), but otherwise remained unchanged in all remaining age group strata (calculated in Appendix S4d). Trends in age-specific mortality ratios were similar between men and women (Figure 5-5 and Appendix S4e–f).

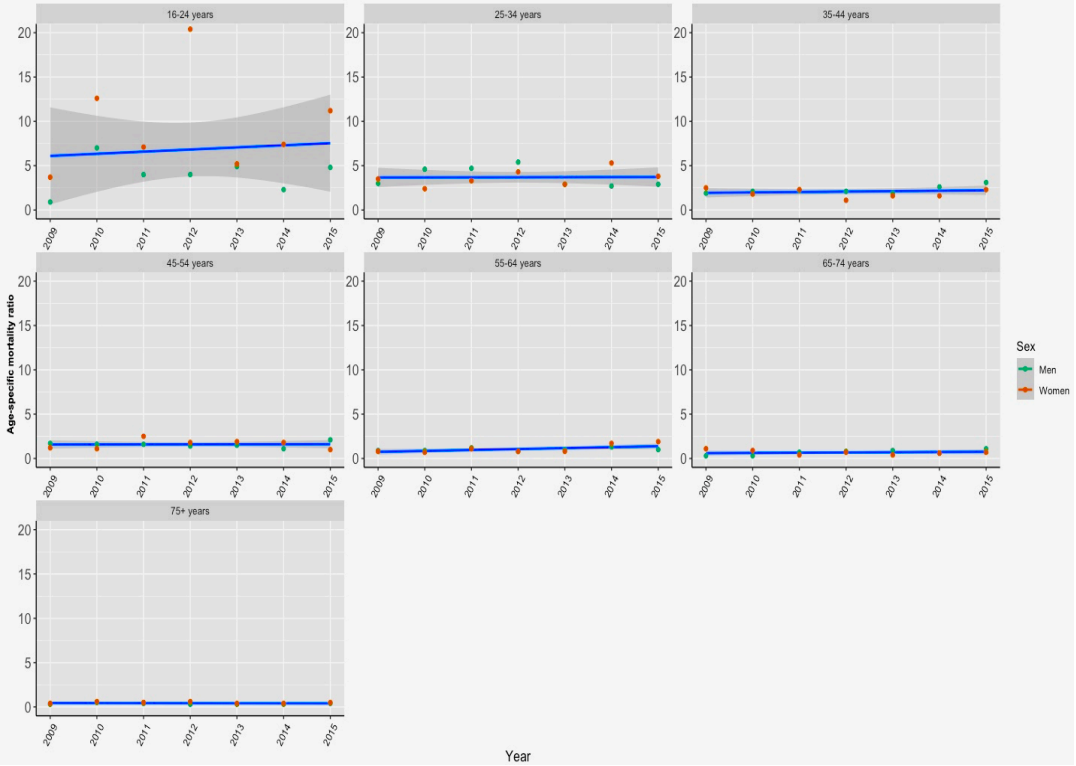


Figure 5-5: Linear regression plots of age-specific mortality ratios each year according to sex

Figure legend: The dot plot is fitted with linear regression line and confidence interval (grey shadow). The figure illustrates little change in epilepsy-related mortality rate over time in both groups.

Table 5-2 summarises the mean age-specific mortality ratios for epilepsy-related deaths throughout the study period, demonstrating that these were significantly more common in young adults aged 16–54 years, peaking at 5.3 (95% CI 1.8–8.8) in age groups 16–24 years, followed by 3.6 (95% CI 1.8–5.4) in age groups 25–34 years. The mean overall mortality ratio for ages 16–54 years remained significantly raised at 2.3 (95% CI 1.8–2.8), and was lower than in the general population in those aged ≥55 years (0.52, 95% CI 0.45–0.60).

Table 5-2: Comparative risks of epilepsy-related death

Age groups	Mean national Scottish mortality rate per 100,000 population per year 2009–16	Mean size of the Scottish epilepsy population each year 2009–16	Expected mean number of deaths in the Scottish epilepsy population each year 2009–16	Observed mean number of epilepsy-related deaths per year in Scotland 2009–16 (study population)	Age-specific mortality ratio (95% CI)	
					By age-group κ	Overall
16-24 years	45	3,740	2	9	5.3 (1.8–8.8)	
25-34 years	87	4,908	4	16	3.6 (1.8–5.4)	2.3 (1.8–2.8) in those aged 16–54 years
35-44 years	173	5,408	9	20	2.2 (1.2–3.1)	
45-54 years	340	6,324	21	34	1.6 (1.1–2.1)	
55-64 years	806	5,126	41	43	1.0 (0.7–1.4)	
65-74 years	2,016	3,927	79	54	0.7 (0.5–0.9)	0.52 (0.45–0.60) in those aged ≥55 years
75+ years	8,142	2,889	235	98	0.4 (0.3–0.5)	

Abbreviations: CI – Confidence interval

5.3.4 Mechanisms involved with epilepsy-related death

Epilepsy-related deaths were significantly increased in young adults aged 16–54 years when compared to age-mated mortality figures in the general population, occurring in 553 young adults in the study. 349/553 were male (i.e. 63%). 62/553 (i.e. 11%) had primary care Read code data available. Mean age in those with primary care data was 40.2 years (±11.3 years SD). This was no different to a mean age of 39.7 years (±10.5 years SD) in the remaining 491 young adults without primary care data (P = 0.74). The proportions of men and women were also the same in both groups (P = 0.60). There was also no difference in proportions living in each SIMD quintile area of deprivation between both groups (P = 0.63), proportion with previous epilepsy-related hospital admission (P = 0.63), and proportion previously seen in a neurology clinic (p = 0.97) – see Table 5-3. The primary care data revealed a median epilepsy duration of 17 years (range 1–55 years). There were Read code indicators of comorbid mental health problems in 50% of the 62 participants,

alcohol and drug abuse in 44% of the 62 participants, and learning disabilities or congenital malformations in 39% of the 62 participants.

Table 5-3: Comparison of the baseline characteristics of young adults who had primary care data available to the remaining cohort of young adults without primary care data

Variable			*Has primary care data*			
			No	**Yes**	**Total**	
Sex (NRS data)	Male	Count	308a	41a	349	Pearson Chi-square p = 0.60
		% within columns	62.7%	66.1%	63.1%	
	Female	Count	183a	21a	204	
		% within columns	37.3%	33.9%	36.9%	
	Total	Count	491	62	553	
		% within columns	100%	100%	100%	
SIMD 2012 (NRS data)	1 - Most deprived quintile	Count	172a	16a	188	Pearson Chi-square p = 0.63
		% within columns	35.0%	25.8%	34.0%	
	2 - Second most deprived quintile	Count	116a	16a	132	
		% within columns	23.6%	25.8%	23.9%	
	3 - Third most deprived quintile	Count	92a	12a	104	
		% within columns	18.7%	19.4%	18.8%	
	4 - Second least deprived quintile	Count	66a	10a	76	
		% within columns	13.4%	16.1%	13.7%	
	5 - Least deprived quintile	Count	45a	8a	53	
		% within columns	9.2%	12.9%	9.6%	
	Total	Count	491	62	553	
		% within columns	100%	100%	100%	
≥1 seizure- or epilepsy-related hospital admission in the study period	No	Count	214a	29a	243	Pearson Chi-square p = 0.63

(SMR01 data)						
		% within columns	43.6%	46.8%	43.9%	
	Yes	Count	277a	33a	310	
		% within columns	56.4%	53.2%	56.1%	
	Total	Count	491	62	553	
		% within columns	100%	100%	100%	
Seen in neurology clinic during the study period (SMR00 data)	No	Count	268a	34a	302	Pearson Chi-square p = 0.97
		% within columns	54.6%	54.8%	54.6%	
	Yes	Count	223a	28a	251	
		% within columns	45.4%	45.2%	45.4%	
	Total	Count	491	62	553	
		% within columns	100%	100%	100%	

Key: *Each subscript letter "a" denotes a pairwise comparison for which the proportions do not differ significantly from each other at the 0.05 level.*

Abbreviations: *NRS – National Records of Scotland (mortality data); SMR01 – Scottish Morbidity Record 01 (hospital admissions data); SMR00 – Scottish Morbidity Record 00 (outpatient hospital attendance data)*

166/553 young adult epilepsy-related deaths were SUDEP (i.e. 30%). 165/166 SUDEP cases underwent investigation by the Procurator Fiscal (i.e. 99%). 149/166 cases (i.e. 90%) were confirmed at autopsy, regarded as *definite* SUDEP.[24] The remaining 17/166 cases (i.e. 10%) did not have an autopsy, regarded as *probable* SUDEP.[24] An additional cause of death was listed alongside SUDEP in 30/166 SUDEP cases (i.e. 18%), regarded as SUDEP *Plus*.[24] 23/30 SUDEP *Plus* cases (i.e. 77%) had autopsy confirmation. 107/166 SUDEP cases were male (i.e. 64%). 144/166 SUDEP cases occurred at home (i.e. 87%), with the remaining small minority occurring in an NHS hospital or other setting. 68/166 SUDEP cases (i.e. 41%) had experienced at least one

seizure- and/or epilepsy-related hospital admission during the study period. A similar number of 79/166 SUDEP cases (i.e. 48%) had been seen at least once in a neurology clinic during the same period. 60% of the 166 SUDEP cases had lived in the two most deprived SIMD quintile areas.[214] All SUDEP cases were identified from the free-text cause-of-death fields containing "sudden unexpected death in epilepsy" within the NRS record. As going on to code these free-text sections is mandatory, but there is no ICD-10 code for SUDEP,[119] it was curious to note that coders universally defaulted to using the G40.9 code for 'epilepsy unspecified' to code these SUDEP free-texts.

Table 5-4 summarises the MCOD data for the 387 young adult epilepsy-related deaths that were not SUDEP. In total, there were 1,365 coded disease mechanisms underlying and contributing to death in this population. By definition, each of these individuals had to have at least one of these causes coded as G40–41/R56.8.[30,31,213] These codes contributed 30% of the 1,365 coded disease mechanisms. By far the most commonly used individual code within this group was G40.9 (epilepsy unspecified), highlighting that this code was not exclusive to SUDEP free-texts. R56.8 (seizures) and G41.9 (status epilepticus) were reported far less frequently. The five most common cause of death mechanisms to accompany epilepsy were J00–J99 diseases of the respiratory system (mainly aspiration pneumonia (J69.0)), I00–I99 diseases of the circulatory system (mainly cardiac arrest (I46.9)), G00–G37 and G42–G99 other non-epilepsy diseases of the nervous system (mainly cerebral palsy (G80.9)), S00–T98 injuries, poisoning, and other external causes (mainly as AED poisoning (T42.6) or drowning (T75.1)), and F00–F99 deaths in relation

to mental and behavioural disorders (mainly in relation to alcohol dependence (F10.2)). The next most common were V01–Y98 external causes of morbidity and mortality (mainly narcotic poisoning (X42)), Q00–Q99 congenital malformations (mainly Down syndrome (Q90.9)), and K00–K93 diseases of the digestive system (mainly alcoholic liver disease (K70.9)). Fatal neoplasms were mainly of the brain (C71.9). 194/387 non-SUDEP epilepsy-related deaths (i.e. 50%) underwent investigation by the Procurator Fiscal, indicating that they were likely to have been accidental, unexpected, unexplained, sudden or suspicious. Of the 194 that underwent investigation, 106 underwent autopsy (i.e. 55%).

Table 5-4: Multiple Cause of Death data (MCOD) in the non-SUDEP epilepsy-related deaths population of 387 cases

ICD-10 chapters	ICD-10-coded causes of death as chapters	Number of times a code from this chapter was used as a cause of death (as a percentage of the total, 95% CI)	ICD-10-coded causes of death as individual codes within chapters	Number of times each individual code was used as a cause of death
VI., XVIII.	**G40–41, R56.8:** Epilepsy, status epilepticus and seizures	404 (29.6%, 27.2–32.1%)	G40.9 – Epilepsy, unspecified	311
			R56.8 – Seizures	46
			G41.9 – Status epilepticus, unspecified	22
			G40.3 – Generalised idiopathic epilepsy and epileptic syndromes	8
			G40.5 – Special epileptic syndromes	6
			G40.6 – Grand mal seizures unspecified (with or without petit mal)	6
X.	**J00–J99:** Diseases of the respiratory system	159 (11.6%, 10.0–13.5%)	J69.0 – Pneumonitis due to food and vomit	53
			J18.9 – Pneumonia, unspecified	35
			J18.0 – Bronchopneumonia, unspecified	24
			J98.8 – Other specified respiratory disorders	9
			J44.9 – Chronic obstructive pulmonary disease, unspecified	8
			J45.9 – Asthma, unspecified	7
			J22 – Unspecified acute lower respiratory infection	6
IX.	**I00–I99:** Diseases of the circulatory system	128 (9.4%, 7.9–11.0%)	I46.9 – Cardiac arrest, unspecified	21
			I25.9 – Chronic ischaemic heart disease, unspecified	15
			I25.1 – Atherosclerotic heart disease	10
			I64 – Stroke, not specified as haemorrhage or infarction	10
			I21.9 – Acute myocardial infarction, unspecified	9
			I26.9 – Pulmonary embolism without mention of acute cor pulmonale	5
			I51.7 – Cardiomegaly	5
VI.	**G00–G37, G42–G99:** Other diseases of the nervous system	102 (7.5%, 6.1–9.0%)	G80.9 – Cerebral palsy, unspecified	36
			G93.1 – Anoxic brain damage not elsewhere classified	19
			G35 – Multiple sclerosis	9
			G30.9 – Alzheimer disease, unspecified	8
			G80.8 – Other cerebral palsy	5
XIX.	**S00–T98:** Injuries, poisoning, and certain other consequences of external causes	101 (7.4%, 6.1–8.9%)	T42.6 – Poisoning: Other antiepileptic and sedative-hypnotic drugs	11
			T75.1 – Drowning and nonfatal submersion	11
			T17.9 – Foreign body in respiratory tract part unspecified	6
			T40.2 – Poisoning: Other opioids	6
			T42.4 – Poisoning: Benzodiazepines	6
			T51.9 – Toxic effect: Alcohol unspecified	5
			T90.9 – Sequelae of unspecified injury of head	5
V.	**F00–F99:** Mental and behavioural disorders	76 (5.6%, 4.4–6.9%)	F10.2 – Mental and behavioural disorders due to use of alcohol (dependence syndrome)	22
			F81.9 – Developmental disorder of scholastic skills unspecified	20
			F19.2 – Mental and behavioural disorders due to multiple drug use and use of other psychoactive substances	11
			F03 – Unspecified dementia	7
XX	V01–Y98: External causes of morbidity and mortality	74 (5.4%, 4.3–6.8%)	X42 – Accidental poisoning by and exposure to narcotics and psychodysleptics [hallucinogens], not elsewhere classified	9
			W74 – Unspecified drowning and submersion	6
			W78 – Inhalation of gastric contents	5
			Y12 – Poisoning by and exposure to narcotics and psychodysleptics [hallucinogens], not elsewhere classified, undetermined intent	5

			W65 – Drowning and submersion while in bath-tub	<5
XVII.	**Q00–Q99:** Congenital malformations, deformations and chromosomal abnormalities	72 (5.3%, 4.1–6.6%)	Q90.9 – Down syndrome, unspecified	27
			Q05.9 – Spina bifida, unspecified	5
XI.	**K00–K93:** Diseases of the digestive system	61 (4.5%, 3.4–5.7%)	K70.9 Alcoholic liver disease, unspecified	11
			K74.6 Other and unspecified cirrhosis of liver	5
II.	**C00–D48:** Neoplasms	47 (3.4%, 2.5–4.6%)	C71.9 Malignant neoplasm: Brain, unspecified	9
			C50.9 Malignant neoplasm: Breast, unspecified	5
IV.	**E00–E90:** Endocrine, nutritional, and metabolic diseases	41 (3.0%, 2.2–4.1%)	E14.9 Unspecified diabetes mellitus, without complications	8
			E11.9 Type 2 diabetes mellitus	5
XVIII.	**R00–R99 (excl. R56.8):** Other symptoms, signs and abnormal clinical/lab findings not elsewhere classified	38 (2.8%, 2.0–3.8%)	R68.8 Other specified general symptoms and signs	13
			R09.0 Asphyxia	6
I.	**A00–B99:** Certain infectious and parasitic diseases	28 (2.1%, 1.4–3.0%)	A41.9 Sepsis, unspecified	15
			B18.2 Chronic viral hepatitis C	5
XIV.	**N00–N99:** Diseases of the genitourinary system	21 (1.5%, 1.0–2.3%)	N19 Unspecified kidney failure	5
III.; XVI.; VII-VIII; XII.; XII.	**D50–D89:** Diseases of blood and blood-forming organs and certain disorders involving immune mechanisms; **P00–P96**: Perinatal diseases; **H00–H95:** Diseases of sense organs; **L00–L99:** Diseases of the skin and subcutaneous tissue	7 (0.5%, 0.2–1.1%)	Each code	<5
XIII.	**M00–M99:** Musculoskeletal and connective tissue diseases	6 (0.4%, 0.2–1.0%)	Each code	<5
XV.	**O00–O99:** Pregnancy, childbirth and puerperium	0 (0.0%)	N/A	N/A
Total		**1,365**		
			Note: All remaining ICD-10-coded causes of death as individual codes within chapters were used less than 5 times and are not shown	

Abbreviations: SUDEP – Sudden unexpected death in epilepsy; ICD-10 – International Classification of Diseases 10; CI – Confidence interval; N/A – Not applicable; Excl. - Excluding

5.3.5 Potentially avoidable epilepsy-related deaths in the young adults

421/553 young adult epilepsy-related deaths (i.e. 76%) had an underlying cause in the RDAM 2016 list, indicating that these were potentially avoidable.[43,45] These deaths are summarised in Table 5-5, demonstrating that the most common reason for them being classified as potentially avoidable was having epilepsy itself (G40–41) coded as the underlying cause of death (contributing 70% of the 421 potentially avoidable cases). Cardiovascular disease then alcohol and drug abuse were the next most common underlying mechanisms to contribute to the potentially avoidable cases.

Table 5-5: Potentially avoidable causes of death according to the ONS RDAM 2016 classification

Potentially avoidable underlying cause of death – ICD-10 codes	Number of young adult epilepsy-related deaths affected (as a percentage of the total, 95% CI)
Neurological disorders • Epilepsy and status epilepticus – G40-G41	294 (69.8%, 65.2–74.2%)
Cardiovascular diseases • Rheumatic and other valvular heart disease – I01-I09 • Hypertensive diseases – I10-I15 • Ischaemic heart disease – I20-I25 • DVT with pulmonary embolism – I26, I80.1-I80.3, I80.9, I82.9 • Cerebrovascular diseases – I60-I69 • Aortic aneurysm and dissection – I71	34 (8.1%, 5.7–11.1%)
Drug use disorders • Alcohol related, excl. external causes – F10, G31.2, G62.1, I42.6, K29.2, K70, K73, K74 (excl. K74.3-K74.5), K86.0 • Illicit drug use disorders – F11-F16, F18-F19	28 (6.7%, 4.5–9.5%)
Neoplasms • Malignant neoplasm of lip, oral cavity and pharynx – C00-C14 • Malignant neoplasm of oesophagus – C15 • Malignant neoplasm of stomach – C16 • Malignant neoplasm of colon and rectum – C18-C21 • Malignant neoplasm of liver – C22 • Malignant neoplasm of trachea, bronchus and lung – C33-C34 • Malignant melanoma of skin – C43 • Mesothelioma – C45 • Malignant neoplasm of breast – C50 • Malignant neoplasm of cervix uteri – C53 • Malignant neoplasm of bladder – C67 • Malignant neoplasm of thyroid gland – C73 • Hodgkin's disease – C81 • Leukaemia – C91, C92.0 • Malignant neoplasm of testis – C62 • Malignant neoplasm of unspecified parts of uterus and body of uterus – C54-C55 • Benign neoplasms – D10-D36	13 (3.1%, 1.7–5.2%)
Respiratory diseases • Influenza (including swine flu) – J09-J11 • Pneumonia – J12-J18 • Chronic obstructive pulmonary disorder – J40-J44 • Asthma – J45-J46 • Selected respiratory diseases – J00-J06, J20-J22, J30-J39	11 (2.6%, 1.3–4.6%)
Maternal and infant • Complications of perinatal period – P00-P96, A33 • Congenital malformations of the circulatory system – Q20-Q28 • Spina Bifida – Q05	11 (2.6%, 1.3–4.6%)
Unintentional injuries • Transport Accidents – V01-V99 • Accidental Injury – W00-X59	9 (2.1%, 1.0–4.0%)
Nutritional, endocrine and metabolic • Diabetes mellitus – E10-E14 • Diseases of the Thyroid – E00-E07 • Addison's disease – E27.1 **Digestive disorders** • Gastric and duodenal ulcer – K25-K28 • Acute abdomen, appendicitis, intestinal obstruction, cholecystitis/lithiasis, pancreatitis, hernia – K35-K38, K40-K46, K80-K83, K85, K86.1-K86.9, K91.5 **Genitourinary disorders** • Nephritis and nephrosis – N00-N07, N17-N19, N25-N27 • Obstructive uropathy and prostatic hyperplasia – N13, N20-N21, N35, N40, N99.1	8 (1.9%, 0.8–3.7%)
Intentional injuries • Suicide and self-inflicted injuries – X60-X84, Y10-Y34 • Homicide/Assault – X85-Y09, U50.9 • Misadventures to patients during surgical and medical care – Y60-Y69, Y83-Y84	8 (1.9%, 0.8–3.7%)
Infections • Tuberculosis – A15-A19, B90 • Selected invasive bacterial and protozoal infections – A38-A41, A46, A48.1, B50-B54, G00, G03, J02, L03 • Hepatitis C – B17.1, B18.2 • Pertussis (whooping cough) – A37 • Measles – B05 • Rubella – B06 • Other infections (Diphtheria, Tetanus, Poliomyelitis and Varicella) – A35, A36, A80, B01 • Intestinal infections – A00-A09 • HIV/AIDS – B20-B24	5 (1.2%, 0.4–2.7%)
Total	**421**

5.4 Discussion

5.4.1 Summary of findings

This study shows that epilepsy-related deaths remain common in Scotland. Despite a high number of epilepsy-related hospital admissions in the period leading up to death, there is relatively little outpatient neurology clinic input. Many of those experiencing epilepsy-related deaths live in Scotland's most deprived areas. The rate of epilepsy-related deaths has failed to reduce over time, and young adults appear to be at the highest risk when compared to age-mated mortality figures in the general population. Although SUDEP is a common mechanism of death in this group, it is only one of many other common mechanisms contributing to the burden of epilepsy-related death. There may be a large burden of potentially avoidable epilepsy-related deaths in young adults, and as many as half of young adults experiencing epilepsy-related death may have had comorbid mental health problems, whilst alcohol or drug abuse may be present in 40%. Although the study findings are limited to 2009–16, they remain relevant as few new treatments or guidelines on epilepsy management have been implemented since this period.[1] This study helps provide a methodological framework for future national study of epilepsy-related mortality in Scotland and other countries with well-linked administrative healthcare datasets. This includes illustrating the utility of pairing ICD-10 cause of death codes with their original free-texts to help identify SUDEP cases, as the lack of a unique ICD-10 code for SUDEP would otherwise generate ambiguity if population screening for SUDEP cases relied on ICD-10 codes alone. The recently introduced ICD-11 system now includes a unique SUDEP code,[222] although many administrative healthcare datasets still rely on ICD-10.[62]

5.4.2 Research in context

The study results complement the only other population-based study investigating mortality in PWE in Scotland.[95] This was limited to using only a regional cohort of PWE attending a specialist epilepsy service in one hospital in Glasgow between 1981–2001.[95] The study estimated standardised mortality ratios (SMRs) of 1.42–2.05, identifying the

greatest risk in those aged <30 years (SMR up to 7.03). These elevated risk figures are similar to those in our study, over a decade after this study and now using data from the entire Scottish population. This suggests little has improved in Scottish epilepsy care despite those early regional warnings of increased premature mortality. More widely, the current study results also complement those of our recent systematic review of mortality in PWE, where a median SMR range of 2.2–3.4 was identified globally across 103 observational studies, with mortality in PWE failing to fall between 1950–present.[12]

The risk of epilepsy-related death in the current study peaked at ages 16–24 years. This is a vulnerable period of transition from paediatric to adult care, and it may be that more needs to be done to support epilepsy transition services in Scotland. Whilst the literature often suggests mortality in PWE is increased largely for children and young adults aged <25 years due to congenital epilepsies,[15] we demonstrate that increased risk can persist well beyond this age-group into the fifth decade, suggesting there may also be important acquired factors. Our primary care dataset indicates that some of these may be comorbid mental health problems including depression, suicidality, and psychiatric problems, and it may be that more needs to be done to manage mental health problems in PWE in Scotland. Healthcare services may also need to be re-evaluated to facilitate acute seizure- and epilepsy-related hospital admissions translating into outpatient neurology care, as we demonstrate a disparity between these areas. The reasons for this remain unclear and require further investigation, particularly as outpatient neurology care is related to improved clinical outcomes for PWE.[223] Furthermore, excess mortality in PWE has been linked to increased social deprivation in England,[224] and we demonstrate similar trends in Scotland, where the majority of epilepsy-related deaths occurred in the two most deprived SIMD quintiles. This finding is also seen in many other conditions in Scotland,[225] and illustrates how widespread the problem of social deprivation is, suggesting that generic strategies for improving social mobility and support could be as effective for epilepsy as other for other conditions.

As most young adults with epilepsy want to know about SUDEP,[226] it is likely they would also want to know about other epilepsy-related causes of death. We show that SUDEP constituted only 30% of young adult epilepsy-related deaths, and mechanisms involved in the remaining epilepsy-related deaths included aspiration pneumonia, cardiac arrest, cerebral palsy, AED poisoning, drowning, alcohol dependence, narcotic poisoning, Down syndrome, and brain neoplasms. Some of these causes were not included in a recently proposed classification system for epilepsy-related deaths,[23] and it would seem plausible to consider their addition in future iterations. It will also be important to further corroborate the estimated burden of potentially avoidable epilepsy-related deaths identified using RDAM 2016 against other sources including PM reports, medical records, GP records, and family/physician interview.

Our finding fewer epilepsy-related deaths in those aged ≥55 years than in the general population should be interpreted with caution as there is known to be an underestimation of epilepsy-related mortality amongst older adults.[23] The alleged incidence of SUDEP, for example, declines markedly after age 50: 8% of all cases in 51–60-year-olds, 3% among 61–70-year-olds, and 1% in individuals >70 years.[23] However, older people are not immune to SUDEP.[23] The lack of reported cases reflects that when elderly PWE die suddenly, autopsies are rarely performed, and, when they are, priority tends to be placed on general medical pathologies such as coexistent cardiovascular disease.[23] In the current study, an autopsy was performed in only 6% of those who died aged ≥55 years. Furthermore, in younger people, there is a preponderance of temporal lobe seizures,[227] the semiology of which is well recognised. In older people, most seizures are of extratemporal onset, diverse in semiology, and convulsions are relatively rare.[227] Lack of awareness that an unusual episodic event might be ictal in origin can prevent diagnosis and therefore cause-of-death attribution.[227] Additionally, particularly in high-income countries where many older people live alone, eyewitness accounts are often unavailable, further compounding the potential for ictal underestimation.[227]

5.4.3 Strengths and limitations

This is the first national study of epilepsy-related deaths in Scotland. We include an MCOD data analysis novel to epilepsy mortality literature in order to identify the various mechanisms of epilepsy-related death in high risk groups. We also use the ONS's RDAM 2016 on epilepsy mortality data for the first time, signalling what proportion were potentially avoidable. The study is strengthened by case ascertainment from two mandatory national datasets, meaning that virtually the entire Scottish population of interest is studied. These are enriched with information from three additional datasets that add to the depth of information studied. The primary care dataset contained more Scottish GP practices than others including The Health Improvement Network and the Clinical Practice Research Datalink.[212] The study also benefitted from using the optimal case-ascertainment codes previously validated, with PPVs of 91%.[212] However, this means that up to 9% of the alleged epilepsy-related deaths we identified may not have had epilepsy. This is a small number that is within the expected false positive margin of any study using administrative data to capture epilepsy,[62] and it is unlikely to change our overall conclusions as these were population-based and frequently underpinned by autopsy confirmation. The study is also likely to have benefitted from the cause-of-death quality assurance processes set up to improve Scottish death certification through NRS,[212] the robustness of which was highlighted recently.[61]

Although 90% of our SUDEP cases were confirmed at autopsy, it remains possible that the study underestimated SUDEP as Procurator Fiscal investigation did not proceed to autopsy in 88/194 (i.e. 45%) of the non-SUDEP epilepsy-related deaths deemed suspicious enough to require Procurator Fiscal involvement. In a "worst-case" scenario in which all 88 of these deaths were missed SUDEP cases, this would increase our SUDEP count to 254 from 166, constituting 45% of the 553 young adult epilepsy-related deaths instead of the 30% reported. Therefore, this would not change our overall conclusion that, taken together, the non-SUDEP epilepsy-related deaths predominated causes in young

adults. This view is also emerging in recent literature,[23] and is supported by our systematic review evidence in Chapter Two.[12] This is an important area to research further as it potentially opens up additional targets for epilepsy mortality advice and prevention beyond the SUDEP-specific targets, which tend to predominate research.[23]

The study limitations included inability to assess a live population at risk dataset, meaning we were unable to calculate MR using person-years at risk as a denominator. We were also limited by lack of access to a national control group dataset without epilepsy to compare our estimates against, including the MCOD data. Finally, we were unable to control for other mortality risk factors such as smoking and obesity because of lack of an appropriate national dataset containing these.[43] Our findings should therefore help stimulate the development of appropriate national datasets to assess these areas. In particular, future studies could use case-control data to assess combinations of causes of death in cases with epilepsy compared with those combinations of death in controls without epilepsy to assess for the relationship. We know that in decreasing order, the 10 most frequently occurring principle causes of premature death in the general Scottish population <75 years old are lung cancer, myocardial infarction, drug poisoning, ischemic heart disease, breast cancer, chronic obstructive pulmonary disease (COPD), atherosclerotic heard disease, infected COPD, oesophageal cancer, and hanging.[228] Our MCOD data signpost a different set of top ten causes to this when epilepsy is involved (Table 5-4), highlighting rationale for a future case-control MCOD study. Our study provides a robust methodological framework for the data linkage and MCOD case-ascertainment methods future researchers could use to undertake such a study in Scotland.

5.5 Conclusion

This study gives an indication that epilepsy-related deaths are potentially a major public health problem in Scotland and highlights several areas in which further work is likely to

help further refine the exact scale of the problem. The study illustrates the validity of using linked administrative data to identify potential trends in disease.

Chapter 6. A case-control study developing the Scottish Epilepsy Deaths Study Score (SEDS Score) as a risk prediction model for epilepsy-related deaths

Student contributions: The bulk of my work in this chapter involved undertaking and interpreting multivariable regression modelling using data I had abstracted from the case and control data collection *pro formas*. This analysis required statistical support. The draft manuscript I have prepared from this work is ready for journal submission.

6.1 Introduction

Epilepsy is common, affecting 70 million people globally.[86,94] Of the 600,000 PWE in the UK, 60,000 live in Scotland.[52] PWE are at significantly increased risk of premature death.[12] Some of those deaths may be entirely unrelated to their epilepsy. However, a substantial proportion relate to the epilepsy itself, its treatment, or comorbidities associated with epilepsy.[12,23] Together, these are termed 'epilepsy-related deaths', and they are likely to contribute a significant burden of total years of potential life lost.[12,23] Examples include deaths occurring directly due to epilepsy (including SUDEP and seizure-related accidents), deaths occurring due to acute symptomatic seizures that would likely have gone on to become epilepsy (including after stroke or traumatic brain injury), deaths indirectly occurring due to epilepsy (including aspiration pneumonia or suicide), and deaths occurring due to the underlying aetiology of epilepsy (including brain tumours and genetic conditions).[23]

In Chapters Two and Five, we showed that it is the epilepsy-related deaths that predominate causes of death in PWE.[12] Furthermore, the avoidable deaths in PWE tend to be those that are epilepsy-related because the greater risk of death over the general population is abolished by, for example, achieving seizure freedom through establishing effective treatment strategies.[14,39] Our data in Chapter Five suggest nearly 80% of epilepsy-related deaths in young adults aged 16–55 years in Scotland between 2009–2016 may have been avoidable. This highlights the importance of trying to identify what

the modifiable risk factors for epilepsy-related death are, and to risk-stratify which individuals are at the highest risk. The risk factors associated with PWE suffering death of any cause (i.e. all-cause mortality) are well-established.[12] As we reported in Chapter Two, they include comorbidities (such as brain tumours, intellectual disabilities/cerebral palsy, and encephalopathies), generalised epilepsies, increased seizure frequency, age (SMRs highest in children and young adults), and being male.[12] The risk factors associated with PWE suffering epilepsy-related death *specifically*, however, remain largely unclear.[12] Whilst risk factors associated with a few individual epilepsy-related causes of death such as SUDEP, drowning, and suicide have been described, few studies have attempted to describe a broad risk factor profile or generate a risk prediction model for epilepsy-related deaths as a group.[12] To help address this, we have now undertaken a Scottish multi-centre case-control study. The study aims to analyse routine clinical information gathered from medical records 1) to describe multiple potential risk factors for suffering epilepsy related death in adults, and 2) to develop a pragmatic risk prediction model for identifying those adults at increased risk of epilepsy-related death.

6.2 Methods

6.2.1 Study design summary

This was an age- and sex-matched case-control study comparing adults who suffered epilepsy-related death to living adults with epilepsy. Data pertaining to potential risk factors for epilepsy-related death were collected from the medical records of cases and controls to allow multivariable analysis and development of a risk prediction model. The study was designed for model development and internal validation.[229] The study is reported according to the transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD) guidelines for model development.[230]

6.2.2 Data sources, and participants

6.2.2.1 Case ascertainment

Cases were screened for inclusion from the dataset described in Chapter Four investigating the diagnostic accuracy of four administrative healthcare datasets used to identify deceased adults with epilepsy in Scotland.[212] In that study, we linked national mortality records, hospital admissions, AED prescriptions, and regional primary care attendances in order to identify adults (aged ≥16 years) who died between 01/01/09–01/01/16. Coded indicators of epilepsy were compared to confirmed epilepsy diagnoses made by a senior epileptologist upon reviewing medical records (*see* Confirming epilepsy diagnosis section). The study identified 614 confirmed epilepsy cases, who had suffered epilepsy-related death during the 2009–16 study period. Epilepsy-related death was defined as death occurring with at least one ICD-10-coded cause listed as G40–41 and/or R56.8 in any position within the sequence of events that lead to death in the mortality records, or with cause of death free texts pertaining to seizures (alongside at least one community-prescribed AED), epilepsy, or SUDEP. Such a definition is consistent with national guidance that the presence of an epilepsy code/indicator anywhere in the cause of death record indicates that epilepsy was thought by the certifying doctor to be either part of the sequence of events leading to death or else a factor contributing to the death.[31,213] Additional case-ascertainment benefit was drawn from the routine cause-of-death quality assurance processes in place to improve the accuracy of Scottish death certification and coding,[212] the robustness of which was highlighted recently.[61]

6.2.2.2 Control ascertainment

Living control participants were screened for inclusion from a nationwide Scottish Health Research Register (SHARE), supplemented by controls recruited directly from NHS Lothian epilepsy clinics. Written consent was obtained from both groups.

Share controls: The SHARE register consists of nearly 300,000 volunteers in Scotland, aged ≥11 years, interested in participating in health research and who agree to allow

SHARE staff to use their health record ICD-10 coding to check whether they might be suitable for clinical research studies.[231,232] SHARE connected us with registrants aged ≥16 years with at least one G40–41-ICD-10-coded activity. We reviewed the medical records of those consenting to participate. Eligible controls required sufficiently complete medical records to allow confirmation of a diagnosis of epilepsy and abstraction of clinical data.

Clinic controls: Clinic controls were recruited in a consecutive manner directly from NHS Lothian epilepsy clinics. We reviewed the medical records of those consenting to participate. Eligible controls required sufficiently complete medical records to allow confirmation of a diagnosis of epilepsy and abstraction of clinical data.

Controls were matched 1:1 by age (± 10 years) and sex to cases of epilepsy-related death after recruitment.

6.2.2.3 Confirming epilepsy diagnosis

For each candidate case and control, the medical records were reviewed by an experienced consultant epileptologist (S.E.D.), who used the medical records to confirm the presence or absence of epilepsy.[62,130] This was based on corroborative evidence such as the presence of two or more unprovoked seizures or clear documentation of a diagnosis of epilepsy from a neurologist.[5,186] The medical records were extensive, and included electronic and paper records from general and/or specialist inpatients, emergency care, outpatients (including neurology clinics), hospital discharge summaries, referral letters from GPs, radiology scans, neurophysiology reports, and medication lists.

6.2.2.4 Data abstraction, outcome and predictors

The outcome for model prediction fell into binary categories: 1 = deceased (i.e. epilepsy-related death) and 0 = alive (living control).

Medical record data abstraction was undertaken by S.E.D during multiple hospital visits made between 11/10/2017–24/02/2019 for cases, and 25/02/2019–15/01/2020 for controls. The results of our systematic review of epilepsy mortality literature in Chapter

Two were used to help guide the development of an *a priori* clinical data abstraction tool that was used both for cases and controls (Appendix S1).[12] This was populated with potential risk factors for epilepsy-related death identified in the systematic review literature. From these, candidate predictor variable selection was further refined through consensus discussion between the epilepsy specialist research team, and also supported by further consultation at research meeting presentations to external experts, patient and charity groups, and parliamentary policymakers.[233-235] Table 6-1 lists the predictor variables that were ultimately selected as candidates for model development. Data for all of the variables were sourced from medical records aside from SIMD quintile, which was already automatically provided within the mortality record for each case. SIMD quintiles split the Scottish population by socioeconomic deprivation, ranked from 1 = most deprived to 5 = least deprived areas of residence.[214] For each control participant, we inserted their home postcode into the Scottish Government's publicly-available SIMD postcode lookup file to reveal their SIMD quintile.[214]

6.2.3 Statistical analysis

All analyses were performed using IBM SPSS Statistics for Macintosh, Version 25.0. Armonk, NY: IBM Corp. Where relevant, categories with less than five events/participants were combined or reported as '<5' to avoid patient identification.[184]

6.2.3.1 Sample size

We estimated that we could capture a convenience sample size of 200–250 controls, matched to a mirrored number of age- and sex-matched cases taken from the 614 deceased adults suffering epilepsy-related death, for whom medical records had already been reviewed during the diagnostic accuracy study.[212] This number of controls was based on our estimated capacity and the time taken to fully recruit living controls and abstract their clinical data from medical records. Using a recommended rule of 50 events per variable for logistic regression to estimate coefficients with adequate precision,[236,237] such a sample size would allow for a multivariable risk prediction model size of four

predictor variables. Using more variables than the sample data could support would increase the risk of overfitting (achieving overly optimistic predictive accuracy and hence resulting in failure to replicate the results in another sample).[238]

6.2.3.2 Missing data

An *a priori* threshold was set to exclude from further analysis any variables with ≥30% of their values missing either within cases or within controls, based on prior literature.[239] We also excluded any variables with >10% difference in the missing data burden between cases compared to controls.[240] These exclusions allowed us to prioritise creating a pragmatic risk prediction model consisting of variables likely to be found complete in medical records, and that were uniformly represented between cases and controls. For variables with <30% of their values missing, an overall summary of the missing data patterns was given and a logistic regression model for multiple imputation was used to replace the missing values. The missing values were imputed using fully conditional specification techniques based on the iterative Markov chain Monte Carlo (MCMC) method.[241] This fits a univariate model using all other available variables in the model as predictors, then imputes missing values for the variable being fit. A total of 20 imputed datasets were generated (which is higher than the standard requirement for 2–10 imputations),[242] and estimates obtained from the imputed datasets were pooled. The missing data handling methods used were based on an assumption data were missing at random (MAR). This was a plausible assumption given that a large number of clinically relevant variables were tested, meaning that the probability of missingness would likely depend on only the observed data already captured within the dataset.[243]

6.2.3.3 Model development

Where possible, we reported and analysed both the original data (unimputed) and the imputed datasets for the models developed. Primary analysis was of the original data. Imputed data were used for sensitivity analysis.

Univariable analysis: Univariable analysis was performed using conditional logistic regression for included variables to provide an odds ratio (OR) and 95% CI.

Correlation: Correlation was examined in a pairwise fashion between the variables, using contingency tables and Chi-square tests with a 2-sided 5% significance. Where correlation was significant ($p < 0.05$ for both unimputed and imputed data), the Lambda statistic was used to assess the strength of association, interpreted as 0.01–0.09 = weak association, 0.10–0.29 = moderate association, 0.30–0.99 = strong association, and 1.00 = perfect association.[244] Correlation was assessed for both cases and controls separately, with the highest Lambda statistic between both groups reported (Table 6-2).

Model selection: Two senior authors (S.E.D. and R.F.M.C.) independently made a clinically-driven decision to select four variables for inclusion in the final risk prediction model,[238] whilst blinded to the univariable results and taking into account variable correlations and missing data burdens (see Table 6-1 for justification of each choice made). Any disagreements in variable selection were resolved by consensus.

Multivariable analysis: Multivariable conditional logistic regression was performed to provide ORs for the four variables selected for risk prediction modelling, adjusted for one another, alongside the 95% CI. Any variables retaining significance in this first model were then combined into a second multivariable conditional logistic regression model to create a risk score. This was calculated, as has been done in other modelling studies,[245] by taking a sum of the factors present, with a point given for each positive factor, to give a total score – which we termed a Scottish Epilepsy-related Deaths Study score (SEDS Score). An internal validation process was performed for each model (i.e. quantifying statistical performance of the model using the data on which the model was developed).[229] This was done by performing 1,000 bootstrap analysis samples to provide more robust p-values, which were then checked for agreement against the original model's p-values for each variable. Where there was poor agreement, the bootstrap result was prioritised.

We used Nagelkerke's R^2 to estimate what proportion of variance in the outcome was explained by the predictor variables in the model, and the Hosmer-Lemeshow goodness-of-fit test to determine whether the model adequately described these outcome data (where $p \geq 0.05$ indicates a good fit). The Hosmer-Lemeshow contingency table was used to assess model calibration by illustrating agreement between subgroups of predicted probabilities and their observed frequencies.[229,246] Discrimination describes how well the model differentiates between patients who experienced the outcome (deceased cases) and those who did not (living controls).[229] This would normally be measured using an area under the receiver operating characteristic (ROC) curve.[246] However, ROC statistics can be misleading for binary or categorical predictors and, as such, they are cautioned against in this scenario.[247] Therefore, we approximated discrimination primarily through indicating the model's ability to capture all of the deceased cases by estimating its sensitivity, and all of the living outcomes by estimating its specificity.[62] PPV, NPV, and model accuracy (sensitivity × prevalence + specificity × (1 − prevalence)) were reported as supplementary material for reference only given that these estimates depend on prevalence of the outcome, which was fixed at 50% in this study because of the 1:1 matched case-control design. 95% CIs were included.

6.2.4 Approvals

This study was approved by South East Scotland Research Ethics Committee 2 (15/SS/0165, IRAS 181131), and the Scottish PBPP.

6.3 Results

6.3.1 Baseline characteristics of cases and controls

The study cohort comprised of 224 deceased cases (114 male, 110 female) and 224 living controls (114 males, 110 female), spread across multiple NHS Health Board regions in Scotland (see Figure 6-1). Mean age at death for cases and controls was 48 years (range 16–90 years for cases, and 19–87 years for controls). 124 controls were recruited from SHARE Scotland,[231] with the remaining 100 recruited from NHS Lothian epilepsy clinics.

6.3.2 Missing data

Several predictor variables were excluded from further analysis owing to their missing data burden (see Table 6-1). The remaining variables were taken forward into multiple imputation and univariable analysis. Figure 6-2 provides an overall summary of the missing data patterns for these remaining variables.

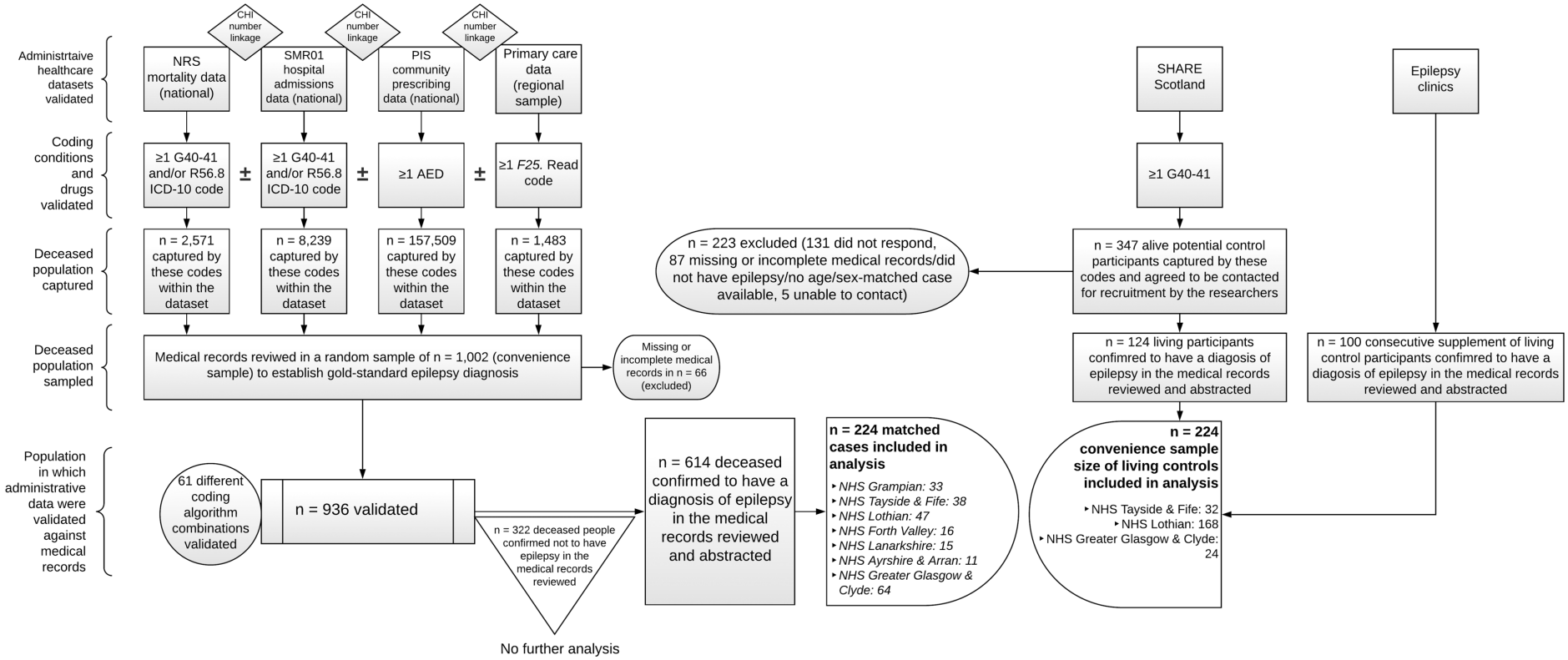


Figure 6-1: Study flow diagram

Abbreviations: AED – antiepileptic drug; *F25* – primary care diagnostic Read codes for epilepsy; G40–41 – International Classification of Disease 10 (ICD-10) codes for epilepsy and status epilepticus; R56.8 – ICD-10 code for seizures; CHI – Scottish Community Health Index patient identification number; GP – General practitioner; NHS – National Health Service

Table 6-1 - Variables considered for analysis (N = 224 deceased cases and 224 alive controls)

Candidate predictor variable	Definition	Functional coding for statistical purposes	Missing data burden		Chosen for multivariable model	Justification for choice
			Cases	Controls		
SIMD Quintile	SIMD quintiles (socioeconomic deprivation markers splitting the population into quintiles ordered as 1 (most deprived) to 5 (least deprived) areas of residence).	⇒ Group 1 = two most deprived quintiles combined ⇒ Group 2 = middle deprived quintile ⇒ Group 3 = two least deprived quintiles combined	0%	3%	Yes	Little missing data, easy to obtain from home postcode, studied little in epilepsy mortality literature, potentially modifiable from improved socioeconomic support
A&E attendance or hospital admission for epilepsy/seizures	Did the patient have a seizure- or epilepsy-related A&E attendance and/or seizure- or epilepsy-related hospital admission in the 12 months prior to death (cases) or the 12 months prior to medical record abstraction (controls)?	⇒ Yes ⇒ No ⇒ Unable to tell (missing data)	4%	1%	Yes	Little missing data, studied little in epilepsy mortality literature, and may help alter the management of PWE attending A&E or hospital for seizures
Aetiology and risk factors for developing epilepsy	Does the patient have an identifiable aetiology or risk factor for developing epilepsy? ***Inherited/Congenital:*** *Febrile convulsions, first degree relative with epilepsy, congenital abnormality/malformation (e.g. cerebral palsy, metabolic infancy syndrome, birth hypoxia), genetic syndrome, ADHD, ASD, developmental/intellectual delay, premature birth, birth/perinatal difficulties, hydrocephalus, neonatal seizures* ***Acquired:*** *History of meningitis/encephalitis, severe head injury, brain tumour, cerebrovascular disease, limbic encephalitis, brain surgery, CNS demyelinating disease, abnormal brain imaging, other hypoxic brain injury, neurodegenerative disease*	⇒ No risk factors for epilepsy ⇒ Inherited/Congenital ⇒ Acquired ⇒ Unable to tell (missing data)	13%	12%	Yes	This would have likely been a major confounding risk factor not to include in a model about epilepsy-related deaths. It correlated strongly with epilepsy type but had fewer missing data.
Number of comorbidities	How many comorbidities did the patient have recorded alongside epilepsy in their medical records when last seen? *This includes any comorbidities listed in the medical record regardless of organ system affected or whether they are physical or mental*	⇒ None ⇒ 1 condition ⇒ ≥2 conditions ⇒ Unable to tell (missing data)	4%	2%	Yes	Little missing data and it would be usual to adjust for comorbidities in a risk prediction model.
Alcohol abuse	Any evidence of alcohol abuse in the 24 months prior to death (cases) or the 24 months prior to medical record abstraction (controls)? *This includes documented alcohol excess history plus/minus deranged liver function tests, liver cirrhosis, or use of alcohol dependency services*	⇒ Yes ⇒ No ⇒ Unable to tell (missing data)	11%	4%	No	Has a moderate correlation with epilepsy risk factors, which were therefore chosen over alcohol abuse to increase clinical generalisability of the model beyond substance misuse.
Epilepsy type	According to the ILAE 2017 Classification, what type of epilepsy did the patient have?[248]	⇒ Focal ⇒ Generalised	20%	9%	No	Has a strong correlation with epilepsy aetiology/risk factors, which were therefore chosen as they had fewer missing data

Mental health problems	Did the patient have evidence of mental health problems in the 24 months prior to death (cases) or the 24 months prior to medical record abstraction (controls)? *This includes contact with psychiatric, psychological, or learning disability services, admission under psychiatry, self-harm or suicide attempts, depression, treated depression, anxiety*	⇒ Yes ⇒ No ⇒ Unable to tell (missing data)	16%	14%		No	Although of importance clinically, this variable had a relatively large amount of missing data compared to other variables included, e.g. A&E attendance or hospital admission, against which this variable was also moderately correlated.
Epilepsy surgery	Has the patient ever had brain surgery to treat or control seizures?	⇒ Yes ⇒ No ⇒ Unable to tell (missing data)	7%	1%	No		It is probable that too few people would have had this highly specialist intervention to make it meaningful as a variable to include in an everyday risk score.
Seen by neurologist	Has the patient been seen by a neurologist in the 12 months prior to death (cases) or the 12 months prior to medical record abstraction (controls)?	⇒ Yes ⇒ No ⇒ Unable to tell (missing data)	1%	1%		No	Although of importance clinically, 45% of the control participants were recruited directly from epilepsy clinics. The potential selection bias introduced by this would make it a less appropriate variable for the final model. It also had a moderate correlation with comorbidities, which were included in the model.
Number of AEDs	Number of AEDs prescribed when last seen	⇒ Not on an AED ⇒ 1 AED ⇒ 2 AEDs ⇒ ≥3 AEDs ⇒ Unable to tell (missing data)	0%	0%		No	Anticipated that potentially very few people were likely to not be on an AED, which would reduce the discriminatory value of this variable in a model.
Seizure frequency	What was their seizure frequency when last seen?	⇒ <1 seizure per year ⇒ ≥1 seizure per year ⇒ Unable to tell (missing data)	40%	11%		No	≥30% of the data were missing, therefore variable excluded from further analysis
Seizure timing	When did they mainly have seizures?	⇒ Daytime ⇒ Night time ⇒ Mixture of both daytime and night time ⇒ Unable to tell (missing data)	70%	56%		No	≥30% of the data were missing, therefore variable excluded from further analysis
Seizure alarm	Did they have a seizure alarm/alert device?	⇒ Yes ⇒ No ⇒ Unable to tell (missing data)	60%	36%		No	≥30% of the data were missing, therefore variable excluded from further analysis
History of status epilepticus	Any history of status epilepticus in the previous 24 months?	⇒ Yes ⇒ No ⇒ Unable to tell (missing data)	13%	1%		No	>10% difference in the missing data burden between cases compared to controls, therefore variable excluded from further analysis
Abbreviations:	SIMD - Scottish Index of Multiple Deprivation; A&E – Accident and Emergency; AEDs – Antiepileptic drugs; ADHD - Attention deficit hyperactivity disorder; ASD - Autism spectrum disorder; ILAE - International League Against Epilepsy; CNS – Central nervous system; PWE – People with epilepsy						

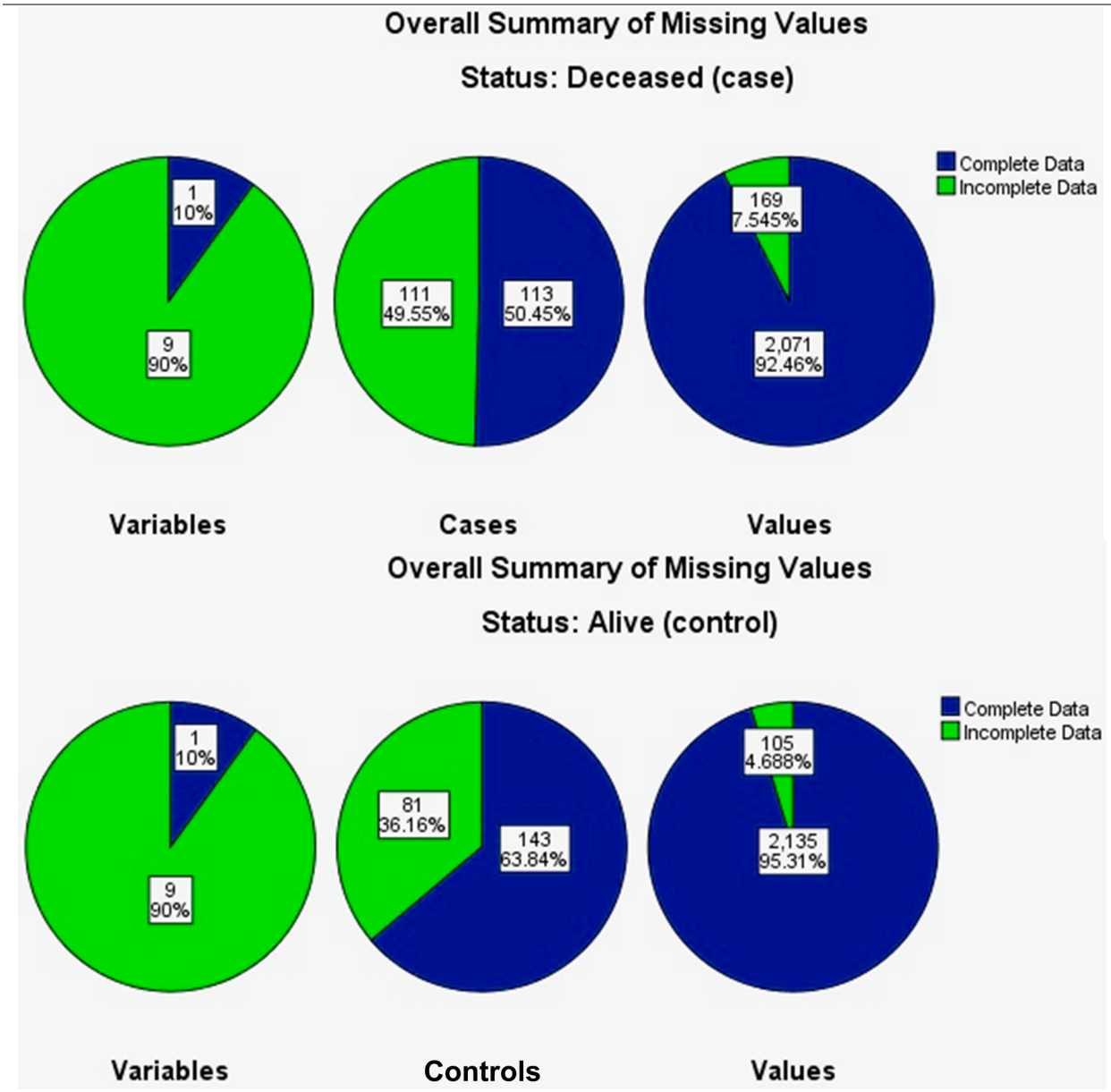


Figure 6-2: Missing data patterns for the 10 variables included in univariable analysis

6.3.3 Correlation

A strong association was seen only between the type of epilepsy (focal or generalised) and the risk factors for developing epilepsy (inherited/congenital or acquired) – see Table 6-2. The remaining pairs demonstrated moderate or weak association, with the greatest association amongst these being between having been seen by a neurologist in the preceding year and the number of AEDs prescribed when last seen.

Table 6-2 - Correlation results for variable pairs

Variable	Variable	Lambda Correlation	
Epilepsy type	Aetiology and risk factors for developing epilepsy	0.31	Strong association
Seen by neurologist	Number of AEDs	0.28	Moderate association
Seen by neurologist	Comorbidities	0.18	Moderate association
Alcohol abuse	Aetiology and risk factors for developing epilepsy	0.15	Moderate association
Aetiology and risk factors for developing epilepsy	Comorbidities	0.13	Moderate association
A&E attendance or hospital admission	Mental health problems	0.12	Moderate association
SIMD Quintile	Aetiology and risk factors for developing epilepsy	0.07	Weak association
SIMD Quintile	Seen by neurologist	0.06	Weak association
Alcohol abuse	Mental health problems	0.05	Weak association
Epilepsy surgery	Comorbidities	0.05	Weak association

Abbreviations: A&E – Accident and Emergency; AEDs – Antiepileptic drugs;
Colour palette key: To help aid interpretation, each colour used represents a single variable and cannot be used for another variable

6.3.4 Univariable analysis

The odds of suffering an epilepsy-related death were significantly in favour of those who lived in the two most deprived SIMD quintiles, those with a history of alcohol abuse in the preceding two years, those with generalised epilepsy, those who had experienced at least one accident and emergency (A&E) attendance or hospital admission in the preceding year related to seizures or epilepsy, those with evidence of mental health problems in the preceding two years, those who had not been seen by a neurologist in the preceding year, and those with an inherited or congenital aetiology for developing epilepsy (see Table 6-3). These variables were significant both using the original data (missing values excluded from analysis), and using imputed data (missing values replaced with imputed values). The odds of suffering an epilepsy-related death were significantly reduced in those who

had undergone previous epilepsy surgery when this was analysed using the original data. However, only a small number of cases and controls had had epilepsy surgery (<10% of the sample, overall), and the difference associated with this variable was lost when imputed values were included. The number of AEDs taken when last seen and the number of comorbidities were not significant individual predictors. Very few cases and controls (<3% of the sample) were not on an AED.

In an exploratory post-hoc analysis, the odds of suffering an epilepsy-related death remained significantly in favour of those who had not been seen by a neurologist in the preceding year even with analysis restricted to using only the SHARE control participants. This indicates that this variable was significant even when excluding the sample of control participants recruited directly from epilepsy clinics.

Table 6-3 – Univariable results

Predictor variable	Grouping	Study status		Total	Original data OR (95% CI)	P Value	Multiple imputation data OR (95% CI)	P Value
		Alive (control)	Deceased (case)					
SIMD Quintile	Two least deprived	85	50	135	Reference		Reference	
	Middle deprivation	42	43	85	1.8 (1.0–3.2)	0.06	1.8 (1.0–3.2)	0.05
	Two most deprived	91	131	222	**2.5 (1.6–4.0)**	**0.00**	**2.5 (1.6–4.0)**	**0.00**
Total		218	224	442				
Alcohol abuse	No	202	157	359	Reference		Reference	
	Yes	13	42	55	**4.4 (2.2–9.2)**	**0.000**	**3.7 (1.8–7.9)**	**0.001**
Total		215	199	414				
Epilepsy type	Focal	153	119	272	Reference		Reference	
	Generalised	71	105	176	**1.9 (1.2–3.0)**	**0.006**	**1.8 (1.1–2.8)**	**0.012**
Total		224	224	448				
A&E attendance or hospital admission	No	181	92	273	Reference		Reference	
	Yes	41	123	164	**5.1 (3.2–8.3)**	**0.000**	**5.1 (3.2–8.2)**	**0.000**
Total		222	215	437				
Mental health problems	No	101	80	181	Reference		Reference	
	Yes	91	109	200	**1.6 (1.0–2.6)**	**0.034**	**1.6 (1.1–2.4)**	**0.027**
Total		192	189	381				
Epilepsy surgery	No	193	203	396	Reference		Reference	
	Yes	29	6	35	**0.2 (0.1–0.5)**	**0.001**	0.4 (0.1–1.4)	0.159
Total		222	209	431				
Seen by neurologist (entire dataset)	Yes	180	115	295	Reference		Reference	
	No	42	106	148	**3.8 (2.4–6.1)**	**0.000**	**3.9 (2.5–6.3)**	**0.000**
Total		222	221	443				
Seen by neurologist (SHARE-only dataset)	**No**	26	61	87	**3.4 (1.8–6.1)**	**0.000**		
	Yes	96	61	157	Reference		Reference	
Total		122	122	244				
Aetiology and risk factors for developing epilepsy	None	92	55	147	Reference		Reference	
	Inherited/congenital	41	73	114	**3.1 (1.7–5.7)**	**0.000**	**3.3 (1.9–5.8)**	**0.000**
	Acquired	64	67	131	1.7 (0.9–3.2)	0.088	**1.9 (1.1–3.4)**	**0.030**
Total		197	195	392				
Comorbidities	No	55	42	97	Reference		Reference	
	1 condition	70	64	134	1.2 (0.7– 2.1)	0.511	1.2 (0.7–2.1)	0.496
	≥2 conditions	94	110	204	1.7 (1.0–3.0)	0.061	1.7 (1.0–2.9)	0.056
Total		219	216	435				
Number of AEDs	Not on an AED	3	8	11	Reference		Reference	
	1 AED	101	96	197	0.4 (0.1–1.3)	0.127	0.3 (0.1–1.3)	0.336
	2 AEDs	67	75	142	0.4 (0.1–1.6)	0.201	0.4 (0.1–1.5)	0.397
	≥3 AEDs	53	44	97	0.3 (0.1–1.2)	0.098	0.3 (0.1–1.2)	0.289
Total		224	223	447				

Abbreviations: SIMD - Scottish Index of Multiple Deprivation; A&E – Accident and Emergency; AEDs – Antiepileptic drugs; SHARE - Scottish Health Research Register; OR – Odds ratio; CI – Confidence interval

6.3.5 Multivariable analysis

The four variables ultimately selected for multivariable analysis are listed below (see Table 6-1 for expanded definitions of each variable):

- SIMD Quintile;
- A&E attendance or hospital admission for epilepsy/seizures
- Aetiology and risk factors for developing epilepsy
- Number of comorbidities.

Table 6-4 summarises results of the multivariable model containing these four variables. With all four variables adjusted for one another, the odds of suffering an epilepsy-related death remained significantly in favour of those who lived in the two most deprived SIMD quintiles (original data only), those who had experienced at least one seizure- or epilepsy-related A&E attendance and/or hospital admission in the preceding year, and those with inherited or congenital aetiology for developing epilepsy. These variables retained significance ($p < 0.05$) through 1,000 bootstrap samples, suggesting internal validity for the findings. The model was a good fit for the data ($p > 0.05$ for the Hosmer-Lemeshow goodness-of-fit test both using the original data and imputed data). 33% of the variance in the outcome was explained by the predictor variables included in this model ($R^2 = 0.332$). The model demonstrated good calibration, with the predicted estimates mirroring the actual estimates closely throughout the Hosmer-Lemeshow contingency table (Table 6-5). In terms of discrimination, the model demonstrated a sensitivity of 68% (95% CI 61–75%) and specificity of 80% (95% CI 74–86%) using the original data, indicating that the majority of cases and controls were identified. There was little difference in these discrimination figures when imputed values were included (Appendix S2).

Table 6-4 – Multivariable results for the four selected variables

Predictor variable	Grouping	Original data OR (95% CI)	P Value	Bootstrap P Value	Multiple imputation data OR (95% CI)	P Value
SIMD Quintile	Two least deprived	Reference			Reference	
	Middle deprivation	1.9 (0.8–4.6)	0.15	0.19	2.9 (1.6–5.4)	0.00
	Two most deprived	**2.7 (1.4–5.3)**	**0.005**	**0.008**	2.0 (0.9–4.3)	0.081
A&E attendance or hospital admission	No	Reference			Reference	
	Yes	**5.1 (2.8–9.2)**	**0.000**	**0.001**	**5.9 (3.5–10.0)**	**0.000**
Aetiology and risk factors for developing epilepsy	No risk factors	Reference			Reference	
	Inherited/congenital	**4.0 (1.9–8.8)**	**0.000**	**0.006**	**4.9 (2.4–10.1)**	**0.000**
	Acquired	1.5 (0.7–3.3)	0.297	0.339	1.7 (0.9–3.4)	0.122
Comorbidities	No	Reference			Reference	
	One condition	1.3 (0.6–2.7)	0.572	0.607	1.3 (0.6–2.5)	0.514
	Two or more conditions	1.4 (0.7–3.2)	0.362	0.475	1.7 (0.8–3.5)	0.153

Abbreviations: SIMD - Scottish Index of Multiple Deprivation; A&E – Accident and Emergency; OR – Odds ratio; CI – Confidence interval

Table 6-5: Contingency Table for Hosmer and Lemeshow Test

	First multivariable model					SEDS Score model				
Data subgroup	**Study status = Alive (control)**		**Study status = Deceased (case)**		**Total**	**Study status = Alive (control)**		**Study status = Deceased (case)**		**Total**
	Observed	Predicted	Observed	Predicted		Observed	Predicted	Observed	Predicted	
1	37	34	≤5	≤5	39	45	44	8	9	53
2	26	28	9	7	35	39	40	13	12	52
3	30	27	8	11	38	26	27	14	13	40
4	26	26	14	14	40	34	35	28	27	62
5	20	22	18	16	38	45	44	45	46	90
6	15	18	24	21	39	12	11	25	26	37
7	12	13	27	26	39	8	8	31	31	39
8	7	8	26	25	33	10	11	52	51	62
9	9	7	34	36	43	45	44	8	9	53
10	≤5	≤5	21	23	25	39	40	13	12	52

Each case and control participant was therefore ascribed a SEDS Score taking a sum of the positive scores of the variables that were significant in the above model as follows:

- 1 point for living in the two most deprived SIMD quintiles (0 points for living in any other SIMD quintile);
- 1 point for having ≥1 A&E attendance or hospital admission for epilepsy/seizures in the preceding year (0 points if not);

- 1 point for having at least one inherited or congenital aetiology or risk factors for developing epilepsy listed in table 1 (0 points if not);
- 0 points for comorbidities (adjusting variable).

Each participant could, therefore, have a SEDS Score of 0–3. These scores were analysed in a multivariable model (SEDS Score model) adjusted for the number of comorbidities, using a score of zero points as the reference category. The results of this model are summarised in Table 6-6.

Table 6-6 – SEDS Score model results

SEDS Score	Grouping	Original data OR (95% CI)	P Value	Bootstrap P Value	Multiple imputation data OR (95% CI)	P Value
0 points		Reference			Reference	
1 point		**3.6 (1.9–6.8)**	**0.000**	**0.001**	**3.9 (1.9–8.1)**	**0.000**
2 points		**17.2 (7.4–39.6)**	**0.000**	**0.001**	**18.1 (7.6–43.2**	**0.000**
3 points		**19.8 (5.1–76.6)**	**0.000**	**0.001**	**24.2 (6.1–95.7)**	**0.000**
Comorbidities (0 points)	No	Reference			Reference	
	One condition	1.4 (0.7–2.7)	0.309	0.364	1.3 (0.7–2.5)	0.462
	Two or more conditions	1.8 (0.9–3.7)	0.083	0.116	1.8 (0.9–3.5)	0.098

Key: Point scoring is available for
⇒ Living in the two most deprived SIMD quintiles (1 point);
⇒ Having experienced at least one A&E attendance or hospital admission related to epilepsy/seizures in the preceding year (1 point);
⇒ Having at least one inherited or congenital aetiology or risk factor for developing epilepsy listed in table 6-1
⇒ 0 points are awarded for comorbidities (adjusting variable).

Abbreviations: OR – Odds ratio; CI – Confidence interval

Compared to having a SEDS Score of zero, the odds of suffering an epilepsy-related death were significantly elevated in those who had a score of 1–3. The suggestion of a dose-response in increasing odds was seen, with OR increasing from 3.5, to 17.1, to 19.8 for a SEDS Score of 1, 2 and 3, respectively. However, there were increasingly fewer cases and controls as SEDS Score increased, which acted to widen the CIs such that it was no longer possible to differentiate OR between a SEDS Score of 2 and a SEDS Score

of 3 using the original data. Both categories were associated with a significantly higher OR than a SEDS Score of 1, but this was lost when imputed values were included. There was otherwise little difference between the original and imputed data models results for SEDS Score.

SEDS Scores 1–3 each remained significant predictors ($p < 0.05$) through 1,000 bootstrap samples, suggesting internal validity for the SEDS Score model (Table 6-6). The model was also a good fit for the data ($p > 0.05$ for the Hosmer-Lemeshow goodness-of-fit test both using the original data and imputed data). 26% of the variance in the outcome was explained by the predictor variables included in this model ($R^2 = 0.264$). The model also demonstrated good calibration, with the predicted estimates mirroring the actual estimates closely throughout the Hosmer-Lemeshow contingency table (Table 6-5). In terms of discrimination, the model demonstrated a sensitivity of 71% (95% CI 64–77%) and specificity of 66% (95% CI 59–72%) using the original data, indicating that the majority of cases and controls were identified. There was little difference in these discrimination figures when imputed values were included (Appendix S2).

6.4 Discussion

In the first reported case-control study investigating epilepsy-related death in Scotland, data from the medical records of 224 deceased cases and living controls from multiple regional centres were analysed to identify multiple potential risk factors for epilepsy-related death and to develop a pragmatic risk prediction model for identifying those at the highest risk.

6.4.1 Risk factors

Seven variables tested in a univariable analysis were found to predict epilepsy-related death. In decreasing order of ORs, these were: recent seizure- or epilepsy-related A&E attendance or hospital admission (OR 5.1, 95% CI 3.2–8.3), a recent history of alcohol abuse (OR 4.4, 95% CI 2.2–9.2), absent recent neurology review (OR 3.8, 95% CI 2.4–6.1), the presence of inherited or congenital aetiology for developing epilepsy (OR 3.1,

95% CI 1.7–5.7), living in the two most deprived Scottish Quintile areas (OR 2.5, 95% CI 1.6–4.0), generalised epilepsy (OR 1.9, 95% CI 1.2–3.0), and a recent history of mental health problems (OR 1.6, 95% 1.0–2.6). The CIs for all of these estimates overlapped, meaning that whilst we could conclude that they are each likely to be significantly associated with the occurrence of epilepsy-related death, we could not confidently report which the strongest predictors were.

Aside from a history of inherited or congenital aetiology and risk factors for developing epilepsy or suffering from generalised epilepsy, the remaining predictors were potentially modifiable. Whilst it is well established in literature that alcohol abuse, congenital or developmental epilepsies, social deprivation, generalised epilepsy, and mental health problems are risk factors for all-cause mortality in PWE,[12] we show that they can be further streamlined into predicting *epilepsy-related* mortality specifically. Furthermore, we test the role of recent A&E attendance or hospital admission for seizures and epilepsy on predicting epilepsy-related mortality, and the role of absent neurology review within a year on this outcome. There is little study of these two variables in epilepsy mortality literature.[12] Our data suggest the odds of suffering epilepsy-related death may be up to five-times increased in those attending A&E or admitted to hospital for seizures/epilepsy within a year, and up to three times increased in those not reviewed by a neurologist within a year. The SUDEP and Seizure Safety Checklist is a free evidence-based tool supporting clinicians in discussing risk with PWE, with over 1,000 clinician subscribers in the UK currently.[51] It lists risk factors linked to epilepsy mortality including (but not restricted to) SUDEP. These risk factors were individually sourced from previous studies of mortality in PWE, and include, for example, active seizures, GTCS, status epilepticus, nocturnal seizures, poor medicine adherence, alcohol or substance misuse, and depression.[51] Following on from the findings of our study, it may be reasonable to consider including A&E attendance or hospital admission within the preceding year for seizures/epilepsy, or

including absent neurology review within a year in future iterations of the SUDEP and Seizure Safety Checklist.

6.4.2 Risk prediction model

To our knowledge, the SEDS Score would be the first multivariable risk prediction model for epilepsy-related deaths proposed in the literature.[12] It takes everyday clinical information that is widely available in medical records (home postcode, prior hospital attendance, and the already established aetiology or risk factors for a patient's epilepsy) and converts it into a simple 0–3-point score clinicians could use to flag up a patient at potentially higher risk of suffering epilepsy-related death. This could be used to help streamline patients into outpatient epilepsy services and to counsel them about risk, perhaps alongside the SUDEP and Seizure Safety Checklist as this was not developed using regression modelling of the proposed risk factors against one another, but rather as expert consensus on risk factors identified in literature. Table 6-7 illustrates the potential avenue to clinical implementation of the SEDS Score as an example risk scoring card for frontline clinicians to use. Although based on a relatively small sample, the SEDS Score modelling gives a first indication that notwithstanding comorbidity burden, patients progressively scoring 1–3 points for any combination of living in the two most deprived SIMD quintiles, having attended A&E or hospital for seizures or epilepsy in the preceding year, and having at least one inherited or congenital aetiology for developing epilepsy may be at substantially high risk of epilepsy-related death. This merits further investigation in a larger sample, as well as external validation in a similar cohort of patients elsewhere (e.g. using the SAIL databank in Wales).[229,249] Following external validation, the results could go on to play a role in helping to target increased resources toward PWE attending hospital because of seizures, from more deprived areas, with or without inherited/congenital aetiology for their epilepsy.

Table 6-7: SEDS Scoring tool (prototype)

INSTRUCTIONS: This is a mortality risk screening tool to help healthcare professionals seeing patients with a known diagnosis of epilepsy. It can be applied when seeing a patient with epilepsy in any setting, whether or not the presentation is related to their epilepsy. Screen the patient with the signalling questions, mark the responses, and award the relevant number of points. Add the points up and follow the points-based management plan suggested.

Signalling questions	Response	Number of points awarded
Does the patient live in one of the two most deprived SIMD quintiles?[A]	No	0
	Yes	1
[A] *SIMD quintiles take a number between 1–5.* *Quintiles 1–2 = the two most deprived.* *To calculate SIMD quintile, simply enter the patient's postcode into column A here - https://bit.ly/3qTxvpC, and look for the result in column F*		
Has the patient experienced at least one A&E attendance or hospital admission because of a seizure or related to their epilepsy within the last 12 months?	No	0
	Yes	1
Does the patient have an inherited or congenital aetiology/risk factor for their epilepsy?[B]	No	0
	Yes	1
[B] *Includes febrile convulsions, first degree relative with epilepsy, congenital abnormality/malformation (e.g. cerebral palsy, metabolic infancy syndrome, birth hypoxia), genetic syndrome, ADHD, ASD, developmental/intellectual delay, premature birth, birth/perinatal difficulties, hydrocephalus, neonatal seizures*		
How many comorbidities does the patient have alongside their epilepsy?[C]	None	0
	1 condition	0
	≥2 conditions	0
[C] *This includes any comorbidities listed in the medical record regardless of organ system affected or whether they are physical or mental*		
	Total number of points	

Interpretation of total number of points:

0 points = reference category

1 point = patient has ~4x increased risk of premature epilepsy-related death within seven years

2 points = patient has ~17x increased risk of premature epilepsy-related death within seven years

3 points = patient has ~20x increased risk of premature epilepsy-related death within seven years

Points-based management plan suggested:

- Email/fax the results of this assessment to your local epilepsy specialist team. **All patients with ≥1 point will have epilepsy clinic follow-up arranged within six months.** All other patients will have epilepsy clinic follow-up arranged within 12 months.
- Provide all patients with a phone number for the local epilepsy nurse specialist team (Edinburgh = 0131 312 0717, Glasgow = 0141 232 7601, Western Isles = 01851 763308)

Abbreviations: SIMD - Scottish Index of Multiple Deprivation; ADHD - Attention deficit hyperactivity disorder; ASD - Autism spectrum disorder

6.4.3 Strengths and limitations

The main strength of this study is that we use every day clinical information routinely available to clinicians in their medical records to risk-stratify patients into those that are potentially at higher risk of epilepsy-related death. We use a well-defined group of people suffering epilepsy-related death captured in our prior diagnostic accuracy validation chapter followed by an in-depth analysis of causes of death.[212] The study continues to benefit from the world-leading mortality review system in place in Scotland in which causes of death are routinely scrutinised for accuracy and amended where necessary.[60,61] To our knowledge, this is the first study to produce an epilepsy-related mortality scoring system for risk stratification (SEDS Score), which, if validated, would be an easy-to-use and pragmatic tool for front-line clinicians.

The study is mainly limited by sample size, which allowed only the development of a four-variable risk prediction model. Sample size was a trade-off against in-depth and quality analysis of individual patient medical records, which was prioritised to allow a breadth of risk factors to be ascertained for each case and control recruited. In order to optimise the quality of clinical information captured, the medical records were reviewed and abstracted only by an experienced consultant epileptologist (S.E.D). There were several downstream effects of smaller sample size. These include that we were, therefore, unable to match patients specifically by centre, but rather pooled them all into one Scottish centre. This increased the chances of identifying suitable matches for each case within the context of a small sample already matched by age and sex. Ascribing one centre is unlikely to have had a major impact on the results as care and services in Scotland are relatively uniform between centres, with neurology training, for example, recruited to as a single Scottish deanery. Given the sample size, we were also unable to stratify groups by additional characteristics such as by type of comorbidity (e.g. by known cardiac disease, which can interact with sudden death) or by type of AED (with some AEDs increasing risks of e.g. weight gain and metabolic syndrome) as these would have led to group sizes too

constricted to allow meaningful comparison in the multivariable models. It would be reasonable to investigate the impact of these potential confounders in future studies dedicated to their investigation.

Another limitation was that controls were drawn from those signing up to SHARE and those attending epilepsy clinics, yet no such selection process applied to cases. This may have created a selection bias, especially towards controls being generally healthier than cases. This is unlikely to have influenced the conclusions we drew from SEDS Scoring as the OR figures for this were markedly high, and would therefore be unlikely to have been entirely attributable to differences in baseline characteristics of the groups.

Our model was a good fit for the data statistically, as demonstrated by a p-value above 0.05 on the Hosmer-Lemeshow goodness-of-fit test. However, the absolute figures for discrimination through sensitivity and specificity estimates were moderate, between 65–80%. This highlights some of the complexities of regression model interpretation, which measures predictive ability in several different ways. When the aim is to use model predictions to inform patients or physicians to make decisions, the most important marker of model performance is calibration,[250] which was demonstrated highly in our model by the predicted estimates mirroring the actual estimates closely throughout the Hosmer-Lemeshow contingency table (Table 6-5). This is reassuring for the clinical conclusions drawn from this model. Related to this was the observation that variables included in the SEDS Score model had differing ORs from each other in the initial multivariable results (Table 6-4). However, we still chose to weight the variables equally in the SEDS Score modelling process (Table 6-6), as opposed ascribing each variable a weight based on its OR. This was done to maximise clinical utility of the SEDS Score given that a single-point scoring system is likely to be much easier for busy clinicians to use than having to calculate statistical weights for each variable used.

Other limitations included a moderate missing data burden meaning we had to include values obtained from multiple imputation as part of the analysis. We minimised the potential impact of this limitation overall by excluding variables with ≥30% of their values missing (or >10% difference in missing data burden between cases and controls), and we therefore saw little difference in the results between analysis with and without imputed values.

Finally, the study was designed to identify those at increased risk of epilepsy-related deaths as a group, but made no attempt to risk stratify according to the type of epilepsy-related death specifically, e.g. SUDEP, drowning, accidents or suicide. Epilepsy-related deaths are seldom investigated together as a group yet we have shown, in Chapter Five, that a substantial proportion of this group may be avoidable. A future study could attempt to further stratify by type of epilepsy-related death following on from our demonstration of increased risks for the overall group.

6.5 Conclusion

We propose the SEDS Score as a potential tool for identifying adults at increased risk of epilepsy-related death based on their recent attendance history to hospital for seizures or epilepsy, area of deprivation, epilepsy aetiology profile, and adjusted for their comorbidity count. This 3-point score needs external validation in another healthcare system.[229] The SEDS Score has the potential to be applied clinically to help streamline frontline care for PWE and to inform clinical discussions with patients about risk. We describe additional risk factors to consider as part of the overall analysis of epilepsy-related mortality risk in PWE, including recent review by a neurologist, alcohol abuse, and mental health problems. There is the potential for a larger study to try to incorporate these into future iterations of the SEDS Score and for them to be included in other tools such as in the SUDEP and Seizure Safety Checklist.[51]

Chapter 7. Discussion

Student contributions: conception, drafting and revision.

7.1 Overview

This thesis explores the concept of epilepsy-related deaths in adults. We review literature on this subject in order to gain insight into the likely burden of epilepsy-related deaths globally, before focusing primary investigation on deceased adults in Scotland. We augment the primary investigation by establishing how accurately epilepsy can be identified within administrative healthcare datasets and propose various algorithms to optimise this accuracy. There are five specific aims we wished to address, highlighted with their summative observations below, alongside suggested areas for future research.

7.2 Estimating the global burden of all-cause as well as epilepsy-related mortality in PWE

7.2.1 Study observations

The method we used to address this aim was to undertake a novel systematic review of epilepsy mortality systematic reviews, reported in Chapter Two. This was designed to evaluate comparative risks, causes, and risk factors for all-cause mortality in PWE in order to help establish the likely burden of epilepsy-related deaths specifically. The systematic review we undertook drew from a large body of 103 primary epilepsy mortality studies, brought together through the evaluation of six moderate- or high-quality systematic reviews. The work demonstrates that the global burden of all-cause mortality in PWE is high, with median SMR ranging 2.2–3.4. Alarmingly, these mortality figures do not seem to have reduced over time, remaining similarly high since the 1950s. Children and young adults with epilepsy appear to be at the highest comparative risk of premature death.

In terms of establishing the burden of epilepsy-related deaths specifically, we applied a novel method of cause-reclassification using a recently published system to help us identify that a substantial proportion of deaths in PWE were probably epilepsy-related. Specifically, epilepsy-related causes constituted 60% (12 out of 20) of the reasons PWE

died. Furthermore, SMR results tended to be higher for epilepsy-related causes (median 3.8, range 0.0–82.4), than for unrelated causes (median 1.7, range 0.7–17.6).

7.2.2 Areas for future research

One of the advantages of the permissive classification system that was applied was that it allowed us to consider the importance of several causes of death perhaps not usually considered to be epilepsy-related in scientific literature. In this way, we were able to suggest that although SUDEP is a common epilepsy-related cause of death, contributing 12% of deaths measured with PMR estimates, other epilepsy-related causes were perhaps as lethal as SUDEP, if not more. These included drowning (PMR 12%) and pneumonia (PMR 30%). The difficulty of generalising this conclusion is that PMRs tended to be measured in low-income countries, whilst high-income countries tended to favour using SMR to report the different causes of death. There were no SMR estimates reported for SUDEP in any study. This makes it difficult, therefore, to know how SUDEP compared with other non-SUDEP epilepsy-related deaths within high-income countries. This is an area that will require further investigation, particularly because the disease-specific nature of SUDEP deaths makes it invalid to compare SUDEP populations to the general population by indirect standardisation to produce an SMR. An innovative approach will be required to help contextualise SUDEP amongst the remaining epilepsy-related causes of death in high-income countries. One approach could be to compare mortality rates using person-years at risk between SUDEP and other non-SUDEP epilepsy-related deaths. However, this is likely to be costly as it would require a long-term prospective design. We have proposed the use of MCOD data analysis as another option. This approach allows more in-depth cross-sectional evaluation of causes of death in order to establish how the burden of SUDEP compares to non-SUDEP epilepsy-related causes within a population. This could be tried in other high-income countries with centralised administrative mortality records.

7.3 Estimating the overall accuracy of using administrative healthcare data to identify epilepsy cases

7.3.1 Study observations

The method we used to address this aim was to undertake a systematic review of studies assessing the accuracy of using administrative healthcare data to identify epilepsy cases. We assessed the performance of ICD-9, ICD-10, and Read codes, with or without AEDs. The primary outcome measures of interest were PPV and sensitivity, as these are the most commonly reported measures in diagnostic accuracy studies of administrative epilepsy data. However, we also reported NPV and specificity, where available. A narrative synthesis of review findings was used primarily for all outcome measures, although meta-analysis of sensitivity and specificity was also reported. Thirty studies were evaluated, spread across several continents.

We show that administrative healthcare data generally capture epilepsy with a high degree of accuracy, with the majority of studies managing to achieve PPV and sensitivity figures above 80%. The NPV and specificity estimates were even higher, mostly above 90%, albeit fewer studies measured these. The high accuracy figures were consistent between different countries, study designs, and coding systems. However, in the context of a known trade-off relationship between achieving high PPV and sensitivity, we were able to show that PPV was optimised by combining disease codes (ICD-10 G40–41, ICD-9 345) with AEDs, whilst sensitivity was optimised by combining symptom codes (ICD-10 R56; ICD-9 780.3, 780.39) with AEDs. AEDs alone were generally ineffective at capturing epilepsy because of a heavy burden of false positive results.

7.3.2 Areas for future research

It is clear most diagnostic accuracy studies of administrative epilepsy data report PPV and sensitivity. We have already discussed that PPV (and NPV) estimates depend on disease prevalence, making it generally invalid to combine these measures into a formal meta-analysis. Another difficulty in assessing PPV alongside sensitivity estimates is their

known trade-off relationship, where one tends to become lower as the other increases. This makes it difficult to be sure what the optimal balance of high PPV and sensitivity readings is in order to call a dataset 'accurate'. One way for a future systematic review of administrative epilepsy validation studies to deal with this tension would be to calculate the critical success index (CSI) of each study algorithm.[251-254] This is defined as the ratio of hits to the sum of hits, false alarms, and misses, i.e.

CSI = TP/(TP + FN + FP),

where TP = true positives, FN = false negatives, and FP = false positives.[251-254] These are the same metrics as are required to calculate PPV and sensitivity, i.e.

PPV = TP/(TP+FP),

Sensitivity = TP/(TP+FN).

However, in CSI, the metrics have been combined to provide a single measure that takes them all into account. CSI is generally used in weather forecasting, as a verification measure of categorical forecast performance.[251-254] In this, observed weather results are compared to forecast weather results. The derived CSI value ranges from 0 to 1, interpreted as 0 = unable to forecast and 1 = perfect forecast.[251-254] Whilst CSI has been used extensively for this purpose in meteorology literature,[255-262] there is little translational work incorporating CSI into diagnostic medical literature, with perhaps only one previous study reported.[263] Given, as we have demonstrated, the heavy reliance on PPV and sensitivity in studies validating administrative epilepsy datasets *particularly*, this would be a good area to help serve as proof-of-concept in applying CSI to diagnostic medical literature. It could also be that CSI would be more appropriate than receiver operating characteristic (ROC) plot analysis for studies validating administrative epilepsy datasets in future. This is because CSI calculation does not take true negatives into account, which are a prerequisite for ROC calculation through plotting sensitivity on the ordinate against 1 – specificity on the abscissa, where specificity = TN/(TN+FP).[254,264,265] The reason CSI does not take true negatives into account is that CSI was developed to avoid over-inflation

(i.e. undue optimism) of test metrics as a consequence of very large numbers of true negatives (i.e. class imbalance), which can occur for metrics such as specificity.[254] Such class imbalance was seen in our systematic review, where specificity estimates were all nearly 100% across most studies, owing to heavy burdens of true negative results. These would likely act to systematically bias any future attempts to determine ROC estimates across administrative epilepsy datasets.

7.4 Identifying the optimal algorithms for identifying epilepsy cases in deceased adults in Scotland using administrative data

7.4.1 Study observations

The method we used to address this aim was to undertake a diagnostic accuracy study of linked administrative data pertaining to mortality records, hospital admissions, AED prescriptions, and primary care attendances in Scotland. We focused on validating cause-of-death ICD-10 codes G40–41 and R56.8, which were tested on their own and in combination with hospital admission with the same codes, AEDs, and *F25* primary care epilepsy Read codes. These other datasets were also validated on their own and in combination with each other. Medical records were used as the reference standard. 159,032 deceased potential epilepsy cases were captured across the four datasets, and medical records were reviewed in a random sample of 936, confirming epilepsy as present and absent in 614 and 322 participants, respectively.

The optimal case-ascertainment strategy involving cause-of-death codes combined ≥1 G40–41 and/or R56.8 code with AEDs, demonstrating PPV of 91% and sensitivity of 81%. This captured 1,921 deceased adults, who we were able to study in the next chapter, which was an in-depth analysis of adult epilepsy-related deaths in Scotland.

Other case-ascertainment strategies to demonstrate high accuracy were, when using only one dataset, to use a carefully selected group of AEDs commonly prescribed for epilepsy. When applied in this manner, these reached PPVs and sensitivities of 90% and 91%, respectively. When using two datasets excluding cause-of-death coding, a combination

of *F25* primary care Read codes combined with this group of AEDs achieved PPVs and sensitivities of 86% and 93%, respectively. Combining three datasets was ineffective at obtaining high sensitivities despite very high PPV, and we therefore cautioned against this approach.

7.4.2 Areas for future research

In a preliminary screen of primary data, we show that CSI scores closely match the choices we made to select ≥1 G40–41 and/or R56.8 code with AEDs as the optimal algorithm involving cause of death codes, and diagnostic primary care codes with AEDs as the optimal algorithms not involving cause of death codes (see table 7-1). These algorithms were associated with the highest CSI scores. This helps to further advocate for the potential role CSI analysis could play in future diagnostic accuracy studies of administrative epilepsy datasets.

Table 7-1: Results of Level 2 validation of algorithms combining two database coding or AED strategies together (CSI results included)[212]

	Database	Coding algorithm	TP	FP	FN	PPV (95% CI)	Sn (95% CI)	CSI
1*	**Primary care + PIS**	**≥1 F25 + ≥1 AED (NL)**	**111**	**18**	**9**	**86% (80–92%)**	**93% (88–97%)**	**0.80**
2	Primary care + PIS	≥1 F25 + ≥1 AED (BL)	112	27	8	81% (74–87%)	93% (89–98%)	0.76
3*	**NRS + PIS**	**≥1 G40-41, R56.8 + ≥1 AED (NL)**	**494**	**47**	**120**	**91% (89–94%)**	**81% (77–84%)**	**0.75**
4	NRS + PIS	≥1 G40-41, R56.8 + ≥1 AED (BL)	507	98	107	84% (81–87%)	83% (80–86%)	0.71
5	NRS + PIS	≥1 G40-41 + ≥1 AED (NL)	419	28	195	94% (92–96%)	68% (65–72%)	0.65
6	NRS +PIS	≥1 G40-41 + ≥1 AED (BL)	426	42	188	91% (88–94%)	69% (66–73%)	0.65
7	SMR01 + PIS	≥1 G40-41, R56.8 + ≥1 AED (NL)	418	34	196	93% (90–95%)	68% (64%–72%)	0.65
8	SMR01 + PIS	≥1 G40-41, R56.8 + ≥1 AED (BL)	424	48	190	90% (87–93%)	69% (65–73%)	0.64
9	NRS + PIS	≥1 G40 + ≥1 AED (BL)	407	31	207	93% (91–95%)	66% (63–70%)	0.63
10	Primary care + SMR01	≥1 F25 + ≥1 G40-41, R56.8	85	15	35	85% (78–92%)	71% (63–79%)	0.63
11	NRS + PIS	≥1 G40 + ≥1 AED (NL)	402	26	212	94% (92–96%)	66% (62–69%)	0.63
12	SMR01 + NRS	≥1 G40-41, R56.8	386	76	228	84% (80–87%)	63% (59–67%)	0.56
13	Primary care + NRS	≥1 F25 + ≥1 G40-41, R56.8	68	8	52	90% (83–96%)	57% (48–66%)	0.53
14	SMR01 + NRS	SMR01 ≥1 G40-41, R56.8 + NRS ≥1 G40-41	347	42	267	89% (86–92%)	57% (53–60%)	0.53
15	SMR01 + NRS	SMR01 ≥1 G40-41, R56.8 + NRS ≥1 G40	325	23	289	93% (91–96%)	53% (49–57%)	0.51
16	Primary care + NRS	≥1 F25 + ≥1 G40-41	63	8	57	89% (81–96%)	53% (44–61%)	0.49

17	Primary care + NRS	≥1 F25 + ≥1 G40	62	8	58	89% (81–96%)	52% (43–61%)	0.48
18	NRS + PIS	≥1 R56.8 + ≥1 AED (BL)	95	52	519	65% (57–72%)	16% (13–18%)	0.14
19	NRS + PIS	≥1 R56.8 + ≥1 AED (NL)	89	17	525	84% (77–91%)	15% (12–17%)	0.14
20	SMR01 + NRS	SMR01 ≥1 G40-41, R56.8 + NRS ≥1 R56.8	58	30	556	66% (56–76%)	9% (7–12%)	0.09
21	Primary care + NRS	≥1 F25 + ≥1 R56.8	8	0	112	100% (100–100%)	7% (2–11%)	0.07
22	SMR01 + NRS	SMR01 ≥1 G40-41, R56.8 + NRS ≥1 G41	36	21	578	63% (51–76%)	6% (4–8%)	0.06
23	NRS + PIS	≥1 G41 + ≥1 AED (BL)	32	13	582	71% (58–84%)	5% (4–7%)	0.05
24	Primary care + NRS	≥1 F25 + ≥1 G41	X	X	X	89% (81–96%)	53% (44–61%)	X
25	NRS + PIS	≥1 G41 + ≥1 AED (NL)	X	X	X	88% (77–99%)	5% (3–7%)	X

Key: * = Optimal case-ascertainment strategy(s); X – categories with five or less events hidden to protect patient identity

Abbreviations: SMR – Scottish Morbidity Record; PIS – Prescription Information Service; NRS – National Records of Scotland; AED – antiepileptic drug; F25 – primary care diagnostic Read Codes for epilepsy; G40–41 – International Classification of Disease 10 (ICD-10) codes for epilepsy and status epilepticus; R56.8 – ICD-10 code for seizures; NL – AEDs on the narrow list (appendix S2b in Chapter 4); BL – AEDs on the broad list (appendix S2a in Chapter 4); TP – true positive; FP – false positives; FN – false negatives, PPV – positive predictive value; Sn – sensitivity; CI – confidence intervals; CSI - critical success index

7.5 Using administrative data from adults to investigate the trends and mechanisms of epilepsy-related deaths in Scotland, highlighting the proportion that were potentially avoidable

7.5.1 Study observations

The method we used to address this aim was to undertake a retrospective observational data-linkage study of administrative data from 2009–2016. Data were linked from mortality records, hospital admissions, outpatient attendance, AED prescriptions, and regional primary care attendances. Cases were ascertained using an optimal algorithm identified in Chapter 4, combining epilepsy cause-of-death coding (≥1 G40–41 and/or R56.8 code) with AEDs and generating 1,921 cases of epilepsy-related death for study. These were studied using MCOD data in order to increase the depth of cause-of-death information available, allowing us to describe all of the mechanisms contributing to the occurrence of epilepsy-related death. We were also able to describe the burden of potentially avoidable epilepsy-related deaths using methods provided by the ONS.

In terms of trend, the rate of epilepsy-related deaths remained statistically unchanged between 2009–2016, ranging from 5.9–8.7 per 100,000 Scottish population. The epilepsy-related deaths were significantly increased in young adults aged 16–54 years, with

mortality ratios of 2.3 (95% CI 1.8–2.8) overall, peaking at ages 16–24 years (mortality ratio 5.3, 95% CI 1.8–8.8). There were 553 young adult epilepsy-related deaths.

In terms of mechanisms, 166/553 young adult epilepsy-related deaths were SUDEP (i.e. 30%). This means that the majority (387/553, i.e. 70%) of young adults suffering an epilepsy-related death did not die from SUDEP. The common non-SUDEP epilepsy-related causes of death identified included aspiration pneumonia, cardiac arrest, AED or narcotic poisoning, drowning, and alcohol dependence.

76% of young adult epilepsy-related deaths were potentially avoidable, largely owing to the involvement of epilepsy itself as a cause of death.

7.5.2 Areas for future research

It may be surprising that the non-SUDEP epilepsy-related causes of death were more common than SUDEP in young adults. This may lead to enquiry as to whether or not SUDEP was accurately captured in this work. There can be little doubt that SUDEP was captured accurately as 90% of the SUDEP cases underwent autopsy. We also took an additional step to highlight the number of non-SUDEP epilepsy-related deaths that were sudden, suspicious, accidental or unexplained (evident because they underwent legal investigation at The Scottish Fatalities Investigation Unit (SFIU) of the Procurator Fiscal Service) but did not proceed to autopsy. We then provided an estimate of what the SUDEP count would have been if these were reclassified as SUDEP had they undergone autopsy. This additional information provides a “worst-case” scenario of what the maximum number of SUDEP cases would have been in the unlikely event there was underestimation occurring (because all sudden/suspicious deaths in Scotland have to go through the SIFU). Even then, the new figure of 254 SUDEP cases out of 553 young adult epilepsy-related deaths (i.e. 45%) does not change our conclusions about a predominance of non-SUDEP epilepsy related deaths in this group. This is an emerging view in literature, as was discussed in Chapter One.[23] The findings of our systematic review in Chapter Two

were also supportive of this idea.[12] This is an important area to research further as it potentially opens up additional targets for epilepsy mortality advice, risk-stratification, and prevention beyond the SUDEP-specific targets, which tend to predominate research.[23]

Further work could also be undertaken in future to try to characterise the nature of deaths in people with a diagnosis of epilepsy but for whom epilepsy did not appear in the cause-of-death record. These would be the 5,980 deceased study participants with G40–41 and/or R56.8 hospital admissions codes for epilepsy in the SMR01 dataset, who were also prescribed ≥1 AED in the PIS dataset, but for whom no G40–41 and/or R56.8 cause-of-death code was reported in the NRS dataset (see Venn diagram 1B in Chapter 4). As intimated by Devinsky *et al.*,[23] some of such deaths could have been missed epilepsy-related deaths (e.g. where they were resultant from aspiration pneumonia or suicide).

Further work could also be done to try to study the 1,333 deceased study participants who were not on any AEDs to try and understand the nature of deaths in untreated people who may have had epilepsy. As shown in Venn diagram 1B in Chapter 4, 321 of these people had G40–41 and/or R56.8 cause of death coding alone, 143 had these codes combined with hospital admissions codes for epilepsy, and 869 had only hospital admissions codes for epilepsy (G40–41 and/or R56.8).

7.6 Identifying multiple risk factors for the occurrence of epilepsy-related death in Scottish adults and developing a pragmatic risk prediction model for identifying which individuals are at the highest risk

7.6.1 Study observations

The method we used to address this aim was to undertake an age- and sex-matched case-control study comparing adults who suffered epilepsy-related deaths to living adults with epilepsy in Scotland. Data pertaining to potential risk factors for epilepsy-related deaths were collected from the medical records of cases and controls to allow multivariable analysis and development of a risk prediction model. 224 cases were compared to 224 controls.

In terms of risk factors, univariable analysis revealed that seven variables were significantly associated with increased risk of suffering epilepsy-related death. These were having a recent A&E attendance or hospital admission for seizures or epilepsy, a recent history of alcohol abuse, absent recent neurology review, the presence of inherited or congenital aetiology for developing epilepsy, living in the two most deprived Scottish quintile areas, generalised epilepsy, and a recent history of mental health problems. ORs ranged between 1.6–5.1 between these variables.

Multivariable modelling was undertaken for the variables SIMD quintile; A&E attendance or hospital admission for epilepsy/seizures; the aetiology for developing epilepsy; and the number of comorbidities suffered. This model was a good fit for the data, and factors that retained significance for predicting epilepsy-related death were living in the two most deprived SIMD quintiles; having experienced at least one A&E attendance or hospital admission related to epilepsy/seizures in the preceding year, and having at least one inherited or congenital aetiology for developing epilepsy. ORs ranged 2.7–5.1 across these positive factors.

Each positive factor was then given one point, if present, to create a risk prediction model termed the SEDS Score. No points were given for the number of comorbidities as these were not significant predictors, but they were kept in the order to allow these to be adjusted for clinically. The SEDS Score modelling demonstrated that progressively scoring 1–3 points for any combination of the three risk factors of interest was associated with a substantially increased risk of suffering epilepsy-related death, with ORs increasing from 3.6 for one point, to 17.2 for two points, to 19.8 for three points. Wide CIs were noted surrounding these estimates, owing to increasingly fewer events as factors were combined.

7.6.2 Areas for future research

The SEDS Score is the first risk prediction model reported in epilepsy mortality literature to quantify the risk of suffering epilepsy-related deaths specifically, and to use multivariable modelling to quantify combinations of the risk factors. With the methods now provided, the SEDS Score can be developed and externally validated in larger datasets to help improve precision of the estimates by narrowing the CIs, and also to help refine or expand the risk factors selected. For example, it was surprising that increasing numbers of comorbidities did not translate into increased risk of suffering epilepsy-related death when certain comorbidities were shown to be risk factors for death in PWE in our systematic review in Chapter Two. However, we were limited in our ability to risk-stratify according to type of comorbidity in the case-control study because of limited sample sizes. Combining all comorbidities listed into a single measure, as we did, may have helped preserve the size of comparison groups to allow valid multivariable modelling but acted to dilute the effect of more relevant comorbidities, such as mental health problems, as these would have been weighted equally with less serious but common comorbidities listed, such as migraine or acid reflux. The univariable analysis demonstrated a clear signal that having recent mental health problems was significantly associated with increased risk of suffering epilepsy-related death. This is perhaps a tip of the iceberg given that our coded primary care dataset revealed large numbers of comorbidity codes for depression in adults suffering epilepsy-related deaths, but this information was probably not reflected consistently in inpatient medical records because specialist mental health records in Scotland are sequestered within psychiatry services, which we did not have access to in the current study. The association data also indicated that having mental health problems correlated with having seizure- and epilepsy-related A&E attendances or hospital admissions. Therefore, our findings help provide justification for the next phase of investigation to seek permission to assess specialist mental health records in Scotland to

try and help optimise the case-ascertainment signal in deceased PWE and potentially incorporate these into future iterations of SEDS Scoring.

Further work could be undertaken to try to capture AED adherence as an outcome for integration into future iterations of the SEDS Score as this is likely to have an important influence on mortality predictions. These data were largely missing from medical records in our study, both for cases and controls. The PIS dataset captures prescription and dispensing dates for all drugs, but we only had access to this dataset for cases rather than controls. It may be that future work will need to focus on trying to capture AED adherence using more standardised methods such calculating medicine possession ratios (MPRs).[266] These are ratios representing the proportion of days in an observation period that a patient is in possession of medicine supply, making an assumption that they are adherent and non-adherent on days with and without supply, respectively. MPRs require access to data on drug supply quantity and refill dates. A regional study has recently been published in Scotland using MPRs to capture AED adherence and mortality of PWE on opiate replacement therapy.[267]

7.7 Final conclusions

This thesis demonstrates that epilepsy-related deaths are a global problem that can be accurately quantified using administrative healthcare data for the purposes of more in-depth analysis. This analysis has been undertaken in Scotland and reveals that epilepsy-related deaths have failed to reduce with time, young adults are particularly vulnerable to these deaths when compared to age-mated mortality figures in the general population, and not all of the deaths are SUDEP. Many of the deaths are potentially avoidable. The risk factors for suffering epilepsy-related death may include living in deprived areas, hospital attendances for seizures and epilepsy, and inherited or congenital aetiologies for developing epilepsy. A novel risk prediction model is indicative that any combination of these may be associated with very high risks of suffering epilepsy-related death. Further work will be required to validate this model externally and to refine the variables used.

Other areas for future research include trying to unify the diagnostic accuracy measures in studies validating administrative epilepsy datasets, and advancing the conversation from investigating SUDEP alone to investigating epilepsy-related deaths as a whole.

References

1. Thijs RD, Surges R, O'Brien TJ, Sander JW. Epilepsy in adults. *Lancet.* 2019;393(10172):689-701.
2. World Health Organization. Epilepsy. *Fact sheet.* 2016. https://bit.ly/2ZGzdin. Accessed 10/02/2017.
3. Murray CJ, Vos T, Lozano R, et al. Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet.* 2012;380(9859):2197-2223.
4. ILAE/IBE/WHO Global Campaign Against Epilepsy. Epilepsy in the WHO European Region: Fostering Epilepsy Care in Europe 2010. https://bit.ly/3ggcsda. Accessed 11/06/2021.
5. Fisher RS, Acevedo C, Arzimanoglou A, et al. ILAE official report: a practical clinical definition of epilepsy. *Epilepsia.* 2014;55(4):475-482.
6. Tomson T, Beghi E, Sundqvist A, Johannessen SI. Medical risks in epilepsy: a review with focus on physical injuries, mortality, traffic accidents and their prevention. *Epilepsy Res.* 2004;60(1):1-16.
7. Gnanavel S. Epilepsy and Depression: A Bidirectional Relationship. *J Neurosci Rural Pract.* 2017;8(Suppl 1):S5-S6.
8. de Souza JL, Faiola AS, Miziara C, de Manreza MLG. The Perceived Social Stigma of People with Epilepsy with regard to the Question of Employability. *Neurol Res Int.* 2018;2018:4140508.
9. DVLA. Epilepsy and driving 2021. https://www.gov.uk/epilepsy-and-driving. Accessed 08/07/2021.
10. Fisher RS, Cross JH, D'Souza C, et al. Instruction manual for the ILAE 2017 operational classification of seizure types. *Epilepsia.* 2017;58(4):531-542.

11. Fisher RS, Cross JH, French JA, et al. Operational classification of seizure types by the International League Against Epilepsy: Position Paper of the ILAE Commission for Classification and Terminology. *Epilepsia.* 2017;58(4):522-530.
12. Mbizvo GK, Bennett K, Simpson CR, Duncan SE, Chin RFM. Epilepsy-related and other causes of mortality in people with epilepsy: A systematic review of systematic reviews. *Epilepsy Res.* 2019;157:106192.
13. Levira F, Thurman DJ, Sander JW, et al. Premature mortality of epilepsy in low- and middle-income countries: A systematic review from the Mortality Task Force of the International League Against Epilepsy. *Epilepsia.* 2016:1-11.
14. Neligan A, Bell GS. The mortality of epilepsy. In: Sander JW, Rugg-Gunn FJ, Smalls JE, eds. *Epilepsy 2009: From Benchside to Bedside - a Practical Guide to Epilepsy.* 12th ed. Chalfont St Peter: National Society for Epilepsy; 2009:pp 347-351.
15. Thurman DJ, Logroscino G, Beghi E, et al. The burden of premature mortality of epilepsy in high-income countries: A systematic review from the Mortality Task Force of the International League Against Epilepsy. *Epilepsia.* 2017;58(1):17-26.
16. Watila MM, Balarabe SA, Ojo O, Keezer MR, Sander JW. Overall and cause-specific premature mortality in epilepsy: A systematic review. *Epilepsy Behav.* 2018.
17. Malzberg B. *Mortality among patients with mental disease.* Utica, NY: State Hospital Press; 1934
18. Alstrom CH. *Mortality in mental hospitals with special regard to tuberculosis.* Uppsala: Appelsbergs Boktryckeriaktieborg; 1942.
19. Alstrom CH. A study of epilepsy in its clinical, social and genetic aspects. *Acta Psychiatr Neurol Suppl.* 1950;63:1-284.
20. Preston T, Clarke R. An investigation into the mortality of impaired lives during the period 1947-63. *Journal of the Institute of Actuaries.* 1966;92(1):27-74.

21. Henriksen PB, Juul-Jensen P, Lund M. The mortality of epileptics. *Acta Neurologica Scandinavica.* 1967;43(S31):164-165.

22. Voute PA. *Epilepsy and life insurance: social studies on epilepsy. No. 5. London.* International Bureau for Epilepsy; 1967.

23. Devinsky O, Spruill T, Thurman D, Friedman D. Recognizing and preventing epilepsy-related mortality: A call for action. *Neurology.* 2016;86(8):779-786.

24. Nashef L, So EL, Ryvlin P, Tomson T. Unifying the definitions of sudden unexpected death in epilepsy. *Epilepsia.* 2012;53(2):227-233.

25. Ali A, Wu S, Issa NP, et al. Association of sleep with sudden unexpected death in epilepsy. *Epilepsy Behav.* 2017;76:1-6.

26. Neligan A, Bell GS, Johnson AL, Goodridge DM, Shorvon SD, Sander JW. The long-term risk of premature mortality in people with epilepsy. *Brain.* 2011;134(Pt 2):388-395.

27. DeToledo JC, Lowe MR. Seizures, lateral decubitus, aspiration, and shoulder dislocation: Time to change the guidelines? *Neurology.* 2001;56(3):290-291.

28. DeToledo JC, Lowe MR, Gonzalez J, Haddad H. Risk of aspiration pneumonia after an epileptic seizure: a retrospective analysis of 1634 adult patients. *Epilepsy Behav.* 2004;5(4):593-595.

29. Piffaretti C, Moreno-Betancur M, Lamarche-Vadel A, Rey G. Quantifying cause-related mortality by weighting multiple causes of death. *Bull World Health Organ.* 2016;94(12):870-879.

30. Redelings MD, Wise M, Sorvillo F. Using multiple cause-of-death data to investigate associations and causality between conditions listed on the death certificate. *Am J Epidemiol.* 2007;166(1):104-108.

31. Chief Medical Officer. Guidance for Doctors Completing Medical Certificates of The Cause of Death (MCCD) and its Quality Assurance 2018. https://bit.ly/2yyMzlC. Accessed 03/03/2020.

32. Thurman DJ, Hesdorffer DC, French JA. Sudden unexpected death in epilepsy: assessing the public health burden. *Epilepsia.* 2014;55(10):1479-1485.

33. World Health Oraganisation. Years of life lost (percentage of total) 2006. https://bit.ly/3wn17O8. Accessed 12/06/2021.

34. Sillanpaa M, Shinnar S. Long-term mortality in childhood-onset epilepsy. *N Engl J Med.* 2010;363(26):2522-2529.

35. Bell GS, Gaitatzis A, Bell CL, Johnson AL, Sander JW. Drowning in people with epilepsy - How great is the risk? *Neurology.* 2008;71(8):578-582.

36. Bell GS, Gaitatzis A, Bell CL, Johnson AL, Sander JW. Suicide in people with epilepsy: how great is the risk? *Epilepsia.* 2009;50(8):1933-1942.

37. Hauser WA, Annegers JF, Elveback LR. Mortality in patients with epilepsy. *Epilepsia.* 1980;21(4):399-412.

38. Kaiboriboon K, Schiltz NK, Bakaki PM, Lhatoo SD, Koroukian SM. Premature mortality in poor health and low income adults with epilepsy. *Epilepsia.* 2014;55(11):1781-1788.

39. Neligan A, Sander JW. Premature mortality in epilepsy: is it preventable? *Expert Review of Neurotherapeutics.* 2011;11(6):767-770.

40. Sander, Bell. Reducing mortality: an important aim of epilepsy management. *J Neurol Neurosurg Psychiatry.* 2004;75(3):349-351.

41. Hanna NJ, Black M, Sander JW, et al. The National Sentinel Clinical Audit of Epilepsy-Related Death: Epilepsy – death in the shadows. *The Stationery Office.* 2002.

42. RCPCH. *Coordinating Epilepsy Care: a UK-wide review of healthcare in cases of mortality and prolonged seizures in children and young people with epilepsies. Child Health Reviews - UK.* London: RCPCH; 2013.

43. Katikireddi SV, Cezard G, Bhopal RS, et al. Assessment of health care, hospital admissions, and mortality by ethnicity: population-based cohort study of health-system performance in Scotland. *Lancet Public Health.* 2018;3(5):e226-e236.

44. Smith GS, Fleming M, Kinnear D, et al. Rates and causes of mortality among children and young people with and without intellectual disabilities in Scotland: a record linkage cohort study of 796 190 school children. *BMJ Open.* 2020;10(8):e034077.

45. Olatunde O, Windsor-Shellard B, Campbell A. Revised Definition of Avoidable Mortality 2016 https://bit.ly/3cggXSK. Accessed 30/10/2019.

46. Novak JL, Miller PR, Markovic D, Meymandi SK, DeGiorgio CM. Risk Assessment for Sudden Death in Epilepsy: The SUDEP-7 Inventory. *Front Neurol.* 2015;6:252.

47. DeGiorgio CM, Miller P, Meymandi S, et al. RMSSD, a measure of vagus-mediated heart rate variability, is associated with risk factors for SUDEP: the SUDEP-7 Inventory. *Epilepsy Behav.* 2010;19(1):78-81.

48. Walczak TS, Leppik IE, D'Amelio M, et al. Incidence and risk factors in sudden unexpected death in epilepsy: a prospective cohort study. *Neurology.* 2001;56(4):519-525.

49. Tarighati Rasekhi R, Devlin KN, Mass JA, et al. Improving prediction of sudden unexpected death in epilepsy: From SUDEP-7 to SUDEP-3. *Epilepsia.* 2021.

50. Jha A, Oh C, Hesdorffer D, et al. Sudden Unexpected Death in Epilepsy: A Personalized Prediction Tool. *Neurology.* 2021;96(21):e2627-e2638.

51. SUDEP Action. SUDEP and Seizure Safety Checklist. https://bit.ly/347i8iY. Published 2017. Accessed 07/11/2018.

52. Joint Epilepsy Council of the UK and Ireland. Epilepsy prevalence, incidence and other statistics. 2011.

53. Steel D, Cylus J. United Kingdom (Scotland): Health system review. *Health Syst Transit.* 2012;14(9):xv-xxii, 1-150.

54. Public Health England. Deaths associated with neurological conditions in England 2001 to 2014. *National Neurology Intelligence Network National End of Life Care Intelligence Network.* 2018. https://bit.ly/3sqym3F. Accessed 28/02/2022.

55. Pavis S, Morris AD. Unleashing the power of administrative health data: the Scottish model. *Public Health Res Pract.* 2015;25(4):e2541541.

56. McCallum G. Annual State of NHSScotland Assets and Facilities Report for 2014. *Health and social care.* https://bit.ly/3x45uOJ. Accessed 29/06/2021.

57. eDRIS Team. How do we link the data? 2020. https://bit.ly/39DJ4ZB. Accessed 27/07/2020.

58. HDRUK. Health Data Research UK. 2021. https://www.hdruk.ac.uk/about-us/. Accessed 08/07/2021.

59. Information Services Division. National Records of Scotland (NRS) - Deaths Data 2016. https://bit.ly/3CiWmYO. Accessed Date accessed: 27/10/2019.

60. National Records of Scotland. Some Specific Checks Carried Out on Deaths Data. *Vital Events – Deaths – Background Information* 2019. https://bit.ly/2F50Gm6. Accessed 05/09/2020.

61. Fernie CGM. Scotland already has world leading mortality review system. *Bmj-Brit Med J.* 2019;364.

62. Mbizvo GK, Bennett KH, Schnier C, Simpson CR, Duncan SE, Chin RFM. The accuracy of using administrative healthcare data to identify epilepsy cases: A systematic review of validation studies. *Epilepsia.* 2020;61(7):1319-1335.

63. Information Services Division. General Acute Inpatient and Day Case - Scottish Morbidity Record (SMR01). *National Data Catalogue.* 2016. https://bit.ly/3Elcc6K. Accessed 27/10/19.

64. Information Services Devision. SMR00. 2020 https://bit.ly/2Dko3ag. Accessed 31/07/2020.

65. Meeraus WH, Petersen I, Chin RF, Knott F, Gilbert R. Childhood epilepsy recorded in primary care in the UK. *Arch Dis Child.* 2013;98(3):195-202.

66. Douglas A, Cezard G, Simpson CR, et al. Pilot study linking primary care records to Census, cardiovascular hospitalization and mortality data in Scotland: feasibility, utility and potential. *J Public Health (Oxf).* 2015.

67. Abraha I, Montedori A, Eusebi P, et al. The Current State of Validation of Administrative Healthcare Databases in Italy: A Systematic Review. *Pharmacoepidemiology and Drug Safety.* 2012;21:400-400.

68. Jette N, Reid AY, Quan H, Hill MD, Wiebe S. How accurate is ICD coding for epilepsy? *Epilepsia.* 2010;51(1):62-69.

69. Loke YK. Use of databases for clinical research. *Archives of Disease in Childhood.* 2014;99(6):587-589.

70. eDRIS Team. Public Benefit and Privacy Panel for Health and Social Care - Why Does it Exist. 2020. https://bit.ly/3vnzFyJ. Accessed 15/06/2021.

71. Medical Research Council. MRC Regulatory Support Centre Learning Management System. 2022. https://bit.ly/3xiOFPA. Accessed 15/06/2021.

72. Becker L, Oxman AD. Overviews of reviews. *Cochrane Handbook for Systematic Reviews of Interventions Version 510 [updated March 2011] Higgins JPT, Green S (editors).* 2011. https://bit.ly/2U1YIKK. Accessed 25/06/2021.

73. Abdel-Mannan O, Taylor H, Donner EJ, Sutcliffe AG. A systematic review of sudden unexpected death in epilepsy (SUDEP) in childhood. *Epilepsy Behav.* 2018;90:99-106.

74. Ackers R, Besag FM, Hughes E, Squier W, Murray ML, Wong IC. Mortality rates and causes of death in children with epilepsy prescribed antiepileptic drugs: a retrospective cohort study using the UK General Practice Research Database. *Drug Saf.* 2011;34(5):403-413.

75. Belhocine M, de Boer HM, Mandlhate C. Epilepsy in the WHO African Region: Bridging the Gap *The Global Campaign Against Epilepsy "Out of the Shadows".* 2004. https://bit.ly/3K7pFB9. Accessed 04/07/2019.

76. Chamorro-Munoz MI, Garcia-Martin G, Perez-Errazquin F, Romero-Acebal M, Garcia-Rodriguez A, Gutierrez-Bedmar M. Epidemiological study of mortality in epilepsy in a Spanish population. *Seizure.* 2017;46:19-23.

77. Chang YH, Ho WC, Tsai JJ, Li CY, Lu TH. Risk of mortality among patients with epilepsy in southern Taiwan. *Seizure-Eur J Epilep.* 2012;21(4):254-259.

78. Christensen J, Pedersen CB, Sidenius P, Olsen J, Vestergaard M. Long-term mortality in children and young adults with epilepsy--A population-based cohort study. *Epilepsy Res.* 2015;114:81-88.

79. Clarke M. Systematic review of reviews of risk factors for intracranial aneurysms. *Neuroradiology.* 2008;50(8):653-664.

80. Corry M, While A, Neenan K, Smith V. A systematic review of systematic reviews on interventions for caregivers of people with chronic conditions. *J Adv Nurs.* 2015;71(4):718-734.

81. Day SM, Wu YW, Strauss DJ, Shavelle RM, Reynolds RJ. Causes of death in remote symptomatic epilepsy. *Neurology.* 2005;65(2):216-222.

82. Ding D, Wang W, Wu J, et al. Premature mortality in people with epilepsy in rural China: a prospective study. *Lancet Neurol.* 2006;5(10):823-827.

83. Ding D, Wang W, Wu J, et al. Premature mortality risk in people with convulsive epilepsy: long follow-up of a cohort in rural China. *Epilepsia.* 2013;54(3):512-517.

84. Donner EJ, Smith CR, Snead OC, 3rd. Sudden unexplained death in children with epilepsy. *Neurology.* 2001;57(3):430-434.

85. Escalaya AL, Tellez-Zenteno JF, Steven DA, Burneo JG. Epilepsy and mortality in Latin America. *Seizure.* 2015;25:99-103.

86. Fazel S, Wolf A, Langstrom N, Newton CR, Lichtenstein P. Premature mortality in epilepsy and the role of psychiatric comorbidity: a total population study. *Lancet.* 2013;382(9905):1646-1654.

87. Granbichler CA, Oberaigner W, Kuchukhidze G, et al. Cause-specific mortality in adult epilepsy patients from Tyrol, Austria: hospital-based study. *J Neurol.* 2015;262(1):126-133.

88. Harden C, Tomson T, Gloss D, et al. Practice guideline summary: Sudden unexpected death in epilepsy incidence rates and risk factors: Report of the Guideline Development, Dissemination, and Implementation Subcommittee of the American Academy of Neurology and the American Epilepsy Society. *Neurology.* 2017;88(17):1674-1680.

89. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 The Cochrane Collaboration; 2011.

90. Keezer MR, Bell GS, Sander JW. Epilepsy-related clinical characteristics and mortality: a systematic review and meta-analysis. *Neurology.* 2015;84(17):1823.

91. Liebenthal JA, Wu SS, Rose S, Ebersole JS, Tao JX. Association of prone position with sudden unexpected death in epilepsy. *Neurology.* 2015;84(7):703-709.

92. Logroscino G, Hesdorffer DC. Methodologic issues in studies of mortality following epilepsy: measures, types of studies, sources of cases, cohort effects, and competing risks. *Epilepsia.* 2005;46 Suppl 11:3-7.

93. Lunny C, Brennan SE, McDonald S, McKenzie JE. Toward a comprehensive evidence map of overview of systematic review methods: paper 1-purpose, eligibility, search and data extraction. *Syst Rev.* 2017;6(1):231.

94. Mbizvo GK, Dixon P, Hutton JL, Marson AG. Levetiracetam add-on for drug-resistant focal epilepsy: an updated Cochrane Review. *Cochrane Database Syst Rev.* 2012;9:CD001901.

95. Mohanraj R, Norrie J, Stephen LJ, Kelly K, Hitiris N, Brodie MJ. Mortality in adults with newly diagnosed and chronic epilepsy: a retrospective comparative study. *Lancet Neurol.* 2006;5(6):481-487.

96. Moher D, Liberati A, Tetzlaff J, Altman DG, Group P. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ.* 2009;339:b2535.

97. Monte CP, Arends JB, Tan IY, Aldenkamp AP, Limburg M, de Krom MC. Sudden unexpected death in epilepsy patients: Risk factors. A systematic review. *Seizure.* 2007;16(1):1-7.

98. Moseley BD, Wirrell EC, Wong-Kisiel LC, Nickels K. Early onset epilepsy is associated with increased mortality: a population-based study. *Epilepsy Res.* 2013;105(3):410-414.

99. Mu J, Liu L, Zhang Q, et al. Causes of death among people with convulsive epilepsy in rural West China: a prospective study. *Neurology.* 2011;77(2):132-137.

100. Mula M, Schmitz B. Depression in epilepsy: mechanisms and therapeutic approach. *Ther Adv Neurol Disord.* 2009;2(5):337-344.

101. Neligan A, Bell GS, Shorvon SD, Sander JW. Temporal trends in the mortality of people with epilepsy: a review. *Epilepsia.* 2010;51(11):2241-2246.

102. Nevalainen O, Ansakorpi H, Simola M, et al. Epilepsy-related clinical characteristics and mortality: a systematic review and meta-analysis. *Neurology.* 2014;83(21):1968-1977.

103. Nevalainen O, Auvinen A, Ansakorpi H, Artama M, Raitanen J, Isojarvi J. Mortality by clinical characteristics in a tertiary care cohort of adult patients with chronic epilepsy. *Epilepsia.* 2012;53(12):e212-214.

104. Nevalainen O, Raitanen J, Ansakorpi H, Artama M, Isojarvi J, Auvinen A. Long-term mortality risk by cause of death in newly diagnosed patients with epilepsy in Finland: a nationwide register-based study. *Eur J Epidemiol.* 2013;28(12):981-990.

105. Ngugi AK, Bottomley C, Fegan G, et al. Premature mortality in active convulsive epilepsy in rural Kenya: causes and associated factors. *Neurology.* 2014;82(7):582-589.

106. Rafnsson V, Olafsson E, Hauser WA, Gudmundsson G. Cause-specific mortality in adults with unprovoked seizures. A population-based incidence cohort study. *Neuroepidemiology.* 2001;20(4):232-236.

107. Ryan R, Santesso N, Lowe D, et al. Interventions to improve safe and effective medicines use by consumers: an overview of systematic reviews. *Cochrane Db Syst Rev.* 2014(4).

108. Samokhvalov AV, Irving H, Mohapatra S, Rehm J. Alcohol consumption, unprovoked seizures, and epilepsy: a systematic review and meta-analysis. *Epilepsia.* 2010;51(7):1177-1184.

109. Shackleton DP, Westendorp RG, Kasteleijn-Nolst Trenite DG, de Craen AJ, Vandenbroucke JP. Survival of patients with epilepsy: an estimate of the mortality risk. *Epilepsia.* 2002;43(4):445-450.

110. Shackleton DP, Westendorp RG, Trenite DG, Vandenbroucke JP. Mortality in patients with epilepsy: 40 years of follow up in a Dutch cohort study. *J Neurol Neurosurg Psychiatry.* 1999;66(5):636-640.

111. Shea BJ, Bouter LM, Peterson J, et al. External validation of a measurement tool to assess systematic reviews (AMSTAR). *PLoS One.* 2007;2(12):e1350.

112. Shea BJ, Grimshaw JM, Wells GA, et al. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. *BMC Med Res Methodol.* 2007;7:10.

113. Shea BJ, Hamel C, Wells GA, et al. AMSTAR is a reliable and valid measurement tool to assess the methodological quality of systematic reviews. *J Clin Epidemiol.* 2009;62(10):1013-1020.

114. Smith V, Devane D, Begley CM, Clarke M. Methodology in conducting a systematic review of systematic reviews of healthcare interventions. *BMC Med Res Methodol.* 2011;11(1):15.

115. Tellez-Zenteno JF, Ronquillo LH, Wiebe S. Sudden unexpected death in epilepsy: evidence-based analysis of incidence and risk factors. *Epilepsy Res.* 2005;65(1-2):101-115.

116. White SJ, McLean AE, Howland C. Anticonvulsant drugs and cancer. A cohort study in patients with severe epilepsy. *Lancet.* 1979;2(8140):458-461.

117. World Health Organization. Drowning. *Fact sheets.* 2018. https://bit.ly/3K1ukon. Accessed 03/07/2019.

118. Young C, Shankar R, Palmer J, et al. Does intellectual disability increase sudden unexpected death in epilepsy (SUDEP) risk? *Seizure-Eur J Epilep.* 2015;25:112-116.

119. Jette N, Beghi E, Hesdorffer D, et al. ICD coding for epilepsy: past, present, and future--a report by the International League Against Epilepsy Task Force on ICD codes in epilepsy. *Epilepsia.* 2015;56(3):348-355.

120. Bossuyt PM, Reitsma JB, Bruns DE, et al. STARD 2015: An Updated List of Essential Items for Reporting Diagnostic Accuracy Studies. *Clin Chem.* 2015;61(12):1446-1452.

121. Mbizvo GK, Bennett K, Simpson CR, Duncan SE, Chin RFM. Accuracy and utility of using administrative healthcare databases to identify people with epilepsy: a protocol for a systematic review and meta-analysis. *BMJ Open.* 2018;8(6):e020824.

122. Benson T. The history of the Read Codes: the inaugural James Read Memorial Lecture 2011. *Inform Prim Care.* 2011;19(3):173-182.

123. Christensen J, Vestergaard M, Olsen J, Sidenius P. Validation of epilepsy diagnoses in the Danish National Hospital Register. *Epilepsy Res.* 2007;75(2-3):162-170.

124. Deeks JJ, Macaskill P, Irwig L. The performance of tests of publication bias and other sample size effects in systematic reviews of diagnostic test accuracy was assessed. *J Clin Epidemiol.* 2005;58(9):882-893.

125. Florkowski CM. Sensitivity, specificity, receiver-operating characteristic (ROC) curves and likelihood ratios: communicating the performance of diagnostic tests. *Clin Biochem Rev.* 2008;29 Suppl 1:S83-87.

126. Fonferko-Shadrach B, Lacey AS, White CP, et al. Validating epilepsy diagnoses in routinely collected data. *Seizure.* 2017;52:195-198.

127. Franchi C, Giussani G, Messina P, et al. Validation of healthcare administrative data for the diagnosis of epilepsy. *J Epidemiol Community Health.* 2013;67(12):1019-1024.

128. Garratt E, Barnes H, Dibben C. Health administrative data: Exploring the potential for academic research. St Andrews: Administrative Data Liaison Service. 2010

129. Hottes TS, Skowronski DM, Hiebert B, et al. Influenza vaccine effectiveness in the elderly based on administrative databases: change in immunization habit as a marker for bias. *PLoS One.* 2011;6(7):e22618.

130. Kee VR, Gilchrist B, Granner MA, Sarrazin NR, Carnahan RM. A systematic review of validated methods for identifying seizures, convulsions, or epilepsy using administrative and claims data. *Pharmacoepidemiology and Drug Safety.* 2012;21:183-193.

131. Krarup LH, Boysen G, Janjua H, Prescott E, Truelsen T. Validity of stroke diagnoses in a National Register of Patients. *Neuroepidemiology.* 2007;28(3):150-154.

132. Lalmohamed A, Vestergaard P, Cooper C, et al. Timing of stroke in patients undergoing total hip replacement and matched controls: a nationwide cohort study. *Stroke.* 2012;43(12):3225-3229.

133. Leeflang MM, Deeks JJ, Takwoingi Y, Macaskill P. Cochrane diagnostic test accuracy reviews. *Syst Rev.* 2013;2:82.

134. Moher D, Shamseer L, Clarke M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Syst Rev.* 2015;4:1.

135. Parko K, Thurman DJ. Prevalence of epilepsy and seizures in the Navajo Nation 1998-2002. *Epilepsia.* 2009;50(10):2180-2185.

136. Pugh MJ, Van Cott AC, Cramer JA, et al. Trends in antiepileptic drug prescribing for older patients with new-onset epilepsy: 2000-2004. *Neurology.* 2008;70(22 Pt 2):2171-2178.

137. Reid AY, St Germaine-Smith C, Liu M, et al. Development and validation of a case definition for epilepsy for use with administrative health data. *Epilepsy Res.* 2012;102(3):173-179.

138. Schneeweiss S, Avorn J. A review of uses of health care utilization databases for epidemiologic research on therapeutics. *J Clin Epidemiol.* 2005;58(4):323-337.

139. Schünemann H, Bro ek J, Oxman A. GRADE handbook for grading quality of evidence and strength of recommendation. 2009;3.2. www.ccims.net/gradepro.

140. Shamseer L, Moher D, Clarke M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. *BMJ.* 2015;349:g7647.

141. Singh S, Chang SM, Matchar DB, Bass EB. Chapter 7: grading a body of evidence on diagnostic tests. *J Gen Intern Med.* 2012;27 Suppl 1:S47-55.

142. St Germaine-Smith C, Metcalfe A, Pringsheim T, et al. Recommendations for optimal ICD codes to study neurologic conditions: a systematic review. *Neurology.* 2012;79(10):1049-1055.

143. Tan M, Wilson I, Braganza V, et al. Development and validation of an epidemiologic case definition of epilepsy for use with routinely collected Australian health data. *Epilepsy Behav.* 2015;51:65-72.

144. Traversa G, Bianchi C, Da Cas R, Abraha I, Menniti-Ippolito F, Venegoni M. Cohort study of hepatotoxicity associated with nimesulide and other non-steroidal anti-inflammatory drugs. *BMJ.* 2003;327(7405):18-22.

145. Tu K, Wang M, Jaakkimainen RL, et al. Assessing the validity of using administrative data to identify patients with epilepsy. *Epilepsia.* 2014;55(2):335-343.

146. Whiting PF, Rutjes AW, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med.* 2011;155(8):529-536.

147. World Health Organization. Classification of Diseases (ICD). 2016. http://www.who.int/classifications/icd/en/. Accessed 13/09/2020.

148. Aaberg KM, Bakken IJ, Lossius MI, et al. Comorbidity and Childhood Epilepsy: A Nationwide Registry Study. *Pediatrics.* 2016;138(3).

149. Aaberg KM, Gunnes N, Bakken IJ, et al. Incidence and Prevalence of Childhood Epilepsy: A Nationwide Cohort Study. *Pediatrics.* 2017;139(5).

150. Bellini I, Policardo L, Zaccara G, et al. Identification of prevalent patients with epilepsy using administrative data: the Tuscany experience. *Neurol Sci.* 2017;38(4):571-577.

151. Chen CC, Chen LS, Yen MF, Chen HH, Liou HH. Geographic variation in the age- and gender-specific prevalence and incidence of epilepsy: Analysis of Taiwanese National Health Insurance-based data. *Epilepsia.* 2012;53(2):283-290.

152. Chubak J, Pocobelli G, Weiss NS. Tradeoffs between accuracy measures for electronic health care data algorithms. *J Clin Epidemiol.* 2012;65(3):343-349 e342.

153. Coulter A, Brown S, Daniels A. Computer held chronic disease registers in general practice: a validation study. *J Epidemiol Community Health.* 1989;43(1):25-28.

154. De Jesus-Alvelo I, Labovitz D. How Reliable Are the ICD9-CM Billing Codes in the Administrative Data To Estimate the Risk of Seizures and Epilepsy after Stroke? *Neurology.* 2013;80.

155. Deeks JJ, Bossuyt PM, Gatsonis C. Handbook for Systematic Reviews of Diagnostic Test Accuracy Version 1.0. 2010. https://methods.cochrane.org/sdt/handbook-dta-reviews. Accessed 19 May 2019.

156. Doebler P, Holling H. Meta-Analysis of Diagnostic Accuracy with mada 2012 https://cran.r-project.org/web/packages/mada/vignettes/mada.pdf. Accessed Accessed 06/01/2019.

157. Engeland A, Bjorge T, Daltveit AK, Vollset SE, Furu K. Validation of disease registration in pregnant women in the Medical Birth Registry of Norway. *Acta Obstet Gynecol Scand.* 2009;88(10):1083-1089.

158. Foebel AD, Hirdes JP, Heckman GA, et al. Diagnostic data for neurological conditions in interRAI assessments in home care, nursing home and mental health care settings: a validity study. *BMC Health Serv Res.* 2013;13:457.

159. Frost FJ, Hurley JS, Petersen HV, Gunter MJ, Gause D. A comparison of two methods for estimating the health care costs of epilepsy. *Epilepsia.* 2000;41(8):1020-1026.

160. Holden EW, Grossman E, Nguyen HT, et al. Developing a computer algorithm to identify epilepsy cases in managed care organizations. *Dis Manag.* 2005;8(1):1-14.

161. Horrocks S, Wilkinson T, Schnier C, et al. Accuracy of routinely-collected healthcare data for identifying motor neurone disease cases: A systematic review. *PLoS One.* 2017;12(2):e0172639.

162. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics.* 1977;33(1):159-174.

163. Lee SY, Chung SE, Kim DW, et al. Estimating the Prevalence of Treated Epilepsy Using Administrative Health Data and Its Validity: ESSENCE Study. *J Clin Neurol.* 2016;12(4):434-440.

164. Marrie RA, Yu BN, Leung S, et al. The utility of administrative data for surveillance of comorbidity in multiple sclerosis: a validation study. *Neuroepidemiology.* 2013;40(2):85-92.

165. Moura LM, Price M, Cole AJ, Hoch DB, Hsu J. Accuracy of claims-based algorithms for epilepsy research: Revealing the unseen performance of claims-based studies. *Epilepsia.* 2017;58(4):683-691.

166. Pickrell WO, Lacey AS, Bodger OG, et al. Epilepsy and deprivation, a data linkage study. *Epilepsia.* 2015;56(4):585-591.

167. Pina-Garza JE, Vekeman F, Cheng W, et al. Development of a claims-based classifier to identify lennox-gastaut syndrome. *Neurology Conference: 67th American Academy of Neurology Annual Meeting, AAN.* 2015;84(SUPPL. 14).

168. Pugh MJ, Parko K. Research using archival health care data: Let the buyer beware. *Epilepsia.* 2015;56(2):321-322.

169. Rehman R, Everhart A, Frontera AT, et al. Implementation of an established algorithm and modifications for the identification of epilepsy patients in the veterans health administration. *Epilepsy Res.* 2016;127:284-290.

170. Reitsma JB, Glas AS, Rutjes AW, Scholten RJ, Bossuyt PM, Zwinderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *J Clin Epidemiol.* 2005;58(10):982-990.

171. Shackleton DP, Westendorp RG, Kasteleijn-Nolst Trenite DG, de Boer A, Herings RM. Dispensing epilepsy medication: a method of determining the frequency of symptomatic individuals with seizures. *J Clin Epidemiol.* 1997;50(9):1061-1068.

172. Syvertsen M, Nakken KO, Edland A, Hansen G, Hellum MK, Koht J. Prevalence and etiology of epilepsy in a Norwegian county-A population based study. *Epilepsia.* 2015;56(5):699-706.

173. Viera AJ, Garrett JM. Understanding interobserver agreement: the kappa statistic. *Fam Med.* 2005;37(5):360-363.

174. Wassenaar M, Carpay JA, Sander JW, Thijs RD. Validity of health insurance data to identify people with epilepsy. *Epilepsy Res.* 2018;139:102-106.

175. Watson N. Using clinical coding systems to best effect in electronic records. *Guidelines in practice.* 2001. https://bit.ly/3tkra8j. Accessed Accessed 09/01/2020.

176. Widdifield J, Ivers NM, Young J, et al. Development and validation of an administrative data algorithm to estimate the disease burden and epidemiology of multiple sclerosis in Ontario, Canada. *Mult Scler.* 2015;21(8):1045-1054.

177. Wilkinson T, Ly A, Schnier C, et al. Identifying dementia cases with routinely collected health data: A systematic review. *Alzheimers Dement.* 2018;14(8):1038-1051.

178. Williamson T, Green ME, Birtwhistle R, et al. Validating the 8 CPCSSN case definitions for chronic disease surveillance in a primary care database of electronic health records. *Ann Fam Med.* 2014;12(4):367-372.

179. Singh S, Fouayzi H, Anzuoni K, et al. Diagnostic Algorithms for Cardiovascular Death in Administrative Claims Databases: A Systematic Review. *Drug Saf.* 2019;42(4):515-527.

180. Albasoft. Albasoft Ltd. http://www.albasoft.co.uk/. Published 2020. Accessed 01/05/2020.

181. Death Certification Review Service. Annual Report 2015–2016 2016. HTTPS://goo.gl/sBDSoU. Accessed 25/06/2020.

182. Donner EJ, Devinsky O. Registries for SUDEP research. In: Panelli R, Hanna J, Jeffs T, Brocett P, eds. *Continuing the global conversation [online].* SUDEP Action & SUDEP Aware; 2018.

183. Du Z, Hao y. Package 'reportROC': An Easy Way to Report ROC Analysis. *CRAN.* 2019. https://bit.ly/2V3X6OA. Accessed 08/04/2020.

184. eDRIS Team. Researcher guide: requesting outputs from Safe Haven and disclosure control 2018. https://bit.ly/2XzD1Bk. Accessed 16/11/2019.

185. England MJ, Liverman CT, Schultz AM, Strawbridge LM. IOM Report 2012: Epilepsy Across the Spectrum: Promoting Health and Understanding Abstracts. *Epilepsy Curr.* 2013;13(1):2-492.

186. Fisher RS, van Emde Boas W, Blume W, et al. Epileptic seizures and epilepsy: definitions proposed by the International League Against Epilepsy (ILAE) and the International Bureau for Epilepsy (IBE). *Epilepsia.* 2005;46(4):470-472.

187. Hamer HM, Dodel R, Strzelczyk A, et al. Prevalence, utilization, and costs of antiepileptic drugs for epilepsy in Germany--a nationwide population-based study in children and adults. *J Neurol.* 2012;259(11):2376-2384.

188. Herrett E, Gallagher AM, Bhaskaran K, et al. Data Resource Profile: Clinical Practice Research Datalink (CPRD). *Int J Epidemiol.* 2015;44(3):827-836.

189. ILAE. Guidelines for Epidemiologic Studies on Epilepsy. *Epilepsia.* 1993;34(4):592-596.

190. Information Services Division. Prescribing Information System (PIS). *National Data Catalogue.* 2016. https://bit.ly/3Ejcj2N. Accessed 27/10/2019.

191. Information Services Division. General Practice - SPIRE 2020. https://bit.ly/3eldK3B. Accessed 03/03/2020.

192. Kerr MP. The impact of epilepsy on patients' lives. *Acta Neurol Scand Suppl.* 2012(194):1-9.

193. Mbizvo GK, Schnier C, Simpson CR, Chin RFM, Duncan SE. The Scottish Epilepsy Deaths Study (Seds): Identifying Avoidable Epilepsy-Related Deaths. *J Neurol Neurosur Ps.* 2019;90(12):E16-E17.

194. Middleton O, Atherton D, Bundock E, et al. National Association of Medical Examiners position paper: Recommendations for the investigation and certification of deaths in people with epilepsy. *Epilepsia.* 2018;59(3):530-543.

195. National Records of Scotland. The Medical Certificate of the Cause of Death 2020. https://bit.ly/3aPkqnZ. Accessed 03/03/2020.

196. National Records Scoltand. Coding the causes of death. *Death Certificates and Coding the Causes of Death.* 2016. https://goo.gl/oWTxQU. Accessed 03/03/2020.

197. O'Malley KJ, Cook KF, Price MD, Wildes KR, Hurdle JF, Ashton CM. Measuring diagnoses: ICD code accuracy. *Health Serv Res.* 2005;40(5):1620-1639.

198. Pickrell WO, Kerr MP. SUDEP and mortality in epilepsy: The role of routinely collected healthcare data, registries, and health inequalities. *Epilepsy Behav.* 2020;103(Pt B):106453.

199. Public Health Scotland. Data and intelligence - Previously ISD Scotland. Public Health Scotland. https://beta.isdscotland.org. Published 2020. Accessed 29/04/2020.

200. *RStudio: Integrated Development for R* [computer program]. Boston, MA RStudio, Inc.; 2015

201. Thiele C. cutpointr: Determine and Evaluate Optimal Cutpoints in Binary Classification Tasks 2020. https://rdrr.io/cran/cutpointr/man/roc.html. Accessed 01/05/2020.

202. Thomas RH, Osland K. Learnings from deaths - the Epilepsy Deaths Register. *Epilepsy Behav.* 2020;103(Pt B):106454.

203. Verducci C, Hussain F, Donner E, et al. SUDEP in the North American SUDEP Registry: The full spectrum of epilepsies. *Neurology.* 2019;93(3):e227-e236.

204. Mbizvo GK, Schnier C, Simpson CR, Chin RFM, Duncan SE. A national study of epilepsy-related deaths in Scotland: Trends, mechanisms, and avoidable deaths. *Epilepsia.* 2021;62(11):2667-2684.

205. Morton B, Richardson A, Duncan S. Sudden unexpected death in epilepsy (SUDEP): don't ask, don't tell? *J Neurol Neurosurg Psychiatry.* 2006;77(2):199-202.

206. Keddie S, Angus-Leppan H, Parker T, et al. Discussing sudden unexpected death in epilepsy: Are we empowering our patients? A questionnaire survey. *JRSM Open.* 2016;7(9):2054270416654358.

207. Nolte E, McKee M. *Does Healthcare Save Lives? Avoidable Mortality Revisited.* London: The Nuffield Trust; 2004.

208. Page A, Tobias M, Glover J, Wright C, Hetzel D, Fisher E. *Australian and New Zealand Atlas of Avoidable Mortality.* Adelaide: PHIDU: University of Adelaide; 2006

209. ONS Mortality Team. Avoidable mortality in the UK QMI 2020. https://bit.ly/2SMgPmY. Accessed 31/05/2021.

210. Crown Office and Procurator Fiscal Service. Reporting deaths to the Procurator Fiscal - Information and Guidance for Medical Practitioners 2015. https://bit.ly/3uUUWPZ. Accessed 06/05/2021.

211. The Royal College of Pathologists. Guidelines on autopsy practice 2002. http://www.ihrdni.org/314-008-1.pdf. Accessed 01/06/2021.

212. Mbizvo GK, Schnier C, Simpson CR, Duncan SE, Chin RFM. Validating the accuracy of administrative healthcare data identifying epilepsy in deceased adults: A Scottish data linkage study. *Epilepsy Res.* 2020;167:106462.

213. National records of Scotland. Death Certificates and Coding the Causes of Death 2017. https://bit.ly/2DrrjR2. Accessed 31/07/2020.

214. Information Services Division. The Scottish Index of Multiple Deprivation (SIMD). *GPD Support - Deprivation.* 2016. https://bit.ly/2V32ilM. Accessed 27/10/2019.

215. Office for National Statistics. *Revised European Standard Population.* 2013.

216. Office for National Statistics. User guide to mortality statistics. *Deaths.* 2019. https://bit.ly/2WydW8I. Accessed 30/10/19.

217. National Records of Scotland. Mid-Year Population Estimates. *Population Estimates.* 2018. https://bit.ly/3hld8vP. Accessed 30/10/2019.

218. Office for National Statistics. Calculation of Standardised Mortality Ratios (SMRs) 2019. https://bit.ly/2WB0j8q. Accessed 30/10/2019.

219. National Records of Scotland. Deaths Time Series Data *Deaths* 2019. https://bit.ly/30mdgUU. Accessed 30/10/2019.

220. Information Services Division. Epilepsy 2013. https://bit.ly/2yXP4y8. Accessed 30/10/2019.

221. World Health Organisation. International Statistical Classification of Diseases and Related Health Problems 10th Revision 1990. https://icd.who.int/browse10/2010/en. Accessed 08/09/2020.

222. World Health Organization. Global burden of epilepsy and the need for coordinated action at the country level to address its health, social and public knowledge implications https://bit.ly/3r66RbK.

223. Bradley PM, Lindsay B, Fleeman N. Care delivery and self management strategies for adults with epilepsy. *Cochrane Database Syst Rev.* 2016;2:CD006244.

224. Steer S, Pickrell WO, Kerr MP, Thomas RH. Epilepsy prevalence and socioeconomic deprivation in England. *Epilepsia.* 2014;55(10):1634-1641.

225. National Records of Scotland. Table 9: All ages age-standardised death rates by SIMD quintile and selected cause of death, Scotland, 2001 to 2019. *Vital Events.* 2020. https://bit.ly/2R4dn6Q. Accessed 30/05/2021.

226. Tonberg A, Harden J, McLellan A, Chin RF, Duncan S. A qualitative study of the reactions of young adults with epilepsy to SUDEP disclosure, perceptions of risks, views on the timing of disclosure, and behavioural change. *Epilepsy Behav.* 2015;42:98-106.

227. Sen A, Jette N, Husain M, Sander JW. Epilepsy in older people. *Lancet.* 2020;395(10225):735-748.

228. The Scottish Public Health Observatory (ScotPHO). Deaths: most frequent causes 2019. https://bit.ly/3gb9x4y. Accessed 06/06/2021.

229. Grant SW, Collins GS, Nashef SAM. Statistical Primer: developing and validating a risk prediction model. *Eur J Cardiothorac Surg.* 2018;54(2):203-208.

230. Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *BMJ.* 2015;350:g7594.

231. McKinstry B, Sullivan FM, Vasishta S, et al. Cohort profile: the Scottish Research register SHARE. A register of people interested in research participation linked to NHS data sets. *BMJ Open.* 2017;7(2):e013351.

232. Sullivan F, McKinstry B, Vasishta S. The "All of Us" Research Program. *N Engl J Med.* 2019;381(19):1883-1884.

233. Scottish Parliament Epilepsy Consortium Scotland. SUDEP Rates remain high in Scotland. *Cross Party Group Meeting.* 2019. http://www.epilepsyconsortiumscotland.co.uk/sudep-rates-scotland/. Accessed 21/09/2020.

234. Fyfe I. Ean Virtual 2020 - the Largest Neurology Conference in History. *Nat Rev Neurol.* 2020;16(8):404.

235. ERUK. ERUK Researcher Wins Award At ABN National Conference 2019. https://bit.ly/3bDanWa. Accessed 19/05/2021.

236. Bujang MA, Sa'at N, Sidik T, Joo LC. Sample Size Guidelines for Logistic Regression from Observational Studies with Large Population: Emphasis on the Accuracy Between Statistics and Parameters Based on Real Life Clinical Data. *Malays J Med Sci.* 2018;25(4):122-130.

237. Peduzzi P, Concato J, Feinstein AR, Holford TR. Importance of events per independent variable in proportional hazards regression analysis. II. Accuracy and precision of regression estimates. *J Clin Epidemiol.* 1995;48(12):1503-1510.

238. Chowdhury MZI, Turin TC. Variable selection strategies and its importance in clinical prediction modelling. *Fam Med Community Health.* 2020;8(1):e000262.

239. McNeish D. Missing data methods for arbitrary missingness with small samples. *J Appl Stat.* 2017;44(1):24-39.

240. Bennett DA. How can I deal with missing data in my study? *Aust Nz J Publ Heal.* 2001;25(5):464-469.

241. IBM. Method (Multiple Imputation). *SPSS Statistics.* 2021. https://ibm.co/3wjcMNw. Accessed 19/05/2021.

242. Rubin DB. *Multiple Imputation for Nonresponse in Surveys.* New York: Wiley; 1987.

243. Dong YR, Peng CYJ. Principled missing data methods for researchers. *Springerplus.* 2013;2.

244. Goodman LA, Kruskal WH. *Measures of Association for Cross Classifications.* New York: Springer-Verlag; 1979.

245. Noe MH, Rosenbach M, Hubbard RA, et al. Development and Validation of a Risk Prediction Model for In-Hospital Mortality Among Patients With Stevens-Johnson Syndrome/Toxic Epidermal Necrolysis-ABCD-10. *JAMA Dermatol.* 2019;155(4):448-454.

246. Dankers F, Traverso A, Wee L, van Kuijk SMJ. Prediction Modeling Methodology. In: Kubben P, Dumontier M, Dekker A, eds. *Fundamentals of Clinical Data Science.* Cham (CH)2019:101-120.

247. Muschelli J. ROC and AUC with a Binary Predictor: a Potentially Misleading Metric. *J Classif.* 2019.

248. Scheffer IE, Berkovic S, Capovilla G, et al. ILAE classification of the epilepsies: Position paper of the ILAE Commission for Classification and Terminology. *Epilepsia.* 2017;58(4):512-521.

249. Lyons RA, Jones KH, John G, et al. The SAIL databank: linking multiple health and social care datasets. *BMC Med Inform Decis Mak.* 2009;9:3.

250. Kappen TH, van Klei WA, van Wolfswinkel L, Kalkman CJ, Vergouwe Y, Moons KGM. Evaluating the impact of prediction models: lessons learned, challenges, and recommendations. *Diagn Progn Res.* 2018;2:11.

251. Schaefer JT. The critical success index as an indicator of warning skill. *Weather Forecast* 1990;5:570-575.

252. Palmer WC, Allen RA. *Note on the accuracy of forecasts concerning the rain problem.* Washington, DC: U.S. Weather Bureau manuscript; 1949.

253. Gilbert GK. Finley's tornado predictions. *Am Meteorol J* 1884;1:166-172.

254. Larner AJ. Assessing cognitive screening instruments with the critical success index. *Prog Neurol Psychiatry* 2021;25(3):in press

255. Chae S, Shin J, Kwon S, Lee S, Kang S, Lee D. PM10 and PM2.5 real-time prediction models using an interpolated convolutional neural network. *Sci Rep.* 2021;11(1):11952.

256. Kim RS, Moon YJ, Gopalswamy N, Park YD, Kim YH. Two-step forecast of geomagnetic storm using coronal mass ejection and solar wind condition. *Space Weather.* 2014;12(4):246-256.

257. Pavlovic R, Chen J, Anderson K, et al. The FireWork air quality forecast system with near-real-time biomass burning emissions: Recent developments and evaluation of performance for the 2015 North American wildfire season. *J Air Waste Manage.* 2016;66(9):819-841.

258. El Jarroudi M, Kouadio L, Delfosse P, Tychon B. Brown rust disease control in winter wheat: I. Exploring an approach for disease progression based on night weather conditions. *Environ Sci Pollut R.* 2014;21(7):4797-4808.

259. Shin D, Kim JH. A New Application of Unsupervised Learning to Nighttime Sea Fog Detection. *Asia-Pac J Atmos Sci.* 2018;54(4):527-544.

260. Skinner PS, Wheatley DM, Knopfmeier KH, et al. Object-Based Verification of a Prototype Warn-on-Forecast System. *Weather and Forecasting.* 2018;33(5):1225-1250.

261. Nguyen P, Ombadi M, Gorooh VA, et al. PERSIANN Dynamic Infrared-Rain Rate (PDIR-Now): A Near-Real-Time, Quasi-Global Satellite Precipitation Dataset. *J Hydrometeorol.* 2020;21(12):2893-2906.

262. Jing JR, Li Q, Peng X. MLC-LSTM: Exploiting the Spatiotemporal Correlation between Multi-Level Weather Radar Echoes for Echo Sequence Extrapolation. *Sensors-Basel.* 2019;19(18).

263. Li GF, Chen Y, Le TM, Wang WY, Tang XY, Li CSR. Neural correlates of individual variation in two-back working memory and the relationship with fluid intelligence. *Sci Rep-Uk.* 2021;11(1).

264. Fawcett T. An introduction to ROC analysis. *Pattern Recogn Lett.* 2006;27(8):861-874.

265. Mbizvo GK, Larner AJ. Receiver operating characteristic plot and area under the curve with binary classifiers: pragmatic analysis of cognitive screening instruments. *Neurodegener Dis Manag.* 2021;11(5):353-360.

266. Sattler EL, Lee JS, Perri M, 3rd. Medication (re)fill adherence measures derived from pharmacy claims data in older Americans: a review of the literature. *Drugs Aging.* 2013;30(6):383-399.

267. Askarieh A, Morrison I, Ross K, et al. Medication adherence, utilization of healthcare services, and mortality of patients with epilepsy on opiate replacement therapy: A retrospective cohort study. *Epilepsy Behav.* 2021;117:107829.

Appendix 1 – Supplementary material to Chapter 2

Appendix A: Search strategy

MEDLINE (PubMed interface) 1946 to 26 December 2018	***Embase Classic+Embase (Ovid interface) 26 December 2018***
(("Epilepsy"[Mesh] OR EPILEPSY[Text Word]) AND ("Mortality"[Mesh] OR "mortality"[Subheading] OR MORTALITY[Text Word] OR "Death"[Mesh])) AND "systematic review"[All Fields]	1. frontal lobe epilepsy/ or "seizure, epilepsy and convulsion"/ or childhood absence epilepsy/ or grand mal epilepsy/ or myoclonic astatic epilepsy/ or focal epilepsy/ or photosensitive epilepsy/ or epilepsy/ or startle epilepsy/ or mesial temporal lobe epilepsy/ or temporal lobe epilepsy/ or catamenial epilepsy/ or myoclonus epilepsy/ or generalized epilepsy/ or lateral temporal lobe epilepsy/ or rolandic epilepsy/ or intractable epilepsy/ or drug resistant epilepsy/ or traumatic epilepsy/ or symptomatic epilepsy/ or reflex epilepsy/ or benign childhood epilepsy/ 2. mortality/ or mortality rate/ or mortality risk/ or standardized mortality ratio/ or hospital mortality/ or mortality risk.mp. 3. "cause of death"/ or death/ or accidental death/ 4. systematic review.mp. or "systematic review"/ 5. meta analysis/ 6. survival analysis/ 7. heterogen*.mp. 8. 2 or 3 9. 4 or 5 or 6 or 7 10. 1 and 8 and 9 11. animal/ not human/ 12. 10 not 11 13. limit 12 to human

Electronic database search strategies

Appendix B: data abstraction pro forma

The abstracted data items were as follows:

A. Systematic review title;

B. Author(s);

C. Journal;

D. Year of publication;

E. Details of funding and any conflicts of interest statements;

F. Research question(s);

G. Inclusion and exclusion criteria;

H. Whether the research question and inclusion criteria were established before conduct of the review;

I. Search strategy, databases and years searched, and whether supplemented by references list searches or contacting of study authors/experts in the field;

J. Whether a list of included and excluded studies was provided, and reasons for exclusion (including whether any reports were excluded because of publication type or language);

K. Whether duplicate study selection was used and details of any consensus procedures;

L. Study design;

M. Populations covered;

N. Interventions/observations;

O. Comparator groups;

P. Whether duplicate data abstraction was performed and details of any consensus procedures;

Q. Risk of bias and study quality assessment methods used;

R. Methods of data synthesis/meta-analysis (including details of significance tests, methods for investigating heterogeneity, and how heterogeneity was dealt with);

S. Whether there was a registered review protocol;

T. Outcomes and outcome measures (including measures of spread such as confidence intervals/standard errors, and p-values) when these were any of standardized mortality ratio, mortality rate, relative risk, hazard ratio, case fatality, odds ratio, proportionate mortality ratio, and rate ratio;

U. Main discussion points and overall conclusions, including whether or not scientific quality of the included studies was taken into consideration when making these;

V. Any review limitations described by the authors.

Appendix C: List of excluded records

Systematic reviews of populations which may not necessarily have epilepsy

Chin, R. F. M., B. G. R. Neville and R. C. Scott (2004). "A systematic review of the epidemiology of status epilepticus." European Journal of Neurology 11(12): 800-810.

Chungath, M. and S. Shorvon (2008). "The mortality and morbidity of febrile seizures." Nature Clinical Practice Neurology 4(11): 610-621.

Hancock, E. C., J. P. Osborne and S. W. Edwards (2013). "Treatment of infantile spasms." Cochrane Database Syst Rev(6): CD001770.

Lv, R. J., Q. Wang, T. Cui, F. Zhu and X. Q. Shao (2017). "Status epilepticus-related etiology, incidence and mortality: A meta-analysis." Epilepsy Res 136: 12-17.

Neligan, A. and S. D. Shorvon (2011). "Prognostic factors, morbidity and mortality in tonic-clonic status epilepticus: A review." Epilepsy Research 93(1): 1-10.

Radaelli, G., F. de Souza Santos, W. V. Borelli, L. Pisani, M. L. Nunes, F. A. Scorza and J. C. da Costa (2018). "Causes of mortality in early infantile epileptic encephalopathy: A systematic review." Epilepsy Behav **85**: 32-36.

Raspall-Chaure, M., R. F. M. Chin, B. G. Neville and R. C. Scott (2006). "Outcome of paediatric convulsive status epilepticus: a systematic review." Lancet Neurology 5(9): 769-779.

Shneker, B. F. and N. B. Fountain (2003). "Assessment of acute morbidity and mortality in nonconvulsive status epilepticus." Neurology 61(8): 1066-U1063.

Not systematic reviews

Diop, A. G., D. C. Hesdorffer, G. Logroscino and W. A. Hauser (2005). "Epilepsy and mortality in Africa: a review of the literature." Epilepsia 46 Suppl 11: 33-35.

Forsgren, L., W. A. Hauser, E. Olafsson, J. W. A. S. Sander, M. Sillanpaa and T. Tomson (2005). "Mortality of epilepsy in developed countries: A review." Epilepsia 46: 18-27.

Guekht, A. (2013). "Epilepsy, AEDs and suicide." Epilepsy & Behavior 28(2): 306.

Hampel, K. G., R. Rocamora Zuniga and C. M. Quesada (2017). "Unravelling the mysteries of sudden unexpected death in epilepsy." Neurologia.

Keezer, M. R., G. S. Bell and J. W. Sander (2015). "Epilepsy-related clinical characteristics and mortality: a systematic review and meta-analysis." Neurology 84(17): 1823.

Nevalainen, O. P., H. Ansakorpi and A. Auvinen (2015). "Epilepsy-related clinical characteristics and mortality: a systematic review and meta-analysis." Neurology 84(17): 1823-1824.

Sethi, N. K. (2015). "Epilepsy-related clinical characteristics and mortality: a systematic review and meta-analysis." Neurology 84(17): 1823.

Did not assess risk factors for the epilepsy-related death reviewed

Pompili, M., P. Girardi, A. Ruberto and R. Tatarelli (2005). "Suicide in the epilepsies: a meta-analytic investigation of 29 cohorts." Epilepsy Behav 7(2): 305-310.

Saetre, E. and M. Abdelnoor (2018). "Incidence rate of sudden death in epilepsy: A systematic review and meta-analysis." Epilepsy Behav **86**: 193-199.

Thurman, D. J., D. C. Hesdorffer and J. A. French (2014). "Sudden unexpected death in epilepsy: assessing the public health burden." Epilepsia 55(10): 1479-1485.

Highly selected subpopulations

Robertson, J., C. Hatton, E. Emerson and S. Baines (2015). "Mortality in people with intellectual disabilities and epilepsy: A systematic review." Seizure 29: 123-133.

Tellez-Zenteno, J. F., R. Dhar, L. Hernandez-Ronquillo and S. Wiebe (2007). "Long-term outcomes in epilepsy surgery: antiepileptic drugs, mortality, cognitive and psychosocial aspects." Brain 130: 334-345.

Previous abstract to an included systematic review

Abdel-Mannan OA, Taylor H and Sutcliffe AG (2017). "A systematic review of Sudden Unexpected Death in Epilepsy (SUDEP) in childhood: What are the current gaps in knowledge?" Developmental Medicine and Child Neurology. Conference: 43rd Annual Conference of the British Paediatric Neurology Association, BPNA 2017.

Watila, M. M. and S. A. Balarabe (2015). "Cause-Specific Mortality in Epilepsy. A Systematic Review." Epilepsia 56: 64-64.

No new studies identified

Neligan, A., G. S. Bell, S. D. Shorvon and J. W. Sander (2010). "Temporal trends in the mortality of people with epilepsy: a review." Epilepsia 51(11): 2241-2246.

Appendix D: List of the primary studies contributing evidence

Studies that did not overlap across systematic reviews (n = 50 studies)

Almeida, A. G., M. L. Nunes, A. L. Palmini and J. C. Costa (2010). "Incidence of SUDEP in a cohort of patients with refractory epilepsy: the role of surgery and lesion localization." Arq Neuropsiquiatr 68(6): 898-902.

Alstrom, C. H. (1942). Mortality in mental hospitals with special regard to tuberculosis. Uppsala, Appelsbergs Boktryckeriaktieborg.

Annegers, J. F., S. P. Coan, W. A. Hauser, J. Leestma, W. Duffell and B. Tarver (1998). "Epilepsy, vagal nerve stimulation by the NCP system, mortality, and sudden, unexpected, unexplained death." Epilepsia 39(2): 206-212.

Bell, G. S., A. Neligan, C. Giavasi, M. R. Keezer, J. Novy, J. L. Peacock, A. L. Johnson, D. M. Goodridge, S. D. Shorvon and J. W. Sander (2016). "Outcome of seizures in the general population after 25 years: a prospective follow-up, observational cohort study." J Neurol Neurosurg Psychiatry 87(8): 843-850.

Benn, E. K., W. A. Hauser, T. Shih, L. Leary, E. Bagiella, P. Dayan, R. Green, H. Andrews, D. J. Thurman and D. C. Hesdorffer (2008). "Estimating the incidence of first unprovoked seizure and newly diagnosed epilepsy in the low-income urban community of Northern Manhattan, New York City." Epilepsia 49(8): 1431-1439.

Benn, E. K., W. A. Hauser, T. Shih, L. Leary, E. Bagiella, P. Dayan, R. Green, H. Andrews, D. J. Thurman and D. C. Hesdorffer (2009). "Underlying cause of death in incident unprovoked seizures in the urban community of Northern Manhattan, New York City." Epilepsia 50(10): 2296-2300.

Berg, A. T., K. Nickels, E. C. Wirrell, A. T. Geerts, P. M. Callenbach, W. F. Arts, C. Rios, P. R. Camfield and C. S. Camfield (2013). "Mortality risks in new-onset childhood epilepsy." Pediatrics 132(1): 124-131.

Callaghan, B., H. Choi, M. Schlesinger, W. Rodemer, J. Pollard, D. C. Hesdorffer, W. A. Hauser and J. French (2014). "Increased mortality persists in an adult drug-resistant epilepsy prevalence cohort." J Neurol Neurosurg Psychiatry 85(10): 1084-1090.

Chamorro-Munoz, M. I., G. Garcia-Martin, F. Perez-Errazquin, M. Romero-Acebal, A. Garcia-Rodriguez and M. Gutierrez-Bedmar (2017). "Epidemiological study of mortality in epilepsy in a Spanish population." Seizure 46: 19-23.

Chen, Z., D. Liew and P. Kwan (2016). "Excess mortality and hospitalized morbidity in newly treated epilepsy patients." Neurology 87(7): 718-725.

Christensen, J., C. B. Pedersen, P. Sidenius, J. Olsen and M. Vestergaard (2015). "Long-term mortality in children and young adults with epilepsy--A population-based cohort study." Epilepsy Res 114: 81-88.

Day, S. M., Y. W. Wu, D. J. Strauss, R. M. Shavelle and R. J. Reynolds (2005). "Causes of death in remote symptomatic epilepsy." Neurology 65(2): 216-222.

Derby, L. E., P. Tennis and H. Jick (1996). "Sudden unexplained death among subjects with refractory epilepsy." Epilepsia 37(10): 931-935.

Diekema, D. S., L. Quan and V. L. Holt (1993). "Epilepsy as a risk factor for submersion injury in children." Pediatrics 91(3): 612-616.

Donner, E. J., C. R. Smith and O. C. Snead, 3rd (2001). "Sudden unexplained death in children with epilepsy." Neurology 57(3): 430-434.

Erilä T (1982). "Mortality in patients with epilepsy in Finland 1967–1973." Acta Universitatis Tamperensis Ser A 145: 1–167.

Ficker, D. M., E. L. So, W. K. Shen, J. F. Annegers, P. C. O'Brien, G. D. Cascino and P. G. Belau (1998). "Population-based study of the incidence of sudden unexplained death in epilepsy." Neurology 51(5): 1270-1274.

Forsgren, L., S. O. Edvinsson, L. Nystrom and H. K. Blomquist (1996). "Influence of epilepsy on mortality in mental retardation: an epidemiologic study." Epilepsia 37(10): 956-963.

Gaitatzis, A., A. L. Johnson, D. W. Chadwick, S. D. Shorvon and J. W. Sander (2004). "Life expectancy in people with newly diagnosed epilepsy." Brain 127(Pt 11): 2427-2432.

Granbichler, C. A., W. Oberaigner, G. Kuchukhidze, J. P. Ndayisaba, A. Ndayisaba, A. Taylor, G. Bauer, G. Luef and E. Trinka (2017). "Decrease in mortality of adult epilepsy patients since 1980: lessons learned from a hospital-based cohort." Eur J Neurol 24(5): 667-672.

Harvey, A. S., T. Nolan and J. B. Carlin (1993). "Community-based study of mortality in children with epilepsy." Epilepsia 34(4): 597-603.

Houinato, D., L. P. Yemadje, G. Glitho, C. Adjien, G. Avode, M. Druet-Cabanac and P. M. Preux (2013). "Epidemiology of epilepsy in rural Benin: prevalence, incidence, mortality, and follow-up." Epilepsia 54(4): 757-763.

Hutchinson JJ (1976). New York Life single medical impairment study: 1972, unpublished report. Medical Risks: Patterns of Mortality and Survival, Part II: tabular abstracts 2. Singer RD and Levinson. Toronto, Lexington Books: 32-49.

Jilek-Aall, L. and H. T. Rwiza (1992). "Prognosis of epilepsy in a rural African community: a 30-year follow-up of 164 patients in an outpatient clinic in rural Tanzania." Epilepsia 33(4): 645-650.

Kaiboriboon, K., N. K. Schiltz, P. M. Bakaki, S. D. Lhatoo and S. M. Koroukian (2014). "Premature mortality in poor health and low income adults with epilepsy." Epilepsia 55(11): 1781-1788.

Keezer, M. R., G. S. Bell, A. Neligan, J. Novy and J. W. Sander (2016). "Cause of death and predictors of mortality in a community-based cohort of people with epilepsy." Neurology 86(8): 704-712.

Lip, G. Y. and M. J. Brodie (1992). "Sudden death in epilepsy: an avoidable outcome?" J R Soc Med 85(10): 609-611.

Loiseau, P., J. Loiseau and M. C. Picot (2005). "One-year mortality in Bordeaux cohort: the value of syndrome classification." Epilepsia 46 Suppl 11: 11-14.

Luhdorf, K., L. K. Jensen and A. M. Plesner (1987). "Epilepsy in the elderly: life expectancy and causes of death." Acta Neurol Scand 76(3): 183-190.

Malzberg B (1934). Mortality among patients with mental disease. Utica, NY, State Hospital Press.

Nashef, L., D. R. Fish, S. Garner, J. W. Sander and S. D. Shorvon (1995). "Sudden death in epilepsy: a study of incidence in a young cohort with epilepsy and learning difficulty." Epilepsia 36(12): 1187-1194.

Nilsson, L., A. Ahlbom, B. Y. Farahmand and T. Tomson (2003). "Mortality in a population-based cohort of epilepsy surgery patients." Epilepsia 44(4): 575-581.

O'Donoghue, M. F. and J. W. Sander (1997). "A historical perspective on the mortality associated with chronic epilepsy." Acta Neurol Scand 96(3): 138-141.

Opeskin, K. and S. F. Berkovic (2003). "Risk factors for sudden unexpected death in epilepsy: a controlled prospective study based on coroners cases." Seizure 12(7): 456-464.

Preston, T. and R. Clarke (1966). "An investigation into the mortality of impaired lives during the period 1947-63." Journal of the Institute of Actuaries 92(1): 27-74.

Ridsdale, L., J. Charlton, M. Ashworth, M. P. Richardson and M. C. Gulliford (2011). "Epilepsy mortality and risk factors for death in epilepsy: a population-based study." Br J Gen Pract 61(586): e271-278.

Sillanpaa, M., M. Jalava, O. Kaleva and S. Shinnar (1998). "Long-term prognosis of seizures with onset in childhood." N Engl J Med 338(24): 1715-1722.

Singh, S., S. M. Chang, D. B. Matchar and E. B. Bass (2012). "Chapter 7: grading a body of evidence on diagnostic tests." J Gen Intern Med 27 Suppl 1: S47-55.

Sperling, M. R., A. Harris, M. Nei, J. D. Liporace and M. J. O'Connor (2005). "Mortality after epilepsy surgery." Epilepsia 46 Suppl 11: 49-53.

Svensson A and Astrand S (1976). Substandardized risk mortality in Sweden. 1955-1965. Medical Risks: Patterns of Mortality and Survival, Part II: tabular abstracts Singer RD and Levinson. Toronto, Lexington Books. 2: 32-49.

Tennis, P., T. B. Cole, J. F. Annegers, J. E. Leestma, M. McNutt and A. Rajput (1995). "Cohort study of incidence of sudden unexplained death in persons with seizure disorder treated with antiepileptic drugs in Saskatchewan, Canada." Epilepsia 36(1): 29-36.

Terra, V. C., F. A. Scorza, E. A. Cavalheiro, L. Wichert-Ana, K. G. Pinto, H. R. Machado and A. C. Sakamoto (2010). "Pediatric epilepsy surgery and sudden unexpected death epilepsy: the contribution of a Brazilian epilepsy surgery program." Childs Nerv Syst 26(8): 1075-1079.

Terra, V. C., F. A. Scorza, A. C. Sakamoto, K. G. Pinto, R. M. Fernandes, R. M. Arida, E. A. Cavalheiro and H. R. Machado (2009). "Does sudden unexpected death in children with epilepsy occur more frequently in those with high seizure frequency?" Arq Neuropsiquiatr 67(4): 1001-1002.

Thomas, S. V., B. Reghunath and P. Sankara Sarma (2001). "Mortality among epilepsy patients attending a tertiary referral center in a developing country." Seizure 10(5): 370-373.

Timmings, P. L. (1998). "Sudden unexpected death in epilepsy: is carbamazepine implicated?" Seizure 7(4): 289-291.

Vickrey, B. G. (1997). "Mortality in a consecutive cohort of 248 adolescents and adults who underwent diagnostic evaluation for epilepsy surgery." Epilepsia 38(11 Suppl): S67-69.

Voute PA (1967). Epilepsy and life insurance: social studies on epilepsy. No. 5. London, International Bureau for Epilepsy.

Wagner, R. G., C. Bottomley, A. K. Ngugi, F. Ibinda, F. X. Gomez-Olive, K. Kahn, S. Tollman, C. R. Newton, S. W. Group, R. Wagner, R. Twine, M. Connor, M. Collinson, H. Masanja, A. Mathew, A. Kakooza, G. Pariyo, S. Peterson, D. Ndyo-mughenyi, R. Odhiambo, E. Chengo, M. Chabi, E. Bauni, G. Kamuyu, V. M. Odera, J. O. Mageto, K. Ae-Ngibise, B. Akpalu, A. Akpalu, F. Agbokey, P. Adjei, S. Owusu-Agyei, I. Kleinschmidt, V. C. Doku, P. Odermatt, B. Neville, J. W. Sander, S. White, T. Nutman, P. Wilkins and J. Noh (2015). "Incidence, Remission and Mortality of Convulsive Epilepsy in Rural Northeast South Africa." PLoS One 10(6): e0129097.

Walczak, T. S., I. E. Leppik, M. D'Amelio, J. Rarick, E. So, P. Ahman, K. Ruggles, G. D. Cascino, J. F. Annegers and W. A. Hauser (2001). "Incidence and risk factors in sudden unexpected death in epilepsy: a prospective cohort study." Neurology 56(4): 519-525.

Weber, P., R. Bubl, U. Blauenstein, B. U. Tillmann and J. Lutschg (2005). "Sudden unexplained death in children with epilepsy: a cohort study with an eighteen-year follow-up." Acta Paediatr 94(5): 564-567.

Studies overlapping two systematic reviews (n = 25 studies)

Alstrom, C. H. (1950). "A study of epilepsy in its clinical, social and genetic aspects." Acta Psychiatr Neurol Suppl 63: 1-284.

Annegers, J. F., S. P. Coan, W. A. Hauser and J. Leestma (2000). "Epilepsy, vagal nerve stimulation by the NCP system, all-cause mortality, and sudden, unexpected, unexplained death." Epilepsia 41(5): 549-553.

Annegers, J. F., W. A. Hauser and S. B. Shirts (1984). "Heart disease mortality and morbidity in patients with epilepsy." Epilepsia 25(6): 699-704.

Callenbach, P. M., R. G. Westendorp, A. T. Geerts, W. F. Arts, E. A. Peeters, C. A. van Donselaar, A. C. Peters, H. Stroink and O. F. Brouwer (2001). "Mortality risk in children with epilepsy: the Dutch study of epilepsy in childhood." Pediatrics 107(6): 1259-1263.

Chang, Y. H., W. C. Ho, J. J. Tsai, C. Y. Li and T. H. Lu (2012). "Risk of mortality among patients with epilepsy in southern Taiwan." Seizure-European Journal of Epilepsy 21(4): 254-259.

Chen, R. C., Y. C. Chang, T. H. Chen, H. M. Wu and H. H. Liou (2005). "Mortality in adult patients with epilepsy in Taiwan." Epileptic Disord 7(3): 213-219.

Chin, R. F., P. M. Cumberland, S. S. Pujar, C. Peckham, E. M. Ross and R. C. Scott (2011). "Outcomes of childhood epilepsy at age 33 years: a population-based birth-cohort study." Epilepsia 52(8): 1513-1521.

Cockerell, O. C., A. L. Johnson, J. W. Sander, Y. M. Hart, D. M. Goodridge and S. D. Shorvon (1994). "Mortality from epilepsy: results from a prospective population-based study." Lancet 344(8927): 918-921.

Cockerell, O. C., A. L. Johnson, J. W. Sander and S. D. Shorvon (1997). "Prognosis of epilepsy: a review and further analysis of the first nine years of the British National General Practice Study of Epilepsy, a prospective population-based study." Epilepsia 38(1): 31-46.

Ding, D., W. Wang, J. Wu, G. Ma, X. Dai, B. Yang, T. Wang, C. Yuan, Z. Hong, H. M. de Boer, L. Prilipko and J. W. Sander (2006). "Premature mortality in people with epilepsy in rural China: a prospective study." Lancet Neurol 5(10): 823-827.

Fazel, S., A. Wolf, N. Langstrom, C. R. Newton and P. Lichtenstein (2013). "Premature mortality in epilepsy and the role of psychiatric comorbidity: a total population study." Lancet 382(9905): 1646-1654.

Geerts, A., W. F. Arts, H. Stroink, E. Peeters, O. Brouwer, B. Peters, L. Laan and C. van Donselaar (2010). "Course and outcome of childhood epilepsy: a 15-year follow-up of the Dutch Study of Epilepsy in Childhood." Epilepsia 51(7): 1189-1197.

Granbichler, C. A., W. Oberaigner, G. Kuchukhidze, G. Bauer, J. P. Ndayisaba, K. Seppi and E. Trinka (2015). "Cause-specific mortality in adult epilepsy patients from Tyrol, Austria: hospital-based study." J Neurol 262(1): 126-133.

Klenerman, P., J. W. Sander and S. D. Shorvon (1993). "Mortality in patients with epilepsy: a study of patients in long term residential care." J Neurol Neurosurg Psychiatry 56(2): 149-152.

Kobulashvili, T., G. Lomidze, S. Kasradze and J. W. Sander (2013). "Premature mortality in a Georgian cohort of people with epilepsy." Epilepsy Research 107(3): 318-322.

Loiseau, J., M. C. Picot and P. Loiseau (1999). "Short-term mortality after a first epileptic seizure: a population-based study." Epilepsia 40(10): 1388-1392.

Moseley, B. D., E. C. Wirrell, L. C. Wong-Kisiel and K. Nickels (2013). "Early onset epilepsy is associated with increased mortality: a population-based study." Epilepsy Res 105(3): 410-414.

Nevalainen, O., A. Auvinen, H. Ansakorpi, M. Artama, J. Raitanen and J. Isojarvi (2012). "Mortality by clinical characteristics in a tertiary care cohort of adult patients with chronic epilepsy." Epilepsia 53(12): e212-214.

Nevalainen, O., J. Raitanen, H. Ansakorpi, M. Artama, J. Isojarvi and A. Auvinen (2013). "Long-term mortality risk by cause of death in newly diagnosed patients with epilepsy in Finland: a nationwide register-based study." Eur J Epidemiol 28(12): 981-990.

Ngugi, A. K., C. Bottomley, G. Fegan, E. Chengo, R. Odhiambo, E. Bauni, B. Neville, I. Kleinschmidt, J. W. Sander and C. R. Newton (2014). "Premature mortality in active convulsive epilepsy in rural Kenya: causes and associated factors." Neurology 82(7): 582-589.

Olesen, J. B., S. Z. Abildstrom, J. Erdal, G. H. Gislason, P. Weeke, C. Andersson, C. Torp-Pedersen and P. R. Hansen (2011). "Effects of epilepsy and selected antiepileptic drugs on risk of myocardial infarction, stroke, and death in patients with or without previous stroke: a nationwide cohort study." Pharmacoepidemiol Drug Saf 20(9): 964-971.

Rafnsson, V., E. Olafsson, W. A. Hauser and G. Gudmundsson (2001). "Cause-specific mortality in adults with unprovoked seizures. A population-based incidence cohort study." Neuroepidemiology 20(4): 232-236.

Terra, V. C., F. A. Scorza, R. M. Arida, R. M. Fernandes, L. Wichert-Ana, H. R. Machado and A. C. Sakamoto (2011). "Mortality in children with severe epilepsy: 10 years of follow-up." Arq Neuropsiquiatr 69(5): 766-769.

White, S. J., A. E. McLean and C. Howland (1979). "Anticonvulsant drugs and cancer. A cohort study in patients with severe epilepsy." Lancet 2(8140): 458-461.

Zielinski, J. J. (1974). "Epilepsy and mortality rate and cause of death." Epilepsia 15(2): 191-201.

Studies overlapping three systematic reviews (n = 20 studies)

Ackers, R., F. M. Besag, E. Hughes, W. Squier, M. L. Murray and I. C. Wong (2011). "Mortality rates and causes of death in children with epilepsy prescribed antiepileptic drugs: a retrospective cohort study using the UK General Practice Research Database." Drug Saf 34(5): 403-413.

Banerjee, T. K., B. K. Ray, S. K. Das, A. Hazra, M. K. Ghosal, A. Chaudhuri, T. Roy and D. K. Raut (2010). "A longitudinal study of epilepsy in Kolkata, India." Epilepsia 51(12): 2384-2391.

Berg, A. T., S. Shinnar, F. M. Testa, S. R. Levy, S. N. Smith and B. Beckerman (2004). "Mortality in childhood-onset epilepsy." Arch Pediatr Adolesc Med 158(12): 1147-1152.

Camfield, C. S., P. R. Camfield and P. J. Veugelers (2002). "Death in children with epilepsy: a population-based study." Lancet 359(9321): 1891-1895.

Ding, D., W. Wang, J. Wu, H. Yang, S. Li, X. Dai, B. Yang, T. Wang, C. Yuan, G. Ma, G. S. Bell, P. Kwan, H. M. de Boer, Z. Hong and J. W. Sander (2013). "Premature mortality risk in people with convulsive epilepsy: long follow-up of a cohort in rural China." Epilepsia 54(3): 512-517.

Hauser, W. A., J. F. Annegers and L. R. Elveback (1980). "Mortality in patients with epilepsy." Epilepsia 21(4): 399-412.

Henriksen, P. B., Jules-Jensen P and Lund M (1970). "The mortality of epileptics. In Brackenridge RDC (Ed) Proceedings of the 10th International Congress of Life Assurance Medicine." Pitman Medical and Scientific Pubishing Co: 139–148.

Holst, A. G., B. G. Winkel, B. Risgaard, J. B. Nielsen, P. V. Rasmussen, S. Haunso, A. Sabers, P. Uldall and J. Tfelt-Hansen (2013). "Epilepsy and risk of death and sudden unexpected death in the young: a nationwide study." Epilepsia 54(9): 1613-1620.

Kaiser, C., G. Asaba, S. Kasoro, T. Rubaale, G. Kabagambe and M. Mbabazi (2007). "Mortality from epilepsy in an onchocerciasis-endemic area in West Uganda." Trans R Soc Trop Med Hyg 101(1): 48-55.

Kamgno, J., S. D. Pion and M. Boussinesq (2003). "Demographic impact of epilepsy in Africa: results of a 10-year cohort study in a rural area of Cameroon." Epilepsia 44(7): 956-963.

Lhatoo, S. D., A. L. Johnson, D. M. Goodridge, B. K. MacDonald, J. W. Sander and S. D. Shorvon (2001). "Mortality in epilepsy in the first 11 to 14 years after diagnosis: multivariate analysis of a long-term, prospective, population-based cohort." Ann Neurol 49(3): 336-344.

Mohanraj, R., J. Norrie, L. J. Stephen, K. Kelly, N. Hitiris and M. J. Brodie (2006). "Mortality in adults with newly diagnosed and chronic epilepsy: a retrospective comparative study." Lancet Neurol 5(6): 481-487.

Morgan, C. L. and M. P. Kerr (2002). "Epilepsy and mortality: a record linkage study in a U.K. population." Epilepsia 43(10): 1251-1255.

Mu, J., L. Liu, Q. Zhang, Y. Si, J. Hu, J. Fang, Y. Gao, J. He, S. Li, W. Wang, J. Wu, J. W. Sander and D. Zhou (2011). "Causes of death among people with convulsive epilepsy in rural West China: a prospective study." Neurology 77(2): 132-137.

Neligan, A., G. S. Bell, A. L. Johnson, D. M. Goodridge, S. D. Shorvon and J. W. Sander (2011). "The long-term risk of premature mortality in people with epilepsy." Brain 134(Pt 2): 388-395.

Nickels, K. C., B. R. Grossardt and E. C. Wirrell (2012). "Epilepsy-related mortality is low in children: a 30-year population-based study in Olmsted County, MN." Epilepsia 53(12): 2164-2171.

Rakitin, A., M. Liik, A. Oun and S. Haldre (2011). "Mortality risk in adults with newly diagnosed and chronic epilepsy: a population-based study." Eur J Neurol 18(3): 465-470.

Shackleton, D. P., R. G. Westendorp, D. G. Trenite and J. P. Vandenbroucke (1999). "Mortality in patients with epilepsy: 40 years of follow up in a Dutch cohort study." J Neurol Neurosurg Psychiatry 66(5): 636-640.

Sillanpaa, M. and S. Shinnar (2010). "Long-term mortality in childhood-onset epilepsy." N Engl J Med 363(26): 2522-2529.

Trinka, E., G. Bauer, W. Oberaigner, J. P. Ndayisaba, K. Seppi and C. A. Granbichler (2013). "Cause-specific mortality among patients with epilepsy: results from a 30-year cohort study." Epilepsia 54(3): 495-501.

Studies that overlapped four systematic reviews (n = 8 studies)

Carpio, A., N. E. Bharucha, P. Jallon, E. Beghi, R. Campostrini, S. Zorzetto and P. P. Mounkoro (2005). "Mortality of epilepsy in developing countries." Epilepsia 46 Suppl 11: 28-32.

Devilat Barros, M., G. Rivera Gomez, V. Gomez Munoz and J. P. Sepulveda Olmos (2004). "[Mortality in children with epilepsy. A clinical prospective study]." Rev Neurol 38(7): 607-614.

Kochen, S. and M. O. Melcon (2005). "Prognosis of epilepsy in a community-based study: 8 years of follow-up in an Argentine community." Acta Neurol Scand 112(6): 370-374.

Lindsten, H., L. Nystrom and L. Forsgren (2000). "Mortality risk in an adult cohort with a newly diagnosed unprovoked epileptic seizure: a population-based study." Epilepsia 41(11): 1469-1473.

Nashef, L., D. R. Fish, J. W. Sander and S. D. Shorvon (1995). "Incidence of sudden unexpected death in an adult outpatient cohort with epilepsy at a tertiary referral centre." J Neurol Neurosurg Psychiatry 58(4): 462-464.

Nicoletti, A., V. Sofia, G. Vitale, S. I. Bonelli, V. Bejarano, F. Bartalesi, D. S. Tran, P. M. Preux, M. Zappia and A. Bartoloni (2009). "Natural history and mortality of chronic epilepsy in an untreated population of rural Bolivia: a follow-up after 10 years." Epilepsia 50(10): 2199-2206.

Nilsson, L., T. Tomson, B. Y. Farahmand, V. Diwan and P. G. Persson (1997). "Cause-specific mortality in epilepsy: A cohort study of more than 9,000 patients once hospitalized for epilepsy." Epilepsia 38(10): 1062-1068.

Olafsson, E., W. A. Hauser and G. Gudmundsson (1998). "Long-term survival of people with unprovoked seizures: a population-based study." Epilepsia 39(1): 89-92.

Appendix E: AMSTAR quality scoring of included systematic reviews

Primary analysis reviews

	Watila 2018	Levira 2017	Thurman 2017	Escalaya 2015	Nevalainen 2014	Shackleton 2002
1. Was an 'a priori' design provided?	No	No	No	No	No	No
2. Was there duplicate study selection and data extraction?	Yes	Yes	Yes	Yes	Yes	No
3. Was a comprehensive literature search performed?	Yes	Yes	Yes	Yes	Yes	Yes
4. Was the status of publication (i.e. grey literature) used as an inclusion criterion?	No	No	No	No	Yes	Yes
5. Was a list of studies (included and excluded) provided?	Yes	No	No	No	No	No
6. Were the characteristics of the included studies provided?	Yes	Yes	Yes	Yes	Yes	Yes
7. Was the scientific quality of the included studies assessed and documented?	Yes	Yes	Yes	No	Yes	No
8. Was the scientific quality of the included studies used appropriately in formulating conclusions?	Yes	Yes	Yes	No	Yes	No
9. Were the methods used to combine the findings of studies appropriate?	Yes	Yes	Yes	Yes	Yes	Yes
10. Was the likelihood of publication bias assessed?	Yes	No	No	No	Yes	No
11. Was the conflict of interest included?	Yes	Yes	Yes	Yes	Yes	Yes
Total AMSTAR Score	**9**	**7**	**7**	**5**	**9**	**5**

Secondary analysis reviews

	Abdel-Mannan 2019	Ali 2017	Bell 2008	Bell 2009	Harden 2017	Liebenthal 2015	Monté 2006	Téllez-Zenteno 2005	Young 2015
1. Was an 'a priori' design provided?	No	No	No	No	No	No	No	No	No
2. Was there duplicate study selection and data extraction?	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	No
3. Was a comprehensive literature search performed?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No
4. Was the status of publication (i.e. grey literature) used as an inclusion criterion?	Yes	No	No	No	No	No	No	No	No
5. Was a list of studies (included and excluded) provided?	Yes	No	No	No	No	No	No	No	No
6. Were the characteristics of the included studies provided?	Yes	No	Yes	Yes	No	No	No	Yes	Yes
7. Was the scientific quality of the included studies assessed and documented?	Yes	No	No	No	Yes	No	Yes	Yes	No
8. Was the scientific quality of the included studies used appropriately in formulating conclusions?	Yes	No	No	No	Yes	No	Yes	Yes	No
9. Were the methods used to combine the findings of studies appropriate?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
10. Was the likelihood of publication bias assessed?	No	No	No	No	No	Yes	No	No	No
11. Was the conflict of interest included?	Yes	Yes	Yes	Yes	Yes	Yes	No	No	Yes
Total AMSTAR Score	**9**	**4**	**5**	**5**	**6**	**5**	**4**	**6**	**3**

Appendix 2 – Supplementary material to chapter 3

Appendix S1: Search strategies for MEDLINE and Embase

Ovid MEDLINE(R) Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Ovid MEDLINE(R) Daily and Ovid MEDLINE(R) 1946 to 03 July 2018	OVID Embase 1974 to 03 July 2018
1 Epilepsy, Complex Partial/ or Epilepsy, Reflex/ or Epilepsy, Absence/ or Drug Resistant Epilepsy/ or Epilepsy/ or Epilepsy, Rolandic/ or Epilepsy, Partial, Motor/ or Epilepsy, Benign Neonatal/ or Epilepsy, Tonic-Clonic/ or Epilepsy, Post-Traumatic/ or Epilepsy, Partial, Sensory/ or Epilepsy.mp. or Epilepsy, Temporal Lobe/ or Epilepsy, Frontal Lobe/ or Myoclonic Epilepsy, Juvenile/ or Epilepsy, Generalized/	1 reflex epilepsy/ or photosensitive epilepsy/ or grand mal epilepsy/ or epilepsy.mp. or drug resistant epilepsy/ or experimental epilepsy/ or severe myoclonic epilepsy in infancy/ or childhood absence epilepsy/ or benign childhood epilepsy/ or catamenial epilepsy/ or symptomatic epilepsy/ or startle epilepsy/ or generalized epilepsy/ or epilepsy/ or mesial temporal lobe epilepsy/ or rolandic epilepsy/ or traumatic epilepsy/ or myoclonic astatic epilepsy/ or temporal lobe epilepsy/ or intractable epilepsy/ or focal epilepsy/ or "seizure, epilepsy and convulsion"/ or myoclonus epilepsy/ or lateral temporal lobe epilepsy/ or frontal lobe epilepsy/
2 Databases, Factual/	2 factual database/ or data base/
3 "Reproducibility of Results"/	3 reproducibility/
4 Algorithms/ or algorithm*.mp.	4 algorithm/ or algorithm*.mp.
5 3 or 4	5 3 or 4
6 2 and 5	6 2 and 5
7 1 and 6	7 1 and 6
8 Administrative Claims, Healthcare/ or administrativ*.mp. or insurance data*.mp. or claims data*.mp. or Veterans Health Administration.mp.	8 "administrative claims (health care)"/ or administrative*.mp. or insurance data*.mp. or claims data*.mp. or Veterans Health Administration.mp.
9 administrat* data*.mp.	9 administrat* data*.mp.
10 routin* data*.mp.	10 routin* data*.mp.
11 big data.mp.	11 big data.mp.
12 8 or 9 or 10 or 11	12 8 or 9 or 10 or 11
13 1 and 12	13 1 and 12
14 Algorithms/ or algorithm*.mp.	14 algorithm/ or algorithm*.mp.
15 "Sensitivity and Specificity"/ or "Predictive Value of Tests"/ or predictive value.mp.	15 (specificity or sensitivity).mp. or "sensitivity and specificity"/ or positive* predict* value*.mp. or negative* predict* value*.mp.
16 positive* predict* value*.mp.	16 area under the curve/ or area* under* curve*.mp.
17 negative* predict* value*.mp.	17 roc curve/ or receiver operating characteristic/ or ROC curve*.mp.
18 sensitivity.mp.	
19 specificity.mp.	
20 area* under* curve*.mp. or Area Under Curve/	
21 ROC Curve/ or ROC curve*.mp.	

22 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21

23 code*.mp.

24 (ICD-9 or ICD-10).mp. or "International Classification of Diseases"/ or Clinical Coding/ or read code*.mp.

25 23 or 24

26 22 and 25

27 1 and 26

28 (validat* or validity).mp. or Validation Studies/

29 Medical Records/ or medical record*.mp. or medical case note*.mp.

30 electronic health records.mp. or Medical Records Systems, Computerized/ or Electronic Health Records/

31 Registries.mp. or Registries/

32 29 or 30 or 31

33 28 and 32

34 1 and 33

35 7 or 13 or 27 or 34

36 limit 35 to yr="1975 -Current"

37 Animals/ not Humans/

38 36 not 37

18 14 or 15 or 16 or 17

19 code*.mp. or "Read code"/

20 ICD-10.mp. or "International Classification of Diseases"/ or ICD-10/ or disease classification/ or ICD-9.mp. or ICD-9/

21 19 or 20

22 18 and 21

23 1 and 22

24 validation study/ or validation process/ or validat*.mp. or validity/ or predictive validity/ or validity.mp.

25 electronic medical record/ or medical record/ or medical record*.mp. or medical case note*.mp.

26 Registries.mp. or register/

27 25 or 26

28 24 and 27

29 1 and 28

30 7 or 13 or 23 or 29

31 limit 30 to yr="1975 -Current"

32 animal/ not human/

33 31 not 32

Appendix S2: Data abstraction tool

- What is the study title?
- Who are the study authors?
- What is the year of study publication?
- What country(s) was the study conducted in?
- Does the study explicitly identify as utilising “administrative data” (yes/no)?
    - If not, how is this identified by the reviewer? – e.g. from descriptions given of the databases utilised and background knowledge about them or from correspondence with authors.
- What are the eligibility criteria for participation in the validation cohort (i.e. the cohort of patients to which the reference standard will be applied)?
- Where, when, and how were potentially eligible validation cohort participants identified? Include within this:
    - What is the name of the administrative database(s) on which the validation cohort was identified?
    - What are the setting and location of the administrative database(s) from which the validation cohort was identified? – e.g. is it primary care, secondary care, tertiary care, outpatient care, emergency care?
    - What were the epilepsy ICD-codes (or other disease classification system codes) used to identify the validation cohort within the administrative database(s)? – i.e. what are the diagnostic epilepsy codes that will be validated by the study?
    - What was the size of the validation cohort identified by the epilepsy codes?
    - What other information was obtained about an individual to help identify epilepsy on the administrative database? For example, describe if they linked an individual’s ICD epilepsy codes with the investigations they underwent (such as EEG) or the AED they were taking.
- What were the demographic and clinical characteristics of the validation cohort? – i.e. age, gender, type of epilepsy, comorbidities, and AEDs.
- Describe the reference standard. Include the following:
    - Name of the reference standard used;

- What type of reference standard it was, e.g. clinical assessment, medical records, or validated disease registry;
- Any rational given for choosing this reference standard (if alternatives exist).

- What were the number, training and expertise of persons reading the reference standard?
- Describe any methods used to blind persons reading the reference standard to how the validation cohort were coded diagnostically on the administrative database, i.e. how a person reviewing the medical records diagnosis of an individual was made unaware of their administrative ICD diagnosis.
- Describe any methods used to blind persons reading the diagnostic administrative data codes to results of the reference standard diagnoses.
- What method was used to estimate the disease-identification accuracy of the administrative database? (i.e. PPV/sensitivity/NPV/specificity, true positives and false positives (allowing calculation of PPV), true positives and false negatives (allowing calculation of sensitivity), true negatives and false negatives (allowing calculation of NPV), and true negatives and false positives (allowing calculation of specificity))
  - What were the results of this?
    - i.e. provide figures for these estimates of diagnostic accuracy for epilepsy and confidence intervals if available (or calculate them from the available data)
- How were indeterminate or missing administrative database diagnoses or reference standard results handled?
- What were the time interval and any clinical interventions given between reference standard diagnosis and administrative database diagnosis?

Appendix S3: Adapted QUADAS-2 assessment form

Study:

1. Patient selection

a) Risk of bias: *Could the selection of patients have introduced bias?*	
Insert/summarise the methods of patient selection: What is the review question?	
Was a consecutive or random sample of patients enrolled?	Yes / No / Unclear
Did the study avoid inappropriate exclusions? *-inappropriate exclusions are e.g. excluding "difficult-to-diagnose" patients as this may result in overestimation of diagnostic accuracy*	Yes / No / Unclear
Could the selection of patients have introduced bias? *- if all YES then risk = LOW, if any NO then risk = HIGH, if any UNCLEAR then risk = UNCLEAR)* **Risk: Low/ High/ Unclear**	

b) Applicability assessment: *Are there concerns that the included patients and setting do not match the review question?*
Concerns about applicability may exist if patients included in the study differ from those targeted by the review question in terms of e.g. severity or type of epilepsy, age/ethnicity, the comorbidities present, or the study setting (e.g. a study on status that used outpatients is probably not an appropriate setting)
Is there concern that the included patients do not match the review question? **Concern: Low / High / Unclear**

2. Index test

a) Risk of bias: *Could the conduct or interpretation of the administrative database have introduced bias?*	
Describe the administrative database and how it was used and interpreted: Where available, include comment on how coding was done, by whom, and whether or not there was reimbursement for coding, and describe any blinding procedures.	
Was blinding used? -if the coding is always conducted and interpreted before review of the medical records/gold standard, this item can be rated "yes."	Yes / No / Unclear
If a diagnostic threshold was used, was it pre-specified? – e.g. was a diagnostic algorithm specified in the methods, e.g. epilepsy = ICD-10 G40-41 + AED	Yes / No / Unclear
Could the conduct or interpretation of the index test have introduced bias? *- if all YES then risk = LOW, if any NO then risk = HIGH, if any UNCLEAR then risk = UNCLEAR)* **Risk: Low / High / Unclear**	

b) Applicability assessment: *Are there concerns that the administrative database, its conduct, or interpretation differ from the review question?*
Does the administrative database conform to the definition of a routinely collected dataset? – i.e. was the coded information collected not for a prior research purpose but rather for routine purposes e.g. health system management or compensation **Concern: Low / High / Unclear**

3. Reference standard

a) Risk of bias: *Could the reference standard, its conduct, or its interpretation have introduced bias?*	
Describe the reference standard and how it was conducted and interpreted: Where available, include comment on quality of the reference standard (e.g. quality of the medical records used to confirm epilepsy), including the level of experience of clinicians making the diagnosis, and whether there was access to diagnostic tests such as electroencephalography and telemetry, and comment on the thresholds/criteria (e.g. ILAE) used to make a diagnosis of epilepsy.	
Is the reference standard likely to correctly classify the epilepsy?	Yes / No / Unclear
Was blinding used? I.e. were the reference standard diagnosis-making steps for each patient done without knowledge of that patient's administrative coded diagnosis?	Yes / No / Unclear
Could the reference standard, its conduct, or its interpretation have introduced bias? - *if all YES then risk = LOW, if any NO then risk = HIGH, if any UNCLEAR then risk = UNCLEAR)* **Risk: Low / High / Unclear**	

b) Applicability assessment: Are there concerns that the target condition as defined by the reference standard does not match the question?
Are there concerns that epilepsy, as defined by the reference standard, does not match the epilepsy proposed to be being investigated in the review question? **Concern: Low / High / Unclear**

4. Flow and timing

a) Risk of bias: *Could the patient flow have introduced bias?*	
Describe whether there were any patients in the validation cohort (i.e. captured by the coding algorithm) who were not found within the reference standard (e.g. missing case notes) or who were excluded from cross-tabulation (calculation of PPV or sensitivity) for whatever reason. Describe the time interval and any interventions (e.g. curative epilepsy surgery) between patients getting an administrative coded diagnosis and their confirmation of disease/no disease in the reference standard review. Things may have changed if code was looked at one time and then the notes looked at many years later, for example.	
Was there an appropriate interval between administrative database diagnosis and reference standard diagnosis? – *i.e. probably not so relevant for lifelong disease like epilepsy, but may be important if diagnosis made in childhood and you suspect some cases may have resolved if study done in adulthood.*	Yes / No / Unclear
Did all patients receive the same reference standard? – *or e.g. were some diagnosis confirmed e.g. by family/GP and others by case notes reviewed by the authors*	Yes / No / Unclear
Were all patients included in the analysis? - *All participants recruited into the study should be included in the analysis (19). A potential for bias exists if the number of patients enrolled differs from the number of patients included in the 2 × 2 table of results/PPV/Sn calculation, because patients lost to follow-up differ systematically from those who remain. Mark as low risk if a best/worst case scenario is done for eg those with missing notes, to say "if they had epilepsy results = x, if they didn't, then results = y)*	Yes / No / Unclear
Could the patient flow have introduced bias? - *if all YES then risk = LOW, if any NO then risk = HIGH, if any UNCLEAR then risk = UNCLEAR)* **Risk: Low / High / Unclear**	

Appendix S4: Kappa statistic interrater agreement: study selection

Reviewers: GKM and KB

There was independent agreement (i.e. without the need for consensus discussion) on 175 records (20 to include, 155 to exclude), and mutual consensus was reached (i.e. without the need for third reviewer adjunction) for the remaining 22 records (kappa = 0.58, 95% CI 0.42–0.74).

	Reviewer 2: INCLUDE	Reviewer 2: EXCLUDE	**Total**
Reviewer 1: INCLUDE	20	15	**35**
Reviewer 1: EXCLUDE	7	155	**162**
Total	**27**	**170**	**197**

Number of observed agreements: 175 (88.83% of the observations)

Number of agreements expected by chance: 144.6 (73.40% of the observations)

Kappa= 0.580

SE of kappa = 0.080

95% confidence interval: From 0.424 to 0.737

Interpretation[1]:

Kappa < 0: No agreement

Kappa between 0.00 and 0.20: Slight agreement

Kappa between 0.21 and 0.40: Fair agreement

Kappa between 0.41 and 0.60: Moderate agreement

Kappa between 0.61 and 0.80: Substantial agreement

Kappa between 0.81 and 1.00: Almost perfect agreement."

1. Landis, J.R.; Koch, G.G. (1977). The measurement of observer agreement for categorical data. Biometrics. 33 (1): 159-174. https://doi.org/10.2307%2F2529310

Appendix S5: Reference list of excluded studies

Aaberg, K. M., I. J. Bakken, M. I. Lossius, C. Lund Soraas, S. E. Haberg, C. Stoltenberg, P. Suren and R. Chin (2016). "Comorbidity and Childhood Epilepsy: A Nationwide Registry Study." Pediatrics 138(3).

Aaberg, K. M., P. Suren, C. L. Soraas, I. J. Bakken, M. I. Lossius, C. Stoltenberg and R. Chin (2017). "Seizures, syndromes, and etiologies in childhood epilepsy: The International League Against Epilepsy 1981, 1989, and 2017 classifications used in a population-based cohort." Epilepsia 58(11): 1880-1891.

Ansari, Z., K. Brown and N. Carson (2008). "Association of epilepsy and burns - a case control study." Aust Fam Physician 37(7): 584-589.

Baca, C. B. and G. L. Barkley (2015). "Can big data help us close an epilepsy care gap?" Neurology 85(12): 1014-1015.

Bakaki, P. M., S. M. Koroukian, L. W. Jackson, J. M. Albert and K. Kaiboriboon (2013). "Defining incident cases of epilepsy in administrative data." Epilepsy Res 106(1-2): 273-279.

Banta, J. E., A. Addison and W. L. Beeson (2015). "Spatial patterns of epilepsy-related emergency department visits in California." Journal of public health research 4(1).

Baviera, M., M. C. Roncaglioni, M. Tettamanti, T. Vannini, I. Fortino, A. Bortolotti, L. Merlino and E. Beghi (2017). "Diabetes mellitus: a risk factor for seizures in the elderly-a population-based study." Acta Diabetol 54(9): 863-870.

Beg, J. M., T. D. Anderson, K. Francis, L. M. Meckley, D. Fitzhenry, T. Foster, S. Sukhtankar, S. J. Kanes and L. M. Moura (2017). "Burden of illness for super-refractory status epilepticus patients." J Med Econ 20(1): 45-53.

Bergin, P. S., T. Ip, R. Sheehan, R. W. Frith, L. G. Sadleir, N. McGrath, A. Ranta and E. B. Walker (2010). "Using the Internet to recruit patients for epilepsy trials: results of a New Zealand pilot study." Epilepsia 51(5): 868-873.

Berlowitz, D. R. and M. J. Pugh (2007). "Pharmacoepidemiology in community-dwelling elderly taking antiepileptic drugs." Int Rev Neurobiol 81: 153-163.

Bjørk, M. H., G. Veiby, O. Spigset and N. E. Gilhus (2014). "Using the Norwegian Mother and Child Cohort Study to determine risk factors for delayed development and neuropsychiatric." Norsk epidemiologi 24(1-2).

Bonthapally, V., A. Eldar-Lissa, S. Schubert-Bast, S. Ansorge and A. Strzelczyk (2017). "Neurological emergencies of childhood: A population-based study of pediatric status epilepticus in Germany." Neurocritical Care 27 (2 Supplement 1): S315.

Bourgeois, F. T., K. L. Olson, A. Poduri and K. D. Mandl (2015). "Comparison of Drug Utilization Patterns in Observational Data: Antiepileptic Drugs in Pediatric Patients." Paediatr Drugs 17(5): 401-410.

Burke, J. P., A. Jain, W. Yang, J. P. Kelly, M. Kaiser, L. Becker, L. Lawer and C. J. Newschaffer (2014). "Does a claims diagnosis of autism mean a true case?" Autism 18(3): 321-330.

Carnahan, R. M. (2012). "Mini-Sentinel's systematic reviews of validated methods for identifying health outcomes using administrative data: summary of findings and suggestions for future research." Pharmacoepidemiology and drug safety 21: 90-99.

Chan, S. and M. Pugh (2016). "The natural history of comorbid conditions in epilepsy." Journal of the American Geriatrics Society 1): S234-S235.

Chang, C. S., C. H. Liao, C. C. Lin, H. Y. Lane, F. C. Sung and C. H. Kao (2014). "Patients with epilepsy are at an increased risk of subsequent stroke: a population-based cohort study." Seizure 23(5): 377-381.

Chang, Y. T., P. C. Chen, I. J. Tsai, F. C. Sung, Z. N. Chin, H. T. Kuo, C. H. Tsai and I. C. Chou (2011). "Bidirectional relation between schizophrenia and epilepsy: a population-based retrospective cohort study." Epilepsia 52(11): 2036-2042.

Charlton, R. A., A. McGrogan, J. Snowball, L. M. Yates, A. Wood, J. Clayton-Smith, W. H. Smithson, J. L. Richardson, N. McHugh, S. H. Thomas, G. A. Baker and R. Bromley (2017). "Sensitivity of the UK Clinical Practice Research Datalink to Detect Neurodevelopmental Effects of Medicine Exposure in Utero: Comparative Analysis of an Antiepileptic Drug-Exposed Cohort." Drug Saf 40(5): 387-397.

Chen, C. H., C. L. Lin and C. H. Kao (2015). "Irritable Bowel Syndrome Increases the Risk of Epilepsy: A Population-Based Study." Medicine (Baltimore) 94(36): e1497.

Chen, L. L., C. B. Baca, J. Choe, J. W. Chen, M. E. Ayad and E. M. Cheng (2014). "Posttraumatic epilepsy in Operation Enduring Freedom/Operation Iraqi Freedom veterans." Mil Med 179(5): 492-496.

Chen, S. Y., N. Wu, L. Boulanger and P. Sacco (2013). "Antiepileptic drug treatment patterns and economic burden of commercially-insured patients with refractory epilepsy with partial onset seizures in the United States." J Med Econ 16(2): 240-248.

Chiang, K. L., F. C. Kuo, J. Y. Lee and C. Y. Huang (2018). "Association of epilepsy and asthma: a population-based retrospective cohort study." PeerJ 6: e4792.

Choi, H., A. Pack, M. S. Elkind, W. T. Longstreth, Jr., T. G. Ton and F. Onchiri (2017). "Predictors of incident epilepsy in older adults: The Cardiovascular Health Study." Neurology 88(9): 870-877.

Chou, I. C., C. H. Wang, W. D. Lin, F. J. Tsai, C. C. Lin and C. H. Kao (2016). "Risk of epilepsy in type 1 diabetes mellitus: a population-based cohort study." Diabetologia 59(6): 1196-1203.

Chung, T. T., C. Y. Lin, W. Y. Huang, C. L. Lin, F. C. Sung and C. H. Kao (2013). "Risks of subsequent epilepsy among patients with hypertensive encephalopathy: a nationwide population-based study." Epilepsy Behav 29(2): 374-378.

Clark, N. M., L. Lachance, M. B. Benedict, R. Little, H. Leo, D. F. Awad and M. K. Wilkin (2015). "The extent and patterns of multiple chronic conditions in low-income children." Clinical pediatrics 54(4): 353-358.

Cookson, R., M. Laudicella and P. L. Donni (2012). "Measuring change in health care equity using small-area administrative data - evidence from the English NHS 2001-2008." Soc Sci Med 75(8): 1514-1522.

Copeland, L. A., A. B. Ettinger, J. E. Zeber, J. M. Gonzalez and M. J. Pugh (2011). "Psychiatric and medical admissions observed among elderly patients with new-onset epilepsy." BMC Health Serv Res 11: 84.

Cramer, J. A., Z. J. Wang, E. Chang, A. Powers, R. Copher, D. Cherepanov and M. S. Broder (2014). "Healthcare utilization and costs in adults with stable and uncontrolled epilepsy." Epilepsy Behav 31: 356-362.

Cramer, J. A., Z. J. Wang, E. Chang, A. Powers, R. Copher, D. Cherepanov and M. S. Broder (2014). "Healthcare utilization and costs in children with stable and uncontrolled epilepsy." Epilepsy Behav 32: 135-141.

Danila, O., J. P. Hirdes, C. J. Maxwell, R. A. Marrie, S. Patten, T. Pringsheim and N. Jette (2014). "Prevalence of neurological conditions across the continuum of care based on interRAI assessments." BMC Health Serv Res 14: 29.

Devine, S. T., E. Weisbart, J. Barron and A. Behm (2010). "Acute epilepsy exacerbations in patients switched between A-rated anti-epileptic drugs." Curr Med Res Opin 26(2): 455-463.

Devinsky, O., C. Dilley, M. Ozery-Flato, R. Aharonov, Y. Goldschmidt, M. Rosen-Zvi, C. Clark and P. Fritz (2016). "Changing the approach to treatment choice in epilepsy using big data." Epilepsy Behav 56: 32-37.

Divino, V., A. A. Petrilla, V. Bollu, F. Velez, A. Ettinger and C. Makin (2015). "Clinical and economic burden of breakthrough seizures." Epilepsy Behav 51: 40-47.

Einarsdottir, K., D. B. Preen, J. D. Emery and C. D. Holman (2010). "Regular primary care decreases the likelihood of mortality in older people with epilepsy." Med Care 48(5): 472-476.

Eldar-Lissai, A., T. Jean-Eric, N. Burke, J. G. Burneo, M. A. Ferro, M. Keezer, O. C. Snead and R. B. Hopkins (2017). "Epidemiology of super refractory status epilepticus in Canada." Neurocritical Care 27 (2 Supplement 1): S344-S345.

Erickson, S. C., L. Le, S. D. Ramsey, B. K. Solow, A. Zakharyan, K. M. Stockl, A. S. Harada and B. Curtis (2011). "Clinical and pharmacy utilization outcomes with brand to generic antiepileptic switches in patients with epilepsy." Epilepsia 52(7): 1365-1371.

Ertl, J., J. Hapfelmeier, T. Peckmann, B. Forth and A. Strzelczyk (2016). "Guideline conform initial monotherapy increases in patients with focal epilepsy: A population-based study on German health insurance data." Seizure 41: 9-15.

Etminan, M., A. Samii and J. M. Brophy (2010). "Statin use and risk of epilepsy: a nested case-control study." Neurology 75(17): 1496-1500.

Ettinger, A. B., L. A. Copeland, J. E. Zeber, A. C. Van Cott and M. J. Pugh (2010). "Are psychiatric disorders independent risk factors for new-onset epilepsy in older individuals?" Epilepsy Behav 17(1): 70-74.

Ettinger, A. B., M. B. Good, R. Manjunath, R. Edward Faught and T. Bancroft (2014). "The relationship of depression to antiepileptic drug adherence and quality of life in epilepsy." Epilepsy Behav 36: 138-143.

Ettinger, A. B., R. Manjunath, S. D. Candrilli and K. L. Davis (2009). "Prevalence and cost of nonadherence to antiepileptic drugs in elderly patients with epilepsy." Epilepsy and Behavior 14(2): 324-329.

Fagnani, F., H. Vespignani, O. Kusnik-Joinville, M. Bertrand, C. Murat, L. Levy-Bachelot and P. Kahane (2013). "[Use of drug reimbursement as markers of disease for epidemiological and cost analysis: The case of severe epilepsy in France]." Presse Med 42(9 Pt 1): e285-292.

Farmen, A. H., J. H. Grundt, T. Tomson, K. O. Nakken, J. Nakling, P. Mowinchel, M. Oie and M. I. Lossius (2016). "Age-specific birth rates in women with epilepsy: a population-based study." Brain Behav 6(8): e00492.

Faught, E., M. S. Duh, J. R. Weiner, A. Guerin and M. C. Cunnington (2008). "Nonadherence to antiepileptic drugs and increased mortality: findings from the RANSOM Study." Neurology 71(20): 1572-1578.

Faught, E., S. Helmers, D. Thurman, H. Kim and L. Kalilani (2018). "Patient characteristics and treatment patterns in patients with newly diagnosed epilepsy: A US database analysis." Epilepsy Behav 85: 37-44.

Faught, E., S. L. Helmers, C. E. Begley, D. J. Thurman, C. Dilley, C. Clark and P. Fritz (2015). "Newer antiepileptic drug use and other factors decreasing hospital encounters." Epilepsy Behav 45: 169-175.

Faught, R., S. Helmers, D. Thurman, H. Kim, T. Durgin and L. Kalilani (2016). "Comparing patient characteristics and treatment patterns in incident patients with or without lesional epilepsy: A us database analysis." Neurology. Conference: 68th American Academy of Neurology Annual Meeting, AAN 86(16 SUPPL. 1).

Faught, R. E., J. R. Weiner, A. Guerin, M. C. Cunnington and M. S. Duh (2009). "Impact of nonadherence to antiepileptic drugs on health care utilization and costs: findings from the RANSOM study." Epilepsia 50(3): 501-509.

Foreman, P., A. Van Cott and M. Jo Pugh (2013). "Increased mortality in elderly veterans diagnosed with epilepsy." Neurology. Conference: 65th American Academy of Neurology Annual Meeting. San Diego, CA United States. Conference Publication: 80(1 MeetingAbstracts).

Foreman, P. J., A. VanCott and M. J. Pugh (2015). "Mortality in elderly veterans with new and chronic epilepsy." Epilepsy Currents 1): 305.

Garnett, W. R., T. D. Gilbert and P. O'Connor (2005). "Patterns of care, outcomes, and direct health plan costs of antiepileptic therapy: a pharmacoeconomic analysis of the available carbamazepine formulations." Clin Ther 27(7): 1092-1103.

Gidal, B. E., J. A. French, P. Grossman and G. Le Teuff (2009). "Assessment of potential drug interactions in patients with epilepsy: impact of age and sex." Neurology 72(5): 419-425.

Gillberg, C., S. Lundstrom, E. Fernell, G. Nilsson and B. Neville (2017). "Febrile Seizures and Epilepsy: Association With Autism and Other Neurodevelopmental Disorders in the Child and Adolescent Twin Study in Sweden." Pediatric Neurology 74: 80-86.e82.

Giussani, G., C. Franchi, P. Messina, A. Nobili, E. Beghi and E. Group (2014). "Prevalence and incidence of epilepsy in a well-defined population of Northern Italy." Epilepsia 55(10): 1526-1533.

Griffiths, R. I., P. N. Schrammel, G. L. Morris, S. H. Wills, D. M. Labiner and M. J. Strauss (1999). "Payer costs of patients diagnosed with epilepsy." Epilepsia 40(3): 351-358.

Groth, A., S. Borghs, P. Gille, L. Joeres and T. Wilke (2017). "Incidence and prevalence of epilepsy in Germany." Value in Health 20 (9): A720.

Groth, A., T. Wilke, S. Borghs, P. Gille and L. Joeres (2017). "Real life pharmaceutical treatment patterns for adult patients with focal epilepsy in Germany: a longitudinal and cross-sectional analysis of recently approved anti-epileptic drugs." Ger Med Sci 15: Doc09.

Guilhoto, L. M. F. F. (2011). "Revised terminology and concepts for organization of seizures and epilepsies: Report of the ILAE commission on classification and terminology, 2005-2009. New paradigms?. [Portuguese]." Journal of Epilepsy and Clinical Neurophysiology 17(3): 100-105.

Helmers, S. L., D. J. Thurman, T. L. Durgin, A. K. Pai and E. Faught (2015). "Descriptive epidemiology of epilepsy in the U.S. population: A different approach." Epilepsia 56(6): 942-948.

Hemminki, K., X. Li, S. E. Johansson, K. Sundquist and J. Sundquist (2006). "Familial risks for epilepsy among siblings based on hospitalizations in Sweden." Neuroepidemiology 27(2): 67-73.

Hernandez, R. L., L. Thorpe, P. Pahwa and Z. J. F. Tellez (2016). "Secular trends in the incidence of epilepsy in the province of Saskatchewan, Canada (2001-2010)." Epilepsia 57 (Supplement 2): 194.

Hernandez-Ronquillo, L., L. Thorpe, P. Pahwa and J. F. Tellez-Zenteno (2018). "Secular trends and population differences in the incidence of epilepsy. A population-based study from Saskatchewan, Canada." Seizure 60: 8-15.

Hixson, J. D., S. L. Van Bebber and K. M. Bertko (2015). "Interest in a digital health tool in veterans with epilepsy: results of a phone survey." Mil Med 180(4): 387-390.

Holden, E. W., H. Thanh Nguyen, E. Grossman, S. Robinson, L. S. Nelson, M. J. Gunter, A. Von Worley and D. J. Thurman (2005). "Estimating prevalence, incidence, and disease-related mortality for patients with epilepsy in managed care organizations." Epilepsia 46(2): 311-319.

Hsieh, L.-P. and C.-Y. Huang (2008). "Prevalence of treated epilepsy in western medicine among the adult population in Taiwan: a study conducted using antiepileptic drug prescription data." Epilepsy research 80(2-3): 114-118.

Hsieh, L. P. and C. Y. Huang (2009). "Antiepileptic drug utilization in Taiwan: analysis of prescription using National Health Insurance database." Epilepsy Res 84(1): 21-27.

Ip, Q., D. C. Malone, J. Chong, R. B. Harris and D. M. Labiner (2018). "An update on the prevalence and incidence of epilepsy among older adults." Epilepsy Res 139: 107-112.

Ishihara, L., D. J. Webb, M. Irizarry and J. Weil (2010). "Exploring differential prescribing between anti-epileptic drugs in epilepsy patients with a history of mood disorders." Pharmacoepidemiol Drug Saf 19(3): 289-295.

Ivanova, J. I., H. G. Birnbaum, Y. Kidolezi, Y. Qiu, D. Mallett and S. Caleo (2010). "Direct and indirect costs associated with epileptic partial onset seizures among the privately insured in the United States." Epilepsia 51(5): 838-844.

Ivanova, J. I., H. G. Birnbaum, Y. Kidolezi, Y. Qiu, D. Mallett and S. Caleo (2010). "Economic burden of epilepsy among the privately insured in the US." Pharmacoeconomics 28(8): 675-685.

Jacobs, K., M. Julyan, M. S. Lubbe, J. R. Burger and M. Cockeran (2016). "Anti-epileptic prescribing patterns in the South African private health sector (2008–2013)." South African Family Practice 58(4): 142-147.

Jang, S. H., B. Kang, S. Lee, K. Jung, S. Yi, Y. Jo and H. Cheong (2012). "Accuracy of epilepsy survey by trained medical record administrators." Epilepsy Currents. Conference: 65th Annual Meeting of the American Epilepsy Society, AES. Baltimore, MD United States. Conference Publication: 12(1 SUPPL. 1).

Jennum, P., J. Christensen, R. Ibsen and J. Kjellberg (2016). "Long-term socioeconomic consequences and health care costs of childhood and adolescent-onset epilepsy." Epilepsia 57(7): 1078-1085.

Jette, N., E. Beghi, D. Hesdorffer, S. L. Moshe, S. M. Zuberi, M. T. Medina and D. Bergen (2015). "ICD coding for epilepsy: past, present, and future--a report by the International League Against Epilepsy Task Force on ICD codes in epilepsy." Epilepsia 56(3): 348-355.

Jette, N., H. Quan, P. Faris, S. Dean, B. Li, A. Fong and S. Wiebe (2008). "Health resource use in epilepsy: Significant disparities by age, gender, and aboriginal status." Epilepsia 49(4): 586-593.

Joyce, N. R., J. Fishman, S. Green, D. M. Labiner, I. Wild and D. C. Grabowski (2018). "Cost sharing for antiepileptic drugs: medication utilization and health plan costs." Am J Manag Care 24(6): e183-e189.

Jung, J., H. Y. Seo, Y. A. Kim, I. H. Oh, Y. H. Lee and S. J. Yoon (2013). "The economic burden of epilepsy in Korea, 2010." J Prev Med Public Health 46(6): 293-299.

Kaiboriboon, K., P. M. Bakaki, S. D. Lhatoo and S. Koroukian (2013). "Incidence and prevalence of treated epilepsy among poor health and low-income Americans." Neurology 80(21): 1942-1949.

Kaiboriboon, K., N. K. Schiltz, P. M. Bakaki, S. D. Lhatoo and S. M. Koroukian (2014). "Premature mortality in poor health and low income adults with epilepsy." Epilepsia 55(11): 1781-1788.

Kang, B. S., H. K. Cheong, K. Y. Jung, S. H. Jang, J. K. Yoo, D. W. Kim, S. E. Chung and S. Y. Lee (2013). "The validity and reliability of characterizing epilepsy based on an external review of medical records." Epidemiol Health 35: e2013006.

Kee, V. R., B. Gilchrist, M. A. Granner, N. R. Sarrazin and R. M. Carnahan (2012). "A systematic review of validated methods for identifying seizures, convulsions, or epilepsy using administrative and claims data." Pharmacoepidemiol Drug Saf 21 Suppl 1: 183-193.

Keller, J., Y. K. Chen and H. C. Lin (2012). "Association between epilepsy and erectile dysfunction: evidence from a population-based study." J Sex Med 9(9): 2248-2255.

Kim, D. W., S. Y. Lee, S. E. Chung, H. K. Cheong, K. Y. Jung and S. Korean Epilepsy (2014). "Clinical characteristics of patients with treated epilepsy in Korea: a nationwide epidemiologic study." Epilepsia 55(1): 67-75.

Kim, H., D. J. Thurman, T. Durgin, E. Faught and S. Helmers (2016). "Estimating Epilepsy Incidence and Prevalence in the US Pediatric Population Using Nationwide Health Insurance Claims Data." J Child Neurol 31(6): 743-749.

Kozyrskyj, A. L. and A. N. Prasad (2004). "The burden of seizures in Manitoba children: a population-based study." Can J Neurol Sci 31(1): 48-52.

Kurth, T., B. E. Lewis and A. M. Walker (2010). "Health care resource utilization in patients with active epilepsy." Epilepsia 51(5): 874-882.

Kwon, C., M. Liu, H. Quan, V. Thoo, S. Wiebe and N. Jette (2011). "Motor vehicle accidents, suicides, and assaults in epilepsy: a population-based study." Neurology 76(9): 801-806.

Kwon, C. S., M. Liu, H. Quan, S. Wiebe, J. McChesney, E. Wirrell, L. Hamiwka and N. Jette (2010). "The incidence of injuries in persons with and without epilepsy--a population-based study." Epilepsia 51(11): 2247-2253.

Lee, S. Y., K. Y. Jung, I. K. Lee, S. D. Yi, Y. W. Cho, D. W. Kim, S. S. Hwang, S. Kim and S. Korean Epilepsy (2012). "Prevalence of treated epilepsy in Korea based on national health insurance data." J Korean Med Sci 27(3): 285-290.

Lee, W. C., S. Arcona, S. K. Thomas, Q. Wang, M. S. Hoffmann and C. L. Pashos (2005). "Effect of comorbidities on medical care use and cost among refractory patients with partial seizure disorder." Epilepsy & Behavior 7(1): 123-126.

LeLorier, J., M. S. Duh, P. E. Paradis, P. Lefebvre, J. Weiner, R. Manjunath and O. Sheehy (2008). "Clinical consequences of generic substitution of lamotrigine for patients with epilepsy." Neurology 70(22 Pt 2): 2179-2186.

Leong, C., M. M. Mamdani, T. Gomes, D. N. Juurlink, E. M. Macdonald and M. Yogendran (2016). "Antiepileptic use for epilepsy and nonepilepsy disorders: A population-based study (1998-2013)." Neurology 86(10): 939-946.

Lossius, M. I., K. Stavem and L. Gjerstad (1999). "Predictors for recurrence of epileptic seizures in a general epilepsy population." Seizure 8(8): 476-479.

MacDonald, S. C., M. A. Hernan, T. F. McElrath and S. Hernandez-Diaz (2018). "Assessment of recording bias in pregnancy studies using health care databases: An application to neurologic conditions." Paediatr Perinat Epidemiol 32(3): 281-286.

Madubata, C. C., M. A. Olsen, D. L. Stwalley, D. H. Gutmann and K. J. Johnson (2015). "Neurofibromatosis type 1 and chronic neurological conditions in the United States: an administrative claims analysis." Genet Med 17(1): 36-42.

Marrie, R. A., S. B. Patten, H. Tremlett, C. Wolfson, S. Warren, L. W. Svenson, N. Jette, J. Fisk, C. T. i. t. Epidemiology and S. Impact of Comorbidity on Multiple (2016). "Sex differences in comorbidity at diagnosis of multiple sclerosis: A population-based study." Neurology 86(14): 1279-1286.

Martin, R. C. (2017). "Expeditions in Epidemiologic Studies of New Onset Epilepsy in Older Adults: Stake Your Claims." Epilepsy currents 17(6): 368-369.

Martin, R. C., E. Faught, J. P. Szaflarski, J. Richman, E. Funkhouser, K. Piper, L. Juarez, C. Dai and M. Pisu (2017). "What does the U.S. Medicare administrative claims database tell us about initial antiepileptic drug treatment for older adults with new-onset epilepsy?" Epilepsia 58(4): 548-557.

Mattsson, P., T. Tomson, K. Edebol Eeg-Olofsson, L. Brannstrom and G. Ringback Weitoft (2012). "Association between sociodemographic status and antiepileptic drug prescriptions in children with epilepsy." Epilepsia 53(12): 2149-2155.

Méndez-Ayala, A., D. Nariño and D. Rosselli (2015). "Burden of epilepsy in Colombia." Neuroepidemiology 44(3): 144-148.

Metcalfe, A. and N. Jette (2010). "Medical and employment-related costs of epilepsy in the USA." Expert Rev Pharmacoecon Outcomes Res 10(6): 645-647.

Miller, W. R., D. Groves, A. Knopf, J. L. Otte and R. D. Silverman (2017). "Word Adjacency Graph Modeling: Separating Signal From Noise in Big Data." West J Nurs Res 39(1): 166-185.

Mittal, M., D. L. Harrison, M. J. Miller, K. C. Farmer, D. M. Thompson and Y. T. Ng (2014). "Have antiepileptic drug prescription claims changed following the FDA suicidality warning? An evaluation in a state Medicaid program." Epilepsy Behav 34: 109-115.

Morris, N. A. and H. Kamel (2018). Cardiac Arrest Does Not Increase the Long-Term Risk of Seizures. STROKE, LIPPINCOTT WILLIAMS & WILKINS TWO COMMERCE SQ, 2001 MARKET ST, PHILADELPHIA ….

Naldi, I., C. Piccinni, B. Mostacci, J. Renzini, G. Accetta, F. Bisulli, M. Tappata, A. Piazza, P. Pagano, S. Bianchi, R. D'Alessandro, P. Tinuper, E. Poluzzi and E. Group (2016). "Prescription patterns of antiepileptic drugs in young women: development of a tool to distinguish between epilepsy and psychiatric disorders." Pharmacoepidemiol Drug Saf 25(7): 763-769.

Ney, J. P., B. Johnson, T. Knabel, K. Craft and J. Kaufman (2016). "Neurologist ambulatory care, health care utilization, and costs in a large commercial dataset." Neurology 86(4): 367-374.

Nilsson, F. M., L. V. Kessing and T. G. Bolwig (2003). "On the increased risk of developing late-onset epilepsy for patients with major affective disorder." J Affect Disord 76(1-3): 39-48.

Oh, A., D. Thurman and H. Kim (2017). "Risk factors for developing epilepsy in a U.S. commercially insured pediatric population." Neurology. Conference: 69th American Academy of Neurology Annual Meeting, AAN 88(16 Supplement 1).

Okamoto, K., M. Fukuda, I. Saito, I. Horiuchi, T. Okazawa and E. Ishii (2018). "Incidence of childhood epilepsy: A population-based study in rural Japan." Brain Dev 40(10): 904-908.

Osterhaus, J., X. Ning, A. Petrilla, B. Agatep and G. Gilmet (2016). "Emergency Department Treatment Patterns of Patients with Epilepsy in the United States from 2008-2011: A Nationally Representative Analysis using Nhamcs Data." Value in Health 19(3): A64.

Pina-Garza, J. E., G. D. Montouris, F. Vekeman, W. Y. Cheng, E. Tuttle, Y. Shen, M. S. Duh, K. Chatamra and S. M. Kymes (2017). "Healthcare costs among patients with lennox-gastaut syndrome treated with clobazam." Value in Health 20 (5): A191.

Pisu, M., J. Richman, K. Piper, R. Martin, E. Funkhouser, C. Dai, L. Juarez, J. P. Szaflarski and E. Faught (2017). "Quality of Antiepileptic Treatment Among Older Medicare Beneficiaries With Epilepsy: A Retrospective Claims Data Analysis." Med Care 55(7): 677-683.

Pisu, M., J. S. Richman, R. C. Martin, Y. Kim, P. Kratt, K. Clements, E. Funkhouser, R. Knowlton and E. Faught (2012). "Diagnostic tests and neurology care for Medicare beneficiaries with seizures: differences across racial groups." Med Care 50(8): 730-736.

Polard, E., A. Degremont, C. Rault, B. Le Nautout, E. Oger, A. Biraben and E. Nowak (2018). "Trends of sodium valproate prescriptions in children 0-14 years of age between 2010 and 2016: Study from the French National Health Insurance Databases." Fundamental and Clinical Pharmacology 32 (Supplement 1): 27.

Polard, E., E. Nowak, A. Happe, A. Biraben, E. Oger and G. S. Group (2015). "Brand name to generic substitution of antiepileptic drugs does not lead to seizure-related hospitalization: a population-based case-crossover study." Pharmacoepidemiol Drug Saf 24(11): 1161-1169.

Powell, G. A., L. J. Bonnett, C. Tudur-Smith, D. A. Hughes, P. R. Williamson and A. G. Marson (2017). "Using routinely recorded data in the UK to assess outcomes in a randomised controlled trial: The Trials of Access." Trials 18(1): 389.

Pugh, M. J., D. R. Berlowitz and L. Kazis (2007). "The impact of epilepsy on older veterans." Int Rev Neurobiol 81: 221-233.

Pugh, M. J., L. A. Copeland, J. E. Zeber, J. A. Cramer, M. E. Amuan, J. E. Cavazos and L. E. Kazis (2005). "The impact of epilepsy on health status among younger and older adults." Epilepsia 46(11): 1820-1827.

Pugh, M. J., J. Cramer, J. Knoefel, A. Charbonneau, A. Mandell, L. Kazis and D. Berlowitz (2004). "Potentially inappropriate antiepileptic drugs for elderly patients with epilepsy." J Am Geriatr Soc 52(3): 417-422.

Pugh, M. J., J. E. Knoefel, E. M. Mortensen, M. E. Amuan, D. R. Berlowitz and A. C. Van Cott (2009). "New-onset epilepsy risk factors in older veterans." J Am Geriatr Soc 57(2): 237-242.

Pugh, M. J., J. A. Orman, C. A. Jaramillo, M. C. Salinsky, B. C. Eapen, A. R. Towne, M. E. Amuan, G. Roman, S. D. McNamee, T. A. Kent, K. K. McMillan, H. Hamid and J. H. Grafman (2015). "The prevalence of epilepsy and association with traumatic brain injury in veterans of the Afghanistan and Iraq wars." J Head Trauma Rehabil 30(1): 29-37.

Pugh, M. J., A. C. Van Cott, M. Amuan, C. Baca, P. Rutecki, M. M. Zack and R. Kobau (2016). "Epilepsy Among Iraq and Afghanistan War Veterans - United States, 2002-2015." MMWR Morb Mortal Wkly Rep 65(44): 1224-1227.

Puka, K., M. L. Smith, R. Moineddin, O. C. Snead and E. Widjaja (2016). "Health resource utilization varies by comorbidities in children with epilepsy." Epilepsy Behav 57(Pt A): 151-154.

Puka, K., M. L. Smith, R. Moineddin, O. C. Snead and E. Widjaja (2016). "The influence of socioeconomic status on health resource utilization in pediatric epilepsy in a universal health insurance system." Epilepsia 57(3): 455-463.

Putignano, D., A. Clavenna, R. Campi, A. Bortolotti, I. Fortino, L. Merlino, A. Vignoli, M. P. Canevini and M. Bonati (2017). "Antiepileptic drug use in Italian children over a decade." Eur J Clin Pharmacol 73(2): 241-248.

Quinlan, S. C., W. Y. Cheng, L. Ishihara, M. C. Irizarry, C. N. Holick and M. S. Duh (2016). "Development and validation of an algorithm for identifying urinary retention in a cohort of patients with epilepsy in a large US administrative claims database." Pharmacoepidemiol Drug Saf 25(4): 413-421.

Rehman, R., P. R. Kelly, A. M. Husain and T. T. Tran (2015). "Characteristics of Veterans diagnosed with seizures within Veterans Health Administration." J Rehabil Res Dev 52(7): 751-762.

Roberts, J., M. Liu, C. Smith, H. Quan, S. Wiebe and N. Jette (2014). "Validation of an epilepsy specific risk adjustment index for mortality research." Epilepsy Currents 1): 275.

Rohde, N. N., C. B. Baca, A. C. Van Cott, K. L. Parko, M. E. Amuan and M. J. Pugh (2015). "Antiepileptic drug prescribing patterns in Iraq and Afghanistan war veterans with epilepsy." Epilepsy Behav 46: 133-139.

Rothmann, L. J., M. S. Lubbe, J. H. P. Serfontein, J. J. Gerber and M. Malik (2016). "Prevalence of chronic diseases in private healthcare sector of South Africa: A threat to public health." Tropical Journal of Pharmaceutical Research 15(6): 1327-1334.

Sanders, L. M., M. Tan, M. J. Cook and W. J. D'Souza (2016). "Increased risk of stroke in patients with epilepsy: Analysis of the tasmanian epilepsy register." European Stroke Journal 1 (1 Supplement 1): 179-180.

Sapkota, S., R. Kobau, D. M. Pastula and M. M. Zack (2018). "People with epilepsy are diagnosed most often with unspecified epilepsy, followed by focal epilepsy, generalized convulsive epilepsy, and generalized nonconvulsive epilepsy—US MarketScan data, 2010–2015." Epilepsy & Behavior 79: 244-246.

Schiariti, V., K. Farrell, J. S. Hoube and S. Lisonkova (2009). "Period prevalence of epilepsy in children in BC: a population-based study." Can J Neurol Sci 36(1): 36-41.

Schiltz, N. K., S. M. Koroukian, S. D. Lhatoo and K. Kaiboriboon (2013). "Temporal trends in pre-surgical evaluations and epilepsy surgery in the U.S. from 1998 to 2009." Epilepsy Res 103(2-3): 270-278.

Shcherbakova, N., K. Rascati, C. Brown, K. Lawson, S. Novak, K. M. Richards and L. Yoder (2014). "Factors associated with seizure recurrence in epilepsy patients treated with antiepileptic monotherapy: A retrospective observational cohort study using US administrative insurance claims." CNS Drugs 28(11): 1047-1058.

Singh, G., A. Mithal and B. Lingala (2016). "Emergency department visits for uncontrolled epilepsy in us: A national perspective." Neurology. Conference: 68th American Academy of Neurology Annual Meeting, AAN 86(16 SUPPL. 1).

Sirven, J. I., P. O. Shafer, L. Kalilani, I. Wild, J. Fishman and S. Owens (2018). "Current state of the union of epilepsy care in the United States: Antiepileptic drugs - An introduction to the Connectors Project." Epilepsy Behav 80: 98-103.

St Germaine-Smith, C., M. Liu, H. Quan, S. Wiebe and N. Jette (2011). "Development of an epilepsy-specific risk adjustment comorbidity index." Epilepsia 52(12): 2161-2167.

St Germaine-Smith, C., A. Metcalfe, T. Pringsheim, J. I. Roberts, C. A. Beck, B. R. Hemmelgarn, J. McChesney, H. Quan and N. Jette (2012). "Recommendations for optimal ICD codes to study neurologic conditions: a systematic review." Neurology 79(10): 1049-1055.

St. Germaine-Smith, C., A. Metcalfe, H. Quan, B. Hemmelgarn, T. Pringsheim, C. Beck and N. Jette (2011). "Validity of administrative data for identifying epilepsy - A systematic review." American Journal of Epidemiology 11): S208.

Strzelczyk, A., S. Ansorge, J. Hapfelmeier, V. Bonthapally, M. H. Erder and F. Rosenow (2017). "Costs, length of stay, and mortality of super-refractory status epilepticus: A population-based study from Germany." Epilepsia 58(9): 1533-1541.

Strzelczyk, A., C. Griebel, W. Lux, F. Rosenow and J. P. Reese (2017). "The Burden of Severely Drug-Refractory Epilepsy: A Comparative Longitudinal Evaluation of Mortality, Morbidity, Resource Use, and Cost Using German Health Insurance Data." Front Neurol 8: 712.

Strzelczyk, A., C. Griebel, W. Lux, F. Rosenow and J. P. Reese (2017). "Long-term course and mortality in patients with severely drug-refractory epilepsy: A population based study on health insurance data." Zeitschrift fur Epileptologie 30 (1 Supplement 1): S18.

Tai, Y. M. and C. S. Gau (2013). "Secular trends of ADHD and developmental delay in Taiwanese children with epilepsy: a 13-year National Health Insurance data retrospective analysis." J Formos Med Assoc 112(5): 293-295.

Tang, D. H., D. C. Malone, T. L. Warholak, J. Chong, E. P. Armstrong, M. K. Slack, C. H. Hsu and D. M. Labiner (2015). "Healthcare resource burden of newly diagnosed epilepsy in the US low-income elderly population." European Geriatric Medicine 6(3): 251-256.

Tang, D. H., D. C. Malone, T. L. Warholak, J. Chong, E. P. Armstrong, M. K. Slack, C. H. Hsu and D. M. Labiner (2015). "Prevalence and Incidence of Epilepsy in an Elderly and Low-Income Population in the United States." J Clin Neurol 11(3): 252-261.

Tatum, W. O., M. E. Langston and E. K. Acton (2016). "Fibromyalgia and seizures." Epileptic Disord 18(2): 148-154.

Thurman, D. J., E. Beghi, C. E. Begley, A. T. Berg, J. R. Buchhalter, D. Ding, D. C. Hesdorffer, W. A. Hauser, L. Kazis and R. Kobau (2011). "Standards for epidemiologic studies and surveillance of epilepsy." Epilepsia 52: 2-26.

Tsai, J. D., C. L. Lin, C. C. Lin, F. C. Sung and K. H. Lue (2014). "Risk of epilepsy in patients with systemic lupus erythematosus - a retrospective cohort study." Neuropsychiatr Dis Treat 10: 1635-1643.

Tseng, C. H., W. S. Huang, C. H. Muo and C. H. Kao (2014). "Increased risk of epilepsy among patients diagnosed with chronic osteomyelitis." Epilepsy Res 108(8): 1427-1434.

Vaccarino, A. L., M. Dharsee, S. Strother, D. Aldridge, S. R. Arnott, B. Behan, C. Dafnas, F. Dong, K. Edgecombe, R. El-Badrawi, K. El-Emam, T. Gee, S. G. Evans, M. Javadi, F. Jeanson, S. Lefaivre, K. Lutz, F. C. MacPhee, J. Mikkelsen, T. Mikkelsen, N. Mirotchnick, T. Schmah, C. M. Studzinski, D. T. Stuss, E. Theriault and K. R. Evans (2018). "Brain-CODE: A Secure Neuroinformatics Platform for Management, Federation, Sharing and Analysis of Multi-Dimensional Neuroscience Data." Front Neuroinform 12: 28.

Van Cott, A., R. Adler, K. Tortorice, M. J. Pugh, B. Gidal, D. Dong and F. Cunningham (2018). "Fracture risk and bone health in veterans treated for epilepsy." Neurology. Conference: 70th Annual Meeting of the American Academy of Neurology, AAN 90(15 Supplement 1).

Velez, F. F. and J. S. Korsnes (2016). "Antiepileptic drug pill burden during the first year following monotherapy initiation associated with increased epilepsy costs in The United States." Value in Health 19 (3): A62.

Wennberg, R., C. Hiploylee, P. Tai and C. H. Tator (2018). "Is Concussion a Risk Factor for Epilepsy?" Can J Neurol Sci 45(3): 275-282.

Wiebe, S. and N. Jette (2009). "Oh data, where are ye? Using administrative databases in epilepsy research." Can J Neurol Sci 36(1): 2-3.

Williamson, T., R. C. Miyagishima, J. D. Derochie and N. Drummond (2017). "Manual review of electronic medical records as a reference standard for case definition development: a validation study." CMAJ Open 5(4): E830-E833.

Wilner, A., B. Sharma, A. Soucy, A. Thompson and A. Krueger (2014). "Common comorbidities in women and men with epilepsy and the relationship between number of comorbidities and health plan paid costs in 2010." Epilepsy & Behavior 32: 15-20.

Wilner, A. N., B. K. Sharma, A. Soucy and A. Krueger (2012). "Health plan paid cost of epilepsy in 2009 in the US." Epilepsy & Behavior 25(3): 412-416.

Wilner, A. N., B. K. Sharma, A. R. Thompson and A. Krueger (2016). "Analgesic opioid use in a health-insured epilepsy population during 2012." Epilepsy Behav 57(Pt A): 126-132.

Wu, C. S., S. C. Wang, I. J. Yeh and S. K. Liu (2016). "Comparative risk of seizure with use of first- and second-generation antipsychotics in patients with schizophrenia and mood disorders." J Clin Psychiatry 77(5): e573-579.

Zarrelli, M. M., E. Beghi, W. A. Rocca and W. A. Hauser (1999). "Incidence of epileptic syndromes in Rochester, Minnesota: 1980–1984." Epilepsia 40(12): 1708-1714.

Zeber, J. E., L. A. Copeland and M. J. Pugh (2010). "Variation in antiepileptic drug adherence among older patients with new-onset epilepsy." Ann Pharmacother 44(12): 1896-1904.

Zelano, J., S. Asberg, P. Redfors and E. Kumlien (2016). "Poststroke epilepsy in Sweden: A population-based registry study." Epilepsia 57 (Supplement 2): 41-42.

Zelano, J., P. Redfors, S. Åsberg and E. Kumlien (2016). "Association between poststroke epilepsy and death: A nationwide cohort study." European stroke journal 1(4): 272-278.

Zhang, T., H. Tremlett, F. Zhu, E. Kingwell, J. D. Fisk, V. Bhan, T. Campbell, K. Stadnyk, R. Carruthers, C. Wolfson, S. Warren, R. A. Marrie, C. T. i. t. Epidemiology and S. Impact of Comorbidity on Multiple (2018). "Effects of physical comorbidities on disability progression in multiple sclerosis." Neurology 90(5): e419-e427.

Appendix S6: QUADAS-2 quality assessment items

A – Kappa statistic agreement (reviewers GKM and KB)

Across all studies overall, there was independent agreement (i.e. without the need for consensus discussion) on 346 quality assessment decisions using the QUADAS-2 form (Appendix S3). Mutual consensus was reached (i.e. without the need for third reviewer adjunction) for the remaining 118 quality assessment decisions (mean Kappa 0.39, 95% CI 0.27–0.51.

Interpretation[1]:

Kappa < 0: No agreement

Kappa between 0.00 and 0.20: Slight agreement

Kappa between 0.21 and 0.40: Fair agreement

Kappa between 0.41 and 0.60: Moderate agreement

Kappa between 0.61 and 0.80: Substantial agreement

Kappa between 0.81 and 1.00: Almost perfect agreement."

1. Landis, J.R.; Koch, G.G. (1977). The measurement of observer agreement for categorical data. Biometrics. 33 (1): 159-174. https://doi.org/10.2307%2F2529310

Kappa statistic interrater agreement data on the 16 QUADAS-2 quality assessment items

First author and year	Kappa results
Aaberg 2017	Number of observed agreements: 13 (81.25% of the observations) Number of agreements expected by chance: 8.7 (54.30% of the observations) Kappa= 0.590 SE of kappa = 0.160 95% confidence interval: From 0.277 to 0.903
Bellini 2017	Number of observed agreements: 9 (60.00% of the observations) Number of agreements expected by chance: 7.3 (48.89% of the observations) Kappa= 0.217 SE of kappa = 0.096 95% confidence interval: From 0.028 to 0.406
Chen 2012	Number of observed agreements: 9 (56.25% of the observations) Number of agreements expected by chance: 8.7 (54.30% of the observations) Kappa= 0.043 SE of kappa = 0.208 95% confidence interval: From -0.365 to 0.451
Christensen 2007	Number of observed agreements: 15 (93.75% of the observations) Number of agreements expected by chance: 15.0 (93.75% of the observations) Kappa= 0.000 SE of kappa = 0.000 95% confidence interval: From 0.000 to 0.000
Coulter 1989	Number of observed agreements: 8 (50.00% of the observations)

	Number of agreements expected by chance: 5.5 (34.38% of the observations) Kappa= 0.238 SE of kappa = 0.139 95% confidence interval: From -0.035 to 0.511
De Jesus-Alvelo 2013	Number of observed agreements: 7 (43.75% of the observations) Number of agreements expected by chance: 7.5 (46.88% of the observations) Kappa= -0.059 SE of kappa = 0.159 95% confidence interval: From -0.371 to 0.253
Engeland 2009	Number of observed agreements: 11 (68.75% of the observations) Number of agreements expected by chance: 9.3 (57.81% of the observations) Kappa= 0.259 SE of kappa = 0.243 95% confidence interval: From -0.217 to 0.736
Foebel 2013	Number of observed agreements: 9 (56.25% of the observations) Number of agreements expected by chance: 6.1 (37.89% of the observations) Kappa= 0.296 SE of kappa = 0.148 95% confidence interval: From 0.006 to 0.585
Fonferko-Shadrach 2017	Number of observed agreements: 12 (75.00% of the observations) Number of agreements expected by chance: 10.5 (65.63% of the observations) Kappa= 0.273 SE of kappa = 0.103 95% confidence interval: From 0.070 to 0.475
Franchi 2013	Number of observed agreements: 11 (68.75% of the observations) Number of agreements expected by chance: 11.4 (71.09% of the observations) Kappa= -0.081 SE of kappa = 0.069 95% confidence interval: From -0.216 to 0.054
Frost 2000	Number of observed agreements: 11 (68.75% of the observations) Number of agreements expected by chance: 9.8 (60.94% of the observations) Kappa= 0.200 SE of kappa = 0.101 95% confidence interval: From 0.002 to 0.398
Holden 2005	Number of observed agreements: 14 (87.50% of the observations) Number of agreements expected by chance: 13.1 (82.03% of the observations) Kappa= 0.304 SE of kappa = 0.151 95% confidence interval: From 0.008 to 0.600
Jette 2010	Number of observed agreements: 16 (100% of the observations) Number of agreements expected by chance: 8.0 (50.00% of the observations) Kappa= 1.000 SE of kappa = 0.000 95% confidence interval: From 1.000 to 1.000
Lee 2016	Number of observed agreements: 12 (75.00% of the observations) Number of agreements expected by chance: 10.5 (65.63% of the observations) Kappa= 0.273 SE of kappa = 0.103 95% confidence interval: From 0.070 to 0.475
Marrie 2013	Number of observed agreements: 6 (37.50% of the observations) Number of agreements expected by chance: 6.5 (40.63% of the observations) Kappa= -0.053 SE of kappa = 0.209 95% confidence interval: From -0.462 to 0.356
Meeraus 2013	Number of observed agreements: 16 (100% of the observations) Number of agreements expected by chance: 12.5 (78.13% of the observations) Kappa= 1.000 SE of kappa = 0.000 95% confidence interval: From 1.000 to 1.000
Moura 2017	Number of observed agreements: 12 (75.00% of the observations) Number of agreements expected by chance: 10.5 (65.63% of the observations) Kappa= 0.273 SE of kappa = 0.103 95% confidence interval: From 0.070 to 0.475
Parko 2009	Number of observed agreements: 12 (75.00% of the observations)

	Number of agreements expected by chance: 10.5 (65.63% of the observations) Kappa= 0.273 SE of kappa = 0.103 95% confidence interval: From 0.070 to 0.475
Pickrell 2015	Number of observed agreements: 16 (100% of the observations) Number of agreements expected by chance: 12.5 (78.13% of the observations) Kappa= 1.000 SE of kappa = 0.000 95% confidence interval: From 1.000 to 1.000
Pina-Garza 2015	Number of observed agreements: 9 (56.25% of the observations) Number of agreements expected by chance: 6.1 (38.28% of the observations) Kappa= 0.291 SE of kappa = 0.165 95% confidence interval: From -0.033 to 0.615
Pugh 2008	Number of observed agreements: 11 (68.75% of the observations) Number of agreements expected by chance: 9.8 (60.94% of the observations) Kappa= 0.200 SE of kappa = 0.114 95% confidence interval: From -0.023 to 0.423
Pugh 2015	Number of observed agreements: 12 (75.00% of the observations) Number of agreements expected by chance: 7.8 (48.44% of the observations) Kappa= 0.515 SE of kappa = 0.191 95% confidence interval: From 0.140 to 0.890
Rehman 2016	Number of observed agreements: 14 (87.50% of the observations) Number of agreements expected by chance: 10.8 (67.19% of the observations) Kappa= 0.619 SE of kappa = 0.228 95% confidence interval: From 0.172 to 1.000
Reid 2012	Number of observed agreements: 15 (93.75% of the observations) Number of agreements expected by chance: 15.0 (93.75% of the observations) Kappa= 0.000 SE of kappa = 0.000 95% confidence interval: From 0.000 to 0.000
Shackleton 1997	Number of observed agreements: 12 (75.00% of the observations) Number of agreements expected by chance: 9.0 (56.25% of the observations) Kappa= 0.429 SE of kappa = 0.232 95% confidence interval: From -0.026 to 0.883
Syvertsen 2015	Number of observed agreements: 16 (100% of the observations) Number of agreements expected by chance: 12.5 (78.13% of the observations) Kappa= 1.000 SE of kappa = 0.000 95% confidence interval: From 1.000 to 1.000
Tan 2015	Number of observed agreements: 16 (100% of the observations) Number of agreements expected by chance: 8.0 (50.00% of the observations) Kappa= 1.000 SE of kappa = 0.000 95% confidence interval: From 1.000 to 1.000
Tu 2014	Number of observed agreements: 14 (87.50% of the observations) Number of agreements expected by chance: 10.8 (67.19% of the observations) Kappa= 0.619 SE of kappa = 0.228 95% confidence interval: From 0.172 to 1.000
Wassenaar 2018	Number of observed agreements: 14 (87.50% of the observations) Number of agreements expected by chance: 10.8 (67.19% of the observations) Kappa= 0.619 SE of kappa = 0.228 95% confidence interval: From 0.172 to 1.000
Williamson 2014	Number of observed agreements: 10 (62.50% of the observations) Number of agreements expected by chance: 7.3 (45.70% of the observations) Kappa= 0.309 SE of kappa = 0.195 95% confidence interval: From -0.072 to 0.691

B

Final QUADAS-2 quality assessment results

	Risk of bias				Applicability		
First author and year	Patient selection	Administrative database (Index test)	Reference standard	Flow & timing	Patient selection	Administrative database (Index test)	Reference standard
Aaberg 2017	✓	✓	?	×	✓	×	✓
Bellini 2017	✓	✓	?	✓	✓	✓	✓
Chen 2012	?	✓	✓	✓	×	✓	✓
Christensen 2007	✓	✓	✓	✓	✓	✓	✓
Coulter 1989	✓	×	×	?	✓	✓	✓
De Jesus-Alvelo 2013	?	✓	?	✓	✓	?	✓
Engeland 2009	✓	×	×	✓	✓	✓	✓
Foebel 2013	✓	✓	×	×	✓	✓	✓
Fonferko-Shadrach 2017	✓	✓	?	×	✓	✓	✓
Franchi 2013	✓	✓	×	✓	✓	✓	✓
Frost 2000	✓	✓	×	×	✓	✓	✓
Holden 2005	✓	✓	?	✓	✓	✓	✓
Jetté 2010	✓	✓	✓	✓	✓	✓	✓
Lee 2016	×	✓	?	✓	✓	✓	✓
Marrie 2013	×	✓	?	×	✓	✓	✓
Meeraus 2013	✓	✓	?	✓	✓	✓	✓
Moura 2017	✓	✓	?	✓	✓	✓	✓
Parko 2009	✓	✓	?	✓	✓	✓	✓
Pickrell 2015	✓	✓	✓	✓	?	✓	✓
Pina-Garza 2015	?	✓	×	?	✓	✓	✓
Pugh 2008	✓	✓	?	?	✓	✓	✓
Pugh 2015	✓	✓	?	?	✓	✓	✓
Rehman 2016	✓	✓	?	✓	✓	✓	✓
Reid 2012	✓	✓	✓	✓	✓	✓	✓
Shackleton 1997	✓	✓	×	×	✓	✓	✓
Syvertsen 2015	✓	✓	?	✓	✓	✓	✓
Tan 2015	✓	✓	✓	✓	✓	✓	✓
Tu 2014	✓	✓	?	✓	✓	✓	✓
Wassenaar 2018	×	✓	✓	✓	✓	?	✓
Williamson 2014	×	✓	✓	×	×	×	✓

Key: ✓ = low risk, × = high risk, ? = unclear risk

Appendix S7: Study algorithm lookup table

Study	Country	Age	Overall quality score	Gold standard	Administrative setting	Coding system	Algorithm	Outcome	Administrative sample size	PPV	PPV lower CI	PPV upper CI	Reference sample size	Sn	Sn lower CI	Sn upper CI	NPV sample size	NPV	NPV lower CI	NPV upper CI	Sp sample size	Sp	Sp lower CI	Sp upper CI	TP	FP	TN	FN
Jetté 2010	Canada	All ages	LOW	Medical records	A&E	ICD10	ICD10: G40-41	Epilepsy	207	100	93.2	100					207	89.50	83.10	93.60								
Jetté 2010	Canada	Unclear	LOW	Medical records	Inpatients, A&E	ICD9	ICD9: 345	Epilepsy	81	100	93.5	100					81	95.80	79.80	99.30								
Jetté 2010	Canada	Unclear	LOW	Medical records	Inpatients, A&E	ICD10	ICD10: G40-41	Epilepsy	77	100	91.2	100					77	70.00	52.10	83.30								
Fonferko-Shadrach 2017	UK	≤16 years	HIGH	Medical records	Primary care	Read Codes, AED	Read Code: F25, 1O30, 667B., SC200 + AEDx2 0.5 yr	Epilepsy	38	100	91.0	100	48	79.0	65.0	90.0	55	81.82	72.16	88.65	45	100	92.00	100	38	0	45	10
Wassenaar 2018	Holland	18-75 years	HIGH	Medical records	Physician claims (neurologist)	ICD9, AED	ICD9: 345 + ≥1AED (1P 3yr)	Epilepsy	269	99.0	98.0	100																
Jetté 2010	Canada	All ages	LOW	Medical records	Inpatients, A&E	ICD9	ICD9: 345	Epilepsy	486	98.9	96.1	99.7					486	97.40	94.80	98.80								
Jetté 2010	Canada	All ages	LOW	Medical records	A&E	ICD9	ICD9: 345	Epilepsy	277	98.9	94.2	99.8					277	97.00	93.20	98.70								
Jetté 2010	Canada	All ages	LOW	Medical records	Inpatients, A&E	ICD10	ICD10: G40-41	Epilepsy	454	98.6	95.1	99.6					454	94.00	90.50	96.30								
Jetté 2010	Canada	Unclear	LOW	Medical records	Inpatients, A&E	ICD9	ICD9: 345	Epilepsy	405	98.4	94.4	99.6					405	97.60	94.90	98.90								
Jetté 2010	Canada	Unclear	LOW	Medical records	Inpatients, A&E	ICD10	ICD10: G40-41	Epilepsy	377	98.1	93.3	100					377	97.00	94.00	98.60								
Fonferko-Shadrach 2017	UK	All ages	HIGH	Medical records	Primary care	Read Codes, AED	Read Code: F25, 1O30, 667B., SC200 + AEDx2 0.5 yr	Epilepsy	124	98.0	94.0	100	145	84.0	77.0	90.0	164	85.98	80.81	89.92	143	99.00	95.00	100	122	2	141	23
Fonferko-Shadrach 2017	UK	>16 years	HIGH	Medical records	Primary care	Read Codes, AED	Read Code: F25, 1O30, 667B., SC200 + AEDx2 0.5 yr	Epilepsy	86	98.0	92.0	100	97	87.0	78.0	93.0	109	88.07	81.65	92.46	98	98.00	93.00	100	84	2	96	13
Fonferko-Shadrach 2017	UK	>16 years	HIGH	Medical records	Primary care	Read Codes	Read Code: F25, 1O30, 667B., SC200	Epilepsy	87	98.0	92.0	100	97	88.0	80.0	93.0	108	88.89	82.48	93.15	98	98.00	93.00	100	85	2	96	12
Fonferko-Shadrach 2017	UK	≤16 years	HIGH	Medical records	Primary care	AED	AEDx2 0.5 yr	Epilepsy	43	98.0	94.0	100	48	88.0	75.0	95.0	50	88.00	77.60	93.95	45	98.00	88.00	100	42	1	44	6
Jetté 2010	Canada	All ages	LOW	Medical records	Inpatients	ICD10	ICD10: G40-41	Epilepsy	247	97.8	92.3	99.4					247	98.50	94.70	99.60								
Jetté 2010	Canada	All ages	LOW	Medical records	Inpatients	ICD9	ICD9: 345	Epilepsy	209	97.7	92.1	99.4					209	99.10	94.80	99.80								
Jetté 2010	Canada	Unclear	LOW	Medical records	Inpatients, A&E	ICD9	ICD9: 345, 780.3	Epilepsy	81	96.5	88.1	99.0																
Fonferko-Shadrach 2017	UK	All ages	HIGH	Medical records	Primary care	Read Codes	Read Code: F25, 1O30, 667B., SC200	Epilepsy	130	96.0	91.0	99.0	145	86.0	80.0	91.0	158	87.34	82.10	91.21	143	97.00	92.00	99.00	125	5	138	20
Rehman 2016	US	Unclear	UNCLEAR	Medical records	Inpatients, outpatients, pharmacy	ICD9, AED	ICD9: 345, 780.39 (3 yr) + seizure + ≥1AED (≥30 days)	Epilepsy	675	96.0	94.5	97.5													648	27		
Wassenaar 2018	Holland	18-75 years	HIGH	Medical records	Physician claims (neurologist)	ICD9, AED	ICD9: 345 (2P 3yr) or + ≥1AED (1P 3yr)	Epilepsy	293	95.0	92.0	97.0																
Pugh 2008	US	≥66 years	UNCLEAR	Medical records	Inpatients, outpatients, pharmacy	ICD9, AED	ICD9: 345, 780.39 + ≥1AED	Epilepsy	126	94.0	89.9	98.2													119	7		

Fonferko-Shadrach 2017	UK	≤16 years	HIGH	Medical records	Primary care	Read Codes	Read Code: F25, 1O30, 667B., SC200	Epilepsy	43	93.0	81.0	99.0	48	83.0	70.0	93.0	50	84.00	73.51	90.85	45	93.00	82.00	99.00	40	3	42	8
Shackleton 1997	Holland	0-75 years	HIGH	GP	Pharmacy	AED	>1AED (> 1 yr)	Epilepsy	106	93.0	88.0	98.0													98	8		
Reid 2012	Canada	≥18 years	LOW	Medical records	Physician claims	ICD9, ICD10	ICD9: 345; ICD10: G40-41 (2P 1yr)	Epilepsy	706	92.1	90.1	94.1	706	92.1	90.1	94.1	827	93.20	91.50	94.90	827	93.20	91.50	94.90	650	56	771	56
Reid 2012	Canada	≥18 years	LOW	Medical records	Inpatients, physician claims	ICD9, ICD10	ICD9: 345; ICD10: G40-41 (2P or 1H 1yr)	Epilepsy	709	92.1	90.1	94.1	706	92.5	90.6	94.4	824	93.60	91.90	95.20	827	93.20	91.50	94.90	653	56	771	53
Meeraus 2013	UK	0-14 years	UNCLEAR	Medical records	Primary care (incl. pharmacy)	Read Codes, AED	Read Code diagnosis or symptoms or AED (0.3 yr)	Epilepsy	183	92.0	88.0	96.0													168	15		
Reid 2012	Canada	≥18 years	LOW	Medical records	Physician claims	ICD9, ICD10	ICD9: 345; ICD10: G40-41 (2P 2yr)	Epilepsy	712	91.9	89.8	93.9	706	92.6	90.7	94.6	821	93.70	92.00	95.30	827	93.00	91.30	94.70	654	58	769	52
Reid 2012	Canada	≥18 years	LOW	Medical records	Inpatients, physician claims	ICD9, ICD10	ICD9: 345; ICD10: G40-41 (2P or 1H 2yr)	Epilepsy	715	91.9	89.9	93.9	706	93.1	91.2	94.9	818	94.00	92.40	95.60	827	93.00	91.30	94.70	657	58	769	49
Moura 2017	US	≥18 years	UNCLEAR	Medical records	Regional professional claims	ICD9, AED	ICD9: 345, 333.2, 779, 780.3, 780.39, 780.2, 780.31 + 311, 300.4, 300, 300.02 + ≥1AED	Epilepsy	875	91.4	88.9	93.4	875	66.5	63.3	69.6	502	60.40	56.80	63.90	502	89.00	86.00	91.60				
Reid 2012	Canada	≥18 years	LOW	Medical records	Physician claims	ICD9, ICD10	ICD9: 345, 780.3; ICD10: G40-41 (2P 1yr)	Epilepsy	735	90.9	88.8	93.0	712	93.8	92.1	95.6	798	94.50	92.90	96.10	821	91.80	90.00	93.70	668	67	754	44
Reid 2012	Canada	≥18 years	LOW	Medical records	Physician claims	ICD9, ICD10	ICD9: 345, 780.3; ICD10: G40-41 (2P 2yr)	Epilepsy	740	90.8	88.7	92.9	712	94.4	92.7	96.1	793	95.00	93.40	96.50	821	91.70	89.80	93.60	672	68	753	40
Reid 2012	Canada	≥18 years	LOW	Medical records	Inpatients, physician claims	ICD9, ICD10	ICD9: 345, 780.3; ICD10: G40-41 (2P or 1H 1yr)	Epilepsy	742	90.6	88.5	92.7	712	94.4	92.7	96.1	791	94.90	93.40	96.50	821	91.50	89.60	93.40	672	70	751	40
Reid 2012	Canada	≥18 years	LOW	Medical records	Inpatients, physician claims	ICD9, ICD10	ICD9: 345, 780.3; ICD10: G40-41 (2P or 1H 2yr)	Epilepsy	747	90.5	88.4	92.6	712	94.9	93.3	96.6	786	95.40	94.00	96.90	821	91.40	89.40	93.30	676	71	750	36
Parko 2009	US	All ages	UNCLEAR	Medical records	Inpatients, outpatients	ICD9	ICD9: 345, 780.3	Epilepsy or seizures	1277	90.0	88.4	91.7																
Tan 2015	Australia	≥16 years	LOW	Medical records	Inpatients, A&E	ICD10	ICD10: G41	Epilepsy	10	90.0	55.5	99.7	158	5.7	2.6	10.5		99.70	99.60	99.70	42602	100	100	100	9	1		
Reid 2012	Canada	≥18 years	LOW	Medical records	Physician claims	ICD9, ICD10	ICD9: 345, 780.3; ICD10: G40-41 (2P 1yr)	Epilepsy	315	89.5	86.1	92.9	310	91.0	87.8	94.2	405	93.10	90.60	95.60	410	92.00	89.30	94.60	282	33	377	28
Reid 2012	Canada	≥18 years	LOW	Medical records	Inpatients, physician claims	ICD9, ICD10	ICD9: 345; ICD10: G40-41 (2P or 1H 1yr)	Epilepsy	291	89.4	85.8	92.9	298	87.3	83.5	91.0	429	91.10	88.50	93.80	422	92.70	90.20	95.10	260	31	391	38
Reid 2012	Canada	≥18 years	LOW	Medical records	Physician claims	ICD9, ICD10	ICD9: 345, 780.3; ICD10: G40-41 (2P 2yr)	Epilepsy	320	89.4	86.0	92.8	310	92.3	89.3	95.2	400	94.00	91.70	96.30	410	91.70	89.00	94.40	286	34	376	24
Reid 2012	Canada	≥18 years	LOW	Medical records	Physician claims	ICD9, ICD10	ICD9: 345; ICD10: G40-41 (2P 1yr)	Epilepsy	288	89.2	85.7	92.8	298	86.2	82.3	90.2	432	90.50	87.80	93.30	422	92.70	90.20	95.10	257	31	391	41
Reid 2012	Canada	≥18 years	LOW	Medical records	Inpatients, physician claims	ICD9, ICD10	ICD9: 345; ICD10: G40-41 (2P or 1H 2yr)	Epilepsy	297	89.2	85.7	92.8	298	88.9	85.4	92.5	423	92.20	89.60	94.80	422	92.40	89.90	94.90	265	32	390	33
Reid 2012	Canada	≥18 years	LOW	Medical records	Physician claims	ICD9, ICD10	ICD9: 345; ICD10: G40-41 (2P 2yr)	Epilepsy	294	89.1	85.6	92.7	298	87.9	84.2	91.6	426	91.60	88.90	94.20	422	92.40	89.90	94.90	262	32	390	36
Reid 2012	Canada	≥18 years	LOW	Medical records	Inpatients, physician claims	ICD9, ICD10	ICD9: 345, 780.3; ICD10: G40-41 (2P or 1H 1yr)	Epilepsy	319	89.0	85.6	92.5	310	91.6	88.5	94.7	401	93.50	91.10	95.90	410	91.50	88.80	94.20	284	35	375	26

Reid 2012	Canada	≥18 years	LOW	Medical records	Inpatients, physician claims	ICD9, ICD10	ICD9: 345, 780.3; ICD10: G40-41 (2P or 1H 2yr)	Epilepsy	324		88.9	85.5	92.3	310	92.9	90.0	95.8	396	94.40	92.20	96.70	410	91.20	88.50	94.00	288	36	374	22
Aaberg 2017	Norway	3-13 years	HIGH	Medical records &/or parent telephone interview	Inpatients, outpatients, private specialist claims	ICD10	ICD10: G40 (≥2R)	Epilepsy	838		88.0	85.0	90.0													737	101		
De Jesus-Alvelo 2013	US	Unclear	UNCLEAR	Medical records	Inpatients	ICD9	ICD9: 345, 780.3	Epilepsy or seizures	241		88.0	84.0	92.0													212	29		
Reid 2012	Canada	≥18 years	LOW	Medical records	Physician claims	ICD9, ICD10	ICD9: 345; ICD10: G40-41 (1P 2yr)	Epilepsy	786		87.5	85.2	89.8	706	97.5	96.3	98.6	747	97.60	96.50	98.70	827	88.20	86.00	90.40	688	98	729	18
Reid 2012	Canada	≥18 years	LOW	Medical records	Inpatients, physician claims	ICD9, ICD10	ICD9: 345; ICD10: G40-41 (1P or 1H 2yr)	Epilepsy	786		87.5	85.2	89.8	706	97.5	96.3	98.6	747	97.60	96.50	98.70	827	88.20	86.00	90.40	688	98	729	18
Moura 2017	US	≥18 years	UNCLEAR	Medical records	Regional professional claims	ICD9, AED, Procedure	ICD9: 345, 333.2, 779, 780.3, 780.39, 780.2, 780.31 + EEG + ≥1AED	Epilepsy	875		86.9	83.8	89.7	875	52.5	49.1	55.8	502	51.00	47.60	54.40	502	86.30	82.90	89.10				
Moura 2017	US	≥18 years	UNCLEAR	Medical records	Regional professional claims	ICD9, AED	ICD9: 345 + ≥1AED (1O)	Epilepsy	875		86.8	83.8	89.4	875	58.5	55.2	61.8	502	53.90	50.30	57.40	502	84.50	81.00	87.50				
Reid 2012	Canada	≥18 years	LOW	Medical records	Physician claims	ICD9, ICD10	ICD9: 345, 780.3; ICD10: G40-41 (1P 2yr)	Epilepsy	354		86.4	82.9	90.0	310	98.7	97.5	100.0	366	98.90	97.80	100	410	88.30	85.20	91.40	306	48	362	4
Reid 2012	Canada	≥18 years	LOW	Medical records	Physician claims	ICD9, ICD10	ICD9: 345, 780.3; ICD10: G40-41 (1P 2yr)	Epilepsy	811		86.4	84.1	88.8	712	98.5	97.6	99.4	722	98.50	97.60	99.40	821	86.60	84.30	88.90	701	110	711	11
Reid 2012	Canada	≥18 years	LOW	Medical records	Inpatients, physician claims	ICD9, ICD10	ICD9: 345, 780.3; ICD10: G40-41 (1P or 1H 2yr)	Epilepsy	814		86.2	83.9	88.6	712	98.6	97.7	99.5	719	98.60	97.80	99.50	821	86.40	84.00	88.70	702	112	709	10
Reid 2012	Canada	≥18 years	LOW	Medical records	Inpatients, physician claims	ICD9, ICD10	ICD9: 345, 780.3; ICD10: G40-41 (1P or 1H 2yr)	Epilepsy	356		86.0	82.4	89.6	310	98.7	97.5	100.0	364	98.90	97.80	100	410	87.80	84.60	91.00	306	50	360	4
Reid 2012	Canada	≥18 years	LOW	Medical records	Inpatients, A&E, physician claims	ICD9, ICD10	ICD9: 345, 780.3; ICD10: G40-41 (2P or 1H or 1ER 1yr)	Epilepsy	350		85.7	82.1	89.4	310	96.8	94.8	98.7	370	97.30	95.60	99.00	410	87.80	84.60	91.00	300	50	360	10
Reid 2012	Canada	≥18 years	LOW	Medical records	Inpatients, A&E, physician claims	ICD9, ICD10	ICD9: 345, 780.3; ICD10: G40-41 (2P or 1H or 1ER 2yr)	Epilepsy	350		85.7	82.1	89.4	310	96.8	94.8	98.7	370	97.30	95.60	99.00	410	87.80	84.60	91.00	300	50	360	10
Williamson 2014	Canada	90% >60 years	HIGH	Medical records	Primary care	ICD9, AED	ICD9: 345 (.2, .3, + ≥1AED), 780.3 + ≥1AED	Epilepsy			85.6	80.2	91.1		98.6	96.6	100.0		99.90	99.80	100		98.70	98.20	99.20				
Jetté 2010	Canada	Unclear	LOW	Medical records	Inpatients, A&E	ICD10	ICD10: G40-41, R56	Epilepsy	77		85.1	72.3	92.6																
Jetté 2010	Canada	Unclear	LOW	Medical records	Inpatients	ICD10	ICD10: G40-41	Epilepsy	127		84.9	76.3	90.8	127	98.8	93.3	99.8	127	97.10	85.10	99.50	127	69.60	55.20	80.90				
Reid 2012	Canada	≥18 years	LOW	Medical records	Physician claims	ICD9, ICD10	ICD9: 345; ICD10: G40-41 (1P 2yr)	Epilepsy	340		84.4	80.6	88.3	298	96.3	94.2	98.5	380	97.10	95.40	98.80	422	87.40	84.30	90.60	287	53	369	11
Reid 2012	Canada	≥18 years	LOW	Medical records	Inpatients, physician claims	ICD9, ICD10	ICD9: 345; ICD10: G40-41 (1P or 1H 2yr)	Epilepsy	340		84.4	80.6	88.3	298	96.3	94.2	98.5	380	97.10	95.40	98.80	422	87.40	84.30	90.60	287	53	369	11
Reid 2012	Canada	≥18 years	LOW	Medical records	Inpatients, A&E, physician claims	ICD9, ICD10	ICD9: 345; ICD10: G40-41 (2P or 1H or 1ER 1yr)	Epilepsy	338		84.3	80.4	88.2	298	95.6	93.3	98.0	382	96.60	94.80	98.40	422	87.40	94.30	90.60	285	53	369	13
Reid 2012	Canada	≥18 years	LOW	Medical records	Inpatients, A&E, physician claims	ICD9, ICD10	ICD9: 345; ICD10: G40-41 (2P or 1H or 1ER 2yr)	Epilepsy	338		84.3	80.4	88.2	298	95.6	93.3	98.0	382	96.60	94.80	98.40	422	87.40	84.30	90.60	285	53	369	13

Reid 2012	Canada	≥18 years	LOW	Medical records	Inpatients, A&E, physician claims	ICD9, ICD10	ICD9: 345; ICD10: G40-41 (2P or 1H or 1ER 2yr)	Epilepsy	832		84.3	81.8	86.7	706	99.3	98.7	99.9	701	99.30	98.70	99.90	827	84.20	81.70	86.70	701	131	696	5
Reid 2012	Canada	≥18 years	LOW	Medical records	Inpatients, A&E, physician claims	ICD9, ICD10	ICD9: 345; ICD10: G40-41 (2P or 1H or 1ER 1yr)	Epilepsy	831		84.2	81.8	86.7	706	99.2	98.5	99.8	702	99.20	98.50	99.80	827	84.20	81.70	86.70	700	131	696	6
Holden 2005	US	All ages	UNCLEAR	Medical records	Inpatients, outpatients, pharmacy	ICD9, AED, Procedure	ICD9: 345, 780.3, 780.31, 780.39, 333.2, 779, 779.1, 780.02, 780.2 + EEG or VNS + ≥1AED	Epilepsy	342		84.1	82.1	86.1	356	81.2	77.1	85.3	919	92.71	91.11	94.04	905	94.10	92.41	95.58	289	53	852	67
Moura 2017	US	≥18 years	UNCLEAR	Medical records	Regional professional claims	ICD9, AED	ICD9: 345, 333.2, 779, 780.3, 780.39, 780.2, 780.31 + ≥1AED	Epilepsy	875		84.1	81.2	86.7	875	69.6	66.4	72.6	502	59.30	55.40	63.10	502	77.10	73.20	80.70				
Christensen 2007	Denmark	Unclear	LOW	Medical records	Inpatients, outpatients	ICD8	ICD8: 345	Epilepsy	83		84.0	75.0	91.0													70	13		
Jetté 2010	Canada	All ages	LOW	Medical records	Inpatients, A&E	ICD9	ICD9: 345, 780.3	Epilepsy	486		84.0	78.4	88.3																
Holden 2005	US	All ages	UNCLEAR	Medical records	Inpatients, outpatients, pharmacy	ICD9, AED	ICD9: 345, 780.3, 780.31, 780.39, 333.2, 779, 779.1, 780.02, 780.2 + ≥1AED	Epilepsy	347		83.9	81.9	85.9	356	81.8	77.8	85.8	914	92.89	91.29	94.21	905	93.80	92.04	95.29	291	56	849	65
Reid 2012	Canada	≥18 years	LOW	Medical records	Inpatients, A&E, physician claims	ICD9, ICD10	ICD9: 345, 780.3; ICD10: G40-41 (1P or 1H or 1ER 2yr)	Epilepsy	367		83.9	80.2	87.7	310	99.4	98.5	100.0	353	99.40	98.70	100	410	85.60	82.20	89.00	308	59	351	2
Moura 2017	US	≥18 years	UNCLEAR	Medical records	Regional professional claims	ICD9	ICD9: 345 (2O and/or 1H)	Epilepsy	875		83.8	80.4	86.9	875	49.8	46.5	53.2	502	48.80	45.40	52.20	502	83.30	79.70	86.40				
Jetté 2010	Canada	All ages	LOW	Medical records	A&E	ICD9	ICD9: 345, 780.3	Epilepsy	277		83.6	75.6	89.4																
Jetté 2010	Canada	All ages	LOW	Medical records	Inpatients	ICD9	ICD9: 345, 780.3	Epilepsy	209		83.5	75.1	89.4																
Reid 2012	Canada	≥18 years	LOW	Medical records	Inpatients, A&E, physician claims	ICD9, ICD10	ICD9: 345, 780.3; ICD10: G40-41 (2P or 1H or 1ER 1yr)	Epilepsy	851		83.1	80.6	85.6	712	99.3	98.7	99.9	682	99.30	98.60	99.90	821	82.50	79.90	85.10	707	144	677	5
Reid 2012	Canada	≥18 years	LOW	Medical records	Inpatients, A&E, physician claims	ICD9, ICD10	ICD9: 345, 780.3; ICD10: G40-41 (2P or 1H or 1ER 2yr)	Epilepsy	852		83.1	80.6	85.6	712	99.4	98.9	100.0	681	99.40	98.80	100	821	82.50	79.90	85.10	708	144	677	4
Christensen 2007	Denmark	Unclear	LOW	Medical records	Inpatients, outpatients	ICD8, ICD10	ICD8: 345; ICD10: G40-41	Epilepsy	150		83.0	76.0	89.0													125	26		
Reid 2012	Canada	≥18 years	LOW	Medical records	Inpatients, A&E, physician claims	ICD9, ICD10	ICD9: 345; ICD10: G40-41 (1P or 1H or 1ER 2yr)	Epilepsy	849		82.8	80.3	85.3	706	99.6	99.1	100.0	684	99.60	99.10	100	827	82.40	79.80	84.90	703	146	681	3
Reid 2012	Canada	≥18 years	LOW	Medical records	Inpatients, A&E, physician claims	ICD9, ICD10	ICD9: 345; ICD10: G40-41 (1P or 1H or 1ER 2yr)	Epilepsy	356		82.3	78.3	86.3	298	98.3	96.9	99.8	364	98.60	97.40	99.80	422	85.10	81.70	88.50	293	63	359	5
Rehman 2016	US	Unclear	UNCLEAR	Medical records	Inpatients, outpatients, pharmacy	ICD9, AED	ICD9: 345, 780.39 (3 yr excluding EEG/VTEM) + ≥1AED (≥30 days)	Epilepsy	500		82.0	78.6	85.4													410	90		
Tu 2014	Canada	≥20 years	UNCLEAR	Medical records	Inpatients, physician claims	ICD8, ICD9, ICD10	ICD8: 345; ICD9: 345 (excluding .2,.3); ICD10:	Epilepsy	82		81.7	73.3	90.1	95	70.5	61.4	79.7	7418	99.60	99.50	99.80	7405	99.80	99.70	99.90	67	15	7390	28

							G40 (H or 4P >30 days 3 yr)																						
Reid 2012	Canada	≥18 years	LOW	Medical records	Inpatients, A&E, physician claims	ICD9, ICD10	ICD9: 345, 780.3; ICD10: G40-41 (1P or 1H or 1ER 2yr)	Epilepsy	870	81.6	79.0	84.2	712	99.7	99.3	100.0	663	99.70	99.30	100	821	80.50	77.80	83.20	710	160	661	2
Tan 2015	Australia	≥16 years	LOW	Medical records	Inpatients, A&E, pharmacy	ICD10, AED	ICD10: G40-41 + ≥1AED	Epilepsy	118	81.4	73.1	87.9	158	60.8	52.7	68.4	42642	99.90	99.80	99.90	42602	99.90	99.90	100	96	22	42580	62
Tu 2014	Canada	≥20 years	UNCLEAR	Medical records	Inpatients, physician claims	ICD8, ICD9, ICD10	ICD8: 345; ICD9: 345 (excluding .2,.3); ICD10: G40 (H or 4P >30 days 2 yr)	Epilepsy	80	81.3	72.7	89.8	95	68.4	59.1	77.8	7420	99.60	99.50	99.70	7405	99.80	99.70	99.90	65	15	7390	30
Franchi 2013	Italy	All ages	HIGH	GP	Pharmacy	Other	EXE: 017	Epilepsy	37	81.1	64.3	91.4	71	42.3	30.8	54.5	15691	99.70	99.60	99.80	15657	99.90	99.90	100	30	7	15650	41
Christensen 2007	Denmark	Unclear	LOW	Medical records	Inpatients, outpatients	ICD8, ICD10	ICD8: 345; ICD10: G40-41	Epilepsy	188	81.0	75.0	87.0													152	36		
Lee 2016	Korea	≥1 year	HIGH	Medical records	Nationwide healthcare & pharmacy claims	ICD10, AED	ICD10: G40-41, R56, F80.3 + ≥1AED	Epilepsy	155307	81.0	80.8	81.2																
Tu 2014	Canada	≥20 years	UNCLEAR	Medical records	Inpatients, physician claims, physician claims (neurologist)	ICD8, ICD9, ICD10	ICD8: 345; ICD9: 345 (excluding .2,.3); ICD10: G40 (H or 3P (2 neuro) >30 days 2 yr)	Epilepsy	73	80.8	71.8	89.9	95	62.1	52.3	71.9	7427	99.50	99.40	99.70	7405	99.80	99.70	99.90	59	14	7391	36
Moura 2017	US	≥18 years	UNCLEAR	Medical records	Regional professional claims	ICD9, AED	2 ICD9: 345, 333.2, 779, 780.3, 780.39, 780.2, 780.31 + ≥1AED	Epilepsy	875	80.7	76.2	84.6	875	33.4	30.3	36.6	502	42.60	39.50	45.70	502	86.10	82.70	89.00				
Syvertsen 2015	Norway	All ages	UNCLEAR	Medical records	Specialist centre (neurology)	ICD10	ICD10: G40	Epilepsy	2662	80.0	78.0	82.0													2130	532		
Tu 2014	Canada	≥20 years	UNCLEAR	Medical records	Inpatients	ICD8, ICD9, ICD10	ICD8: 345; ICD9: 345 (excluding .2,.3); ICD10: G40 (H)	Epilepsy	40	80.0	67.6	92.4	95	33.7	24.2	43.2	7460	99.20	98.90	99.40	7405	99.90	99.80	100	32	8	7397	63
Tan 2015	Australia	≥16 years	LOW	Medical records	Inpatients, A&E, pharmacy	ICD10, AED	ICD10: G40 + ≥1AED	Epilepsy	300	79.8	71.1	86.9	158	55.1	47.0	63.0		99.80	99.80	99.90	42602	99.90	99.90	100				
Tu 2014	Canada	≥20 years	UNCLEAR	Medical records	Inpatients, physician claims	ICD8, ICD9, ICD10	ICD8: 345; ICD9: 345 (excluding .2,.3); ICD10: G40 (H or 3P >30 days 1 yr)	Epilepsy	84	79.8	71.2	88.4	95	70.5	61.4	79.7	7416	99.60	99.50	99.80	7405	99.80	99.70	99.90	67	17	7388	28
Tu 2014	Canada	≥20 years	UNCLEAR	Medical records	Inpatients, physician claims	ICD8, ICD9, ICD10	ICD8: 345; ICD9: 345 (excluding .2,.3); ICD10: G40 (H or 3P >30 days 2 yr)	Epilepsy	88	79.5	71.1	88.0	95	73.7	64.8	82.5	7412	99.70	99.50	99.80	7405	99.80	99.60	99.90	70	18	7387	25
Tu 2014	Canada	≥20 years	UNCLEAR	Medical records	Inpatients, physician claims, physician claims (neurologist)	ICD8, ICD9, ICD10	ICD8: 345; ICD9: 345 (excluding .2,.3); ICD10: G40 (H or 3P (1 neuro) >30 days 2 yr)	Epilepsy	78	79.5	70.5	88.4	95	65.3	55.7	74.8	7422	99.60	99.40	99.70	7405	99.80	99.70	99.90	62	16	7389	33
Jetté 2010	Canada	Unclear	LOW	Medical records	Inpatients, A&E	ICD9	ICD9: 345, 780.3	Epilepsy	405	79.4	72.3	85.0																
Moura 2017	US	≥18 years	UNCLEAR	Medical records	Regional professional claims	ICD9, AED	ICD9: 345, 333.2, 779, 780.3, 780.39, 780.2, 780.31	Epilepsy	875	79.3	71.4	85.8	875	12.2	10.1	14.6	502	38.20	35.50	40.90	502	94.40	92.00	96.30				

							+ ≥1AED + ≥1AED used																			
Holden 2005	US	All ages	UNCLEAR	Medical records	Inpatients, outpatients, pharmacy	ICD9	ICD9: 345, 780.3, 780.31, 780.39, 333.2, 779, 779.1, 780.02, 780.2	Epilepsy	342	79.2	77.0	81.4	356	76.9	72.5	81.3	919	91.08	89.40	92.51	905	92.50	90.57	94.12	274	68	837	82
Tu 2014	Canada	≥20 years	UNCLEAR	Medical records	Inpatients, physician claims (neurologist)	ICD8, ICD9, ICD10	ICD8: 345; ICD9: 345 (excluding .2,.3); ICD10: G40 (H or 3P (3 neuro) >30 days 2 yr)	Epilepsy	67	79.1	69.4	88.8	95	55.8	45.8	65.8	7433	99.40	99.30	99.60	7405	99.80	99.70	99.90	53	14	7391	42
Christensen 2007	Denmark	Unclear	LOW	Medical records	Inpatients, outpatients	ICD10	ICD10: G40	Epilepsy	105	79.0	70.0	86.0													83	22		
Moura 2017	US	≥18 years	UNCLEAR	Medical records	Regional professional claims	AED, Procedure	EEG + ≥1AED	Epilepsy	875	78.6	74.1	82.5	875	34.7	31.6	38.0	502	42.30	39.20	45.50	502	83.50	79.90	86.60				
Jetté 2010	Canada	All ages	LOW	Medical records	Inpatients	ICD10	ICD10: G40-41, R56	Epilepsy	247	78.1	69.6	84.7																
Tu 2014	Canada	≥20 years	UNCLEAR	Medical records	Inpatients, physician claims	ICD8, ICD9, ICD10	ICD8: 345; ICD9: 345 (excluding .2,.3); ICD10: G40 (H or 3P >30 days 3 yr)	Epilepsy	90	77.8	69.2	86.4	95	73.7	64.8	82.5	7410	99.70	99.50	99.80	7405	99.70	99.60	99.80	70	20	7385	25
Fonferko-Shadrach 2017	UK	All ages	HIGH	Medical records	Primary care	AED	AEDx2 0.5 yr	Epilepsy	288	77.0	70.0	83.0	145	92.0	86.0	96.0	116	89.66	83.32	93.76	143	73.00	65.00	80.00	133	39	104	12
Shackleton 1997	Holland	0-75 years	HIGH	GP	Pharmacy	AED	1AED (> 1 yr)	Epilepsy	132	77.0	70.0	84.0													102	30		
Jetté 2010	Canada	All ages	LOW	Medical records	Inpatients, A&E	ICD10	ICD10: G40-41, R56	Epilepsy	454	75.5	68.9	81.1																
Tu 2014	Canada	≥20 years	UNCLEAR	Medical records	Inpatients, physician claims (neurologist)	ICD8, ICD9, ICD10	ICD8: 345; ICD9: 345 (excluding .2,.3); ICD10: G40 (H or 2P (2 neuro) >30 days 2 yr)	Epilepsy	81	75.3	65.9	84.7	95	64.2	54.6	73.9	7419	99.50	99.40	99.70	7405	99.70	99.60	99.80	61	20	7385	34
Jetté 2010	Canada	Unclear	LOW	Medical records	Inpatients, A&E	ICD10	ICD10: G40-41, R56	Epilepsy	377	72.3	64.4	79.1																
Jetté 2010	Canada	All ages	LOW	Medical records	A&E	ICD10	ICD10: G40-41, R56	Epilepsy	207	71.6	60.5	80.6																
Tan 2015	Australia	≥16 years	LOW	Medical records	Inpatients, A&E	ICD10	ICD10: G40-41	Epilepsy	149	71.3	63.4	78.4	158	67.7	59.8	74.9		99.90	99.80	99.90	42602	99.90	99.90	99.90	106	43		
Tan 2015	Australia	≥16 years	LOW	Medical records	Inpatients, A&E	ICD10	ICD10: G40-41, R56.8	Epilepsy	300	71.3	63.4	78.4	158	67.7	59.8	74.9		99.90	99.80	99.90	42602	99.90	99.90	99.90				
Fonferko-Shadrach 2017	UK	>16 years	HIGH	Medical records	Primary care	AED	AEDx2 0.5 yr	Epilepsy	129	71.0	63.0	78.0	97	94.0	87.0	98.0	66	90.91	81.93	95.66	98	61.00	51.00	71.00	91	38	60	6
Tu 2014	Canada	≥20 years	UNCLEAR	Medical records	Inpatients, physician claims, physician claims (neurologist)	ICD8, ICD9, ICD10	ICD8: 345; ICD9: 345 (excluding .2,.3); ICD10: G40 (H or 2P (1 neuro) >30 days 2 yr)	Epilepsy	92	70.7	61.3	80.0	95	68.4	59.1	77.8	7408	99.60	99.50	99.70	7405	99.60	99.50	99.80	65	27	7378	30
Tan 2015	Australia	≥16 years	LOW	Medical records	Inpatients, A&E	ICD10	ICD10: G40	Epilepsy	139	69.5	61.2	77.0	158	62.0	54.0	69.6		99.90	99.90	99.90	42602	99.90	99.90	99.90	97	42		
Tan 2015	Australia	≥16 years	LOW	Medical records	Inpatients, A&E	ICD10	ICD10: G40, R56.8	Epilepsy	290	69.5	61.2	77.0	158	62.0	54.0	69.6		99.90	99.80	99.90	42602	99.90	99.90	99.90				
Tu 2014	Canada	≥20 years	UNCLEAR	Medical records	Inpatients, physician claims	ICD8, ICD9, ICD10	ICD8: 345; ICD9: 345 (excluding .2,.3); ICD10: G40 (H or 3P 1 yr)	Epilepsy	103	68.9	60.0	77.9	95	74.7	66.0	83.5	7397	99.70	99.50	99.80	7405	99.60	99.40	99.70	71	32	7373	24

Tu 2014	Canada	≥20 years	UNCLEAR	Medical records	Inpatients, physician claims	ICD8, ICD9, ICD10	ICD8: 345; ICD9: 345 (excluding .2,.3); ICD10: G40 (H or 2P >30 days 1 yr)	Epilepsy	109	68.8	60.1	77.5	95	78.9	70.7	87.1	7391	99.70	99.60	99.80	7405	99.50	99.40	99.70	75	34	7371	20
Tu 2014	Canada	≥20 years	UNCLEAR	Medical records	Inpatients, physician claims	ICD8, ICD9, ICD10	ICD8: 345; ICD9: 345 (excluding .2,.3); ICD10: G40 (H or 3P 2 yr)	Epilepsy	107	68.2	59.4	77.0	95	76.8	68.4	85.3	7393	99.70	99.60	99.80	7405	99.50	99.40	99.70	73	34	7371	22
Tu 2014	Canada	≥20 years	UNCLEAR	Medical records	Inpatients, physician claims	ICD8, ICD9, ICD10	ICD8: 345; ICD9: 345 (excluding .2,.3); ICD10: G40 (H or 3P 2 yr)	Epilepsy	107	68.2	59.4	77.0	95	76.8	68.4	85.3	7393	99.70	99.60	99.80	7405	99.50	99.40	99.70	73	34	7371	22
Aaberg 2017	Norway	3-13 years	HIGH	Medical records &/or parent telephone interview	Inpatients, outpatients, private specialist claims	ICD10	ICD10: G40 (position 1)	Epilepsy	838	68.0	65.0	72.0													570	269		
Shackleton 1997	Holland	0-75 years	HIGH	GP	Pharmacy	AED	≥1AED	Epilepsy	372	67.0	62.0	72.0													251	121		
Aaberg 2017	Norway	3-13 years	HIGH	Medical records &/or parent telephone interview	Inpatients, outpatients, private specialist claims	ICD10	ICD10: G40 (1 or ≥2R)	Epilepsy	838	66.0	63.0	69.0													553	285		
Tu 2014	Canada	≥20 years	UNCLEAR	Medical records	Inpatients, physician claims	ICD8, ICD9, ICD10	ICD8: 345; ICD9: 345 (excluding .2,.3); ICD10: G40 (H or 2P >30 days 2 yr)	Epilepsy	118	65.3	56.7	73.8	95	81.1	73.2	88.9	7382	99.80	99.60	99.90	7405	99.40	99.30	99.60	77	41	7364	18
Tu 2014	Canada	≥20 years	UNCLEAR	Medical records	Inpatients, physician claims	ICD8, ICD9, ICD10	ICD8: 345; ICD9: 345 (excluding .2,.3); ICD10: G40 (H or 2P >30 days 3 yr)	Epilepsy	120	65.0	56.5	73.5	95	82.1	74.4	89.8	7380	99.80	99.70	99.90	7405	99.40	99.30	99.60	78	42	7363	17
Franchi 2013	Italy	All ages	HIGH	GP	Pharmacy	AED, Procedure	EEG + ≥1AED	Epilepsy	95	64.2	53.7	73.6	71	85.9	76.0	92.2	15633	99.90	99.90	100	15657	99.80	99.70	99.90	61	34	15623	10
Franchi 2013	Italy	All ages	HIGH	GP	Pharmacy	ICD9, AED, Procedure	ICD9: 345, 333.2, 780.3, 779.0, 781.0 + EEG + ≥1AED	Epilepsy	95	64.2	53.7	73.6	71	85.9	76.0	92.2	15633	99.90	99.90	100	15657	99.80	99.70	99.90	61	34	15623	10
Franchi 2013	Italy	All ages	HIGH	GP	Pharmacy	AED, Procedure, Other	EXE: 017 + EEG + ≥1AED	Epilepsy	95	64.2	53.7	73.6	71	85.9	76.0	92.2	15633	99.90	99.90	100	15657	99.80	99.70	99.90	61	34	15623	10
Chen 2012	Taiwan	≥30 years	HIGH	Previous epilepsy prevalence study	Physician claims, pharmacy	ICD9, AED	ICD9: 345 + ≥1AED	Epilepsy	56	62.5	49.8	75.2	43	81.4	69.8	93.0	12657	99.94	99.88	99.97	12670	99.83	99.75	99.90	35	21	12649	8
Parko 2009	US	All ages	UNCLEAR	Medical records	Inpatients, outpatients	ICD9	ICD9: 345, 780.3	Epilepsy	1277	62.0	59.3	64.7																
Franchi 2013	Italy	All ages	HIGH	GP	Pharmacy	ICD9, AED, Procedure, Other	ICD9: 345, 333.2, 780.3, 779.0, 781.0 + EXE: 017 + EEG + ≥1AED	Epilepsy	103	61.2	51.0	70.5	71	88.7	79.3	96.2	15625	99.90	99.90	100	15657	99.70	99.70	99.80	63	40	15617	8
Tu 2014	Canada	≥20 years	UNCLEAR	Medical records	Inpatients, physician claims	ICD8, ICD9, ICD10	ICD8: 345; ICD9: 345 (excluding .2,.3); ICD10: G40 (H or 2P 1yr)	Epilepsy	132	60.6	52.3	68.9	95	84.2	76.9	91.5	7368	99.80	99.70	99.90	7405	99.30	99.10	99.50	80	52	7353	15
Tan 2015	Australia	≥16 years	LOW	Medical records	Inpatients, A&E, pharmacy	ICD10, AED	ICD10: R56.8 + ≥1AED	Epilepsy	300	59.8	53.2	66.2	158	88.6	82.6	93.1		100	99.90	100	42602	99.80	99.70	99.80				
Tu 2014	Canada	≥20 years	UNCLEAR	Medical records	Physician claims (neurologist)	ICD8, ICD9, ICD10	ICD8: 345; ICD9: 345 (excluding	Epilepsy	117	59.8	50.9	68.7	95	73.7	64.8	82.5	7383	99.70	99.50	99.80	7405	99.40	99.20	99.50	70	47	7358	25

							.2,.3); ICD10: G40 (P neuro)																							
Rehman 2016	US	Unclear	UNCLEAR	Medical records	Inpatients, outpatients, pharmacy	ICD9, AED	ICD9: 345, 780.39 (1 yr) + ≥1AED (≥30 days)	Epilepsy	527	58.4	54.2	62.7															308	219		
Tu 2014	Canada	≥20 years	UNCLEAR	Medical records	Inpatients, physician claims	ICD8, ICD9, ICD10	ICD8: 345; ICD9: 345 (excluding .2,.3); ICD10: G40 (H or 2P 2yr)	Epilepsy	139	58.3	50.1	66.5	95	85.3	78.1	92.4	7361	99.80	99.70	99.90	7405	99.20	99.00	99.40	81	58	7347	14
Tu 2014	Canada	≥20 years	UNCLEAR	Medical records	Inpatients, physician claims	ICD8, ICD9, ICD10	ICD8: 345; ICD9: 345 (excluding .2,.3); ICD10: G40 (H or 2P 3 yr)	Epilepsy	141	58.2	50.0	66.3	95	86.3	79.4	93.2	7359	99.80	99.70	99.90	7405	99.20	99.00	99.40	82	59	7346	13
Shackleton 1997	Holland	0-75 years	HIGH	GP	Pharmacy	AED	1AED (> 1 yr)	Epilepsy	266	58.0	52.0	64.0															153	113		
Franchi 2013	Italy	All ages	HIGH	GP	Pharmacy	Procedure	EEG	Epilepsy	109	56.0	46.1	65.4	71	85.9	76.0	92.2	15619	99.90	99.90	100	15657	99.70	99.60	99.80	61	48	15609	10
Franchi 2013	Italy	All ages	HIGH	GP	Pharmacy	ICD9, Procedure	ICD9: 345, 333.2, 780.3, 779.0, 781.0 + EEG	Epilepsy	109	56.0	46.1	65.4	71	85.9	76.0	92.2	15619	99.90	99.90	100	15657	99.70	99.60	99.80	61	48	15609	10
Franchi 2013	Italy	All ages	HIGH	GP	Pharmacy	Procedure, Other	EXE: 017 + EEG	Epilepsy	109	56.0	46.1	65.4	71	85.9	76.0	92.2	15619	99.90	99.90	100	15657	99.70	99.60	99.80	61	48	15609	10
Franchi 2013	Italy	All ages	HIGH	GP	Pharmacy	ICD9, Procedure, Other	ICD9: 345, 333.2, 780.3, 779.0, 781.0 + EXE: 017 + EEG	Epilepsy	117	53.8	44.4	63.0	71	88.7	79.3	94.2	15611	99.90	99.90	100	15657	99.60	99.60	99.80	63	54	15603	8
Pugh 2015	US	Unclear	UNCLEAR	Medical records	Inpatients, outpatients, pharmacy	ICD9	1 ICD9: 345, ≥2 780.3	Epilepsy	25	52.0	32.0	72.0															13	12		
Shackleton 1997	Holland	0-75 years	HIGH	GP	Pharmacy	AED	Carbamazepine (> 1 yr)	Epilepsy	94	52.0	42.0	62.0															49	45		
Marrie 2013	Canada	≥20 years	HIGH	Medical records	Inpatients, physician claims, pharmacy	ICD9, ICD10	ICD9; ICD10 (≥1 H or P) + ≥1AED 1 yr	Comorbid epilepsy		50.0	6.7	93.2		22.2	2.8	60.0		98.30	96.50	99.30		99.50	98.20	99.90					
Wassenaar 2018	Holland	18-75 years	HIGH	Medical records	Physician claims (neurologist)	ICD9	ICD9: 345 (2P 3yr)	Epilepsy	24	50.0	30.0	70.0																		
Pugh 2015	US	Unclear	UNCLEAR	Medical records	Inpatients, outpatients, pharmacy	ICD9	≥2 345	Epilepsy	25	48.0	28.0	68.0															12	13		
Coulter 1989	UK	Unclear	HIGH	Antiepileptic drugs	Primary care	Chronic disease register	Presence on the register	Epilepsy	427	47.0	42.0	52.0															201	226		
Franchi 2013	Italy	All ages	HIGH	GP	Pharmacy	ICD9, Other	ICD9: 345, 333.2, 780.3, 779.0, 781.0 + EXE: 017	Epilepsy	97	46.4	36.3	56.8	71	63.4	51.8	73.6	15631	99.80	99.80	99.90	15657	99.70	99.60	99.80	45	52	15605	26
Franchi 2013	Italy	All ages	HIGH	GP	Pharmacy	ICD9, AED, Other	ICD9: 345, 333.2, 780.3, 779.0, 781.0 + EXE: 017 + ≥1AED	Epilepsy	97	46.4	36.3	56.8	71	63.4	51.8	73.6	15631	99.80	99.80	99.90	15657	99.70	99.60	99.80	45	52	15605	26
Marrie 2013	Canada	≥20 years	HIGH	Medical records	Inpatients, physician claims, pharmacy	ICD9, ICD10	ICD9; ICD10 (≥1 H or (≥1 P + ≥2 AED 3 yr)	Comorbid epilepsy		38.5	13.9	68.4		55.6	21.2	86.3		98.90	97.40	99.70		97.90	96.10	99.10					
Engeland 2009	Norway	<20-40+ years	HIGH	Antiepileptic drugs	National maternal disease registry	ICD10	ICD10	Epilepsy	855	37.0	34.0	40.0	426	74.0	69.0	78.0	107634	99.90	99.88	99.91	108063	100	99.00	100	314	541	107522	112
Tu 2014	Canada	≥20 years	UNCLEAR	Medical records	Inpatients, physician claims	ICD8, ICD9, ICD10	ICD8: 345; ICD9: 345 (excluding .2,.3); ICD10: G40 (H or P)	Epilepsy	234	36.8	30.6	42.9	95	90.5	84.6	96.4	7266	99.90	99.80	100	7405	98.00	97.70	98.30	86	148	7257	9
Marrie 2013	Canada	≥20 years	HIGH	Medical records	Inpatients, physician claims, pharmacy	ICD9, ICD10	ICD9; ICD10 (≥1 H or P) + ≥1AED 2 yr	Comorbid epilepsy		36.4	10.9	69.2		44.4	13.7	78.8		98.70	97.10	99.60		98.20	96.40	99.30					

Franchi 2013	Italy	All ages	HIGH	GP	Pharmacy	ICD9	ICD9: 345, 333.2, 780.3, 779.0, 781.0	Epilepsy	75	34.7	24.3	46.6	71	36.6	25.7	49.0	15653	99.70	99.60	99.80	15657	99.70	99.60	99.80	26	49	15608	45
Tan 2015	Australia	≥16 years	LOW	Medical records	Inpatients, A&E	ICD10	ICD10: R56.8	Epilepsy	151	34.4	26.9	42.6	158	32.9	25.7	40.8		99.80	99.70	99.80	42602	99.80	99.70	99.80	52	99		
Marrie 2013	Canada	≥20 years	HIGH	Medical records	Inpatients, physician claims, pharmacy	ICD9, ICD10	ICD9; ICD10 (≥1 H or P) + ≥1AED 3 yr	Comorbid epilepsy		33.3	11.8	61.6		55.6	21.2	86.3		98.90	97.40	99.70		97.50	95.40	98.80				
Frost 2000	US	All ages	HIGH	GP plus/minus medical records	Inpatients, outpatients, pharmacy	ICD9, AED, Procedure	ICD9: 345, 780.3 or EEG or ≥1AED	Epilepsy	2474	32.7	30.9	34.6													808	1666		
Foebel 2013	Canada	All ages	HIGH	Administrative data	Psychiatric inpatients	ICD10	ICD10: G40	Epilepsy	406	27.0	23.0	32.0	273	41.0	35.0	47.0	13406	98.78	98.66	98.90	13539	98.00	98.00	98.00	110	296	13243	163
Marrie 2013	Canada	≥20 years	HIGH	Medical records	Inpatients, physician claims, pharmacy	ICD9, ICD10	ICD9; ICD10 (≥1 H or P 1 yr)	Comorbid epilepsy		25.0	3.2	65.1		22.2	2.8	60.0		98.30	96.40	99.30		98.50	96.70	99.40				
Marrie 2013	Canada	≥20 years	HIGH	Medical records	Inpatients, physician claims, pharmacy	ICD9, ICD10	ICD9; ICD10 (≥1 H or ≥2P 5yr)	Comorbid epilepsy		23.5	6.8	49.9		44.4	13.7	78.8		98.70	97.00	99.60		96.70	94.40	98.20				
Marrie 2013	Canada	≥20 years	HIGH	Medical records	Inpatients, physician claims, pharmacy	ICD9, ICD10	ICD9; ICD10 (≥1 H or P 2 yr)	Comorbid epilepsy		22.2	6.4	47.6		44.4	13.7	78.8		98.70	97.00	99.60		96.50	94.20	98.10				
Marrie 2013	Canada	≥20 years	HIGH	Medical records	Inpatients, physician claims, pharmacy	ICD9, ICD10	ICD9; ICD10 (≥1 H or P 3 yr)	Comorbid epilepsy		21.7	7.5	43.7		55.6	21.2	86.3		98.90	97.40	99.70		95.50	93.00	97.30				
Marrie 2013	Canada	≥20 years	HIGH	Medical records	Inpatients, physician claims, pharmacy	ICD9, ICD10	ICD9; ICD10 (≥1 H or P) + ≥1 AED 5 yr)	Comorbid epilepsy		21.7	7.5	97.3		55.6	21.2	86.3		98.90	97.40	99.70		95.50	93.00	97.30				
Bellini 2017	Italy	All ages	UNCLEAR	Epilepsy centre patient list	Inpatients, pharmacy, procedures	ICD9, AED, Procedure	ICD9: 345 and/or EEG + ≥1AED (2P 1yr) and/or ≥1 specific AED (2P 1yr)	Epilepsy	3117	21.2	19.6	22.4	756	87.3	84.9	89.7	2172076	100	99.99	100	2174437	99.89	99.88	99.89	660	2457	2171980	96
Marrie 2013	Canada	≥20 years	HIGH	Medical records	Inpatients, physician claims, pharmacy	ICD9, ICD10	ICD9; ICD10 (≥1 H or P 5 yr)	Comorbid epilepsy		15.2	5.1	31.9		55.6	21.2	86.3		98.90	97.30	99.70		92.90	90.00	95.30				
Foebel 2013	Canada	All ages	HIGH	Administrative data	Long-term care and complex continuing care facilities	ICD10, Other	ICD10: G40; RAI 2.0 pick list items	Epilepsy	4681	14.0	13.0	15.0	910	72.0	69.0	75.0	95180	99.73	99.70	99.76	98951	96.00	96.00	96.00	655	4026	94925	255
Franchi 2013	Italy	All ages	HIGH	GP	Pharmacy	AED	≥1AED	Epilepsy	469	13.6	10.7	17.2	71	90.1	81.0	95.1	15259	99.90	99.90	100	15657	97.40	97.20	97.70	64	405	15252	7
Franchi 2013	Italy	All ages	HIGH	GP	Pharmacy	ICD9, AED	ICD9: 345, 333.2, 780.3, 779.0, 781.0 + ≥1AED	Epilepsy	469	13.6	10.7	17.2	71	90.1	81.0	95.1	15259	100	99.90	100	15657	97.40	97.20	97.70	64	405	15252	7
Franchi 2013	Italy	All ages	HIGH	GP	Pharmacy	AED, Other	EXE: 017 + ≥1AED	Epilepsy	469	13.6	10.7	17.2	71	90.1	81.0	95.1	15259	100	99.90	100	15657	97.40	97.20	97.70	64	405	15252	7
Lee 2016	Korea	≥1 year	HIGH	Medical records	Pharmacy	AED	≥1AED	Epilepsy	964053	6.6	6.6	6.7																
Marrie 2013	Canada	≥20 years	HIGH	Medical records	Inpatients, physician claims, pharmacy	ICD9, ICD10	ICD9; ICD10 (≥2 H or P or AED 2 yr)	Comorbid epilepsy		5.6	2.1	11.9		66.7	29.9	92.5		99.00	97.10	99.80		74.90	70.40	79.10				
Marrie 2013	Canada	≥20 years	HIGH	Medical records	Inpatients, physician claims, pharmacy	ICD9, ICD10	ICD9; ICD10 (≥2 H or P or AED 1 yr)	Comorbid epilepsy		5.2	1.4	12.8		44.4	13.7	78.8		98.50	96.50	99.50		81.70	77.50	85.40				
Shackleton 1997	Holland	0-75 years	HIGH	GP	Pharmacy	AED	Clonazepam (> 1 yr)	Epilepsy	40	5.0	0.0	12.0													2	38		
Pickrell 2015	UK	All ages	UNCLEAR	Specialist epilepsy database	Primary care	Read Codes, AED	Read Code: F25, 1O30, 667B., SC200 + AEDx2 1yr	Epilepsy					1834	90.5	89.2	91.8									1660			174

Pina-Garza 2015	US	Unclear	HIGH	Rufinamide	Regional claims	ICD9	Random forest classifier	Lennox-gastaut syndrome		97.3	95.60

Abbreviations: UK – United Kingdom; US – United States; A&E – Accident and emergency; ICD – International classification of disease; ICD10 G40 – epilepsy; ICD10 G41 – status epilepticus; ICD9 345 – epilepsy; F25, 1O30, 667B., SC200 – epilepsy Read Codes; AED – antiepileptic drug; ICD9 780.39 – convulsions; Yr – year; P – Physician claim; H – Hospitalisation; ICD9 333.2 – Myoclonus; ICD9 779 – Perinatal illness; ICD9 780.3 – Convulsions; ICD9 780.2 – Syncope and collapse; ICD9 780.31 – Febrile convulsions; ICD9 311, 300.4, 300, 300.02 – Anxiety depression codes; EEG – Electroencephalogram; VNS – Vagal nerve stimulation; ICD8 345 – Epilepsy; EXE – Exception code; ICD10 F80.3 – Acquired aphasia with epilepsy; RAI - Resident Assessment Instrument; PPV – Positive predictive value (%); NPV – Negative predictive value (%); Sn – Sensitivity (%); Sp – Specificity (%); CI – 95% Confidence interval (%). Overall quality score = Quality Assessment of Diagnostic Accuracy Studies 2 (QUADAS-2)[146] score for overall risk of study bias or applicability concern for that study

Appendix S8: Summary of results from countries with a single study

Overview

The Danish study[123] (with gold standard medical records) showed little difference in accuracy between ICD-8 345 and ICD-10 G40–41 (PPV range 79.0–84.0%). In the Australian study (with gold standard medical records),[143] the highest sensitivity came from a combination of R56.8 plus one or more AEDs (88.6%, 95% CI 82.6–93.1%, Figure 4B). Surprisingly, the study demonstrated a highest PPV for G41 (90.0%, 95% CI 55.5–99.7%, Figure 4A) as compared to PPV 69.5% (95% CI 61.2–77.0%) for G40. This might suggest the dramatic nature of a status epilepticus admission in people with epilepsy tended to result in their epilepsy being coded correctly more often. However, G41 had the lowest sensitivity of all included studies: 5.7%, 95% CI 2.6–10.5%, Figure 5B. This is intuitive, as not all persons with epilepsy develop status epilepticus and therefore G41 generates many false negatives. The South Korean study (with gold standard medical records) had an optimal diagnostic algorithm consisting of ICD-10 G40–41, R56, or F80.3 (acquired aphasia with epilepsy) plus one or more AEDs (PPV 81.0%, 95% CI 80.8–81.2%, Figure 4A).[163] One or more AEDs alone generated a worst PPV of 6.6% (95% CI 6.6–6.7%, Figure 5A), reinforcing the potential problems of using AEDs alone. The Taiwan study[151] (with an epilepsy prevalence study gold standard) validated ICD-9 345 plus one or more AEDs and this demonstrated a good sensitivity (81.4%, 95% CI 69.8–93.0, Figure 4B) with a lower PPV trade-off of 62.5% (95% CI 49.8–75.2, Figure 4A).

Denmark

There were four PPV algorithms from a single Danish study validating data from a national hospital register consisting of 57 departments and 41 hospitals, therefore providing good coverage of the country (Appendix S7).[123] Risks of bias and applicability concern were low (Appendix S6B). Medical records were used as the diagnostic gold standard. The study validated ICD-8 345 and ICD-10 G40–41 on their own and in combination, with PPV ranging 79.0–84.0% and little difference in estimates between the two coding systems.

Australia

A single Australian study had low risks of bias and applicability concern (Appendix S6B), and used medical records as the diagnostic gold standard.[143] This study was the only study to allow direct comparison of G40 (epilepsy), G41 (status epilepticus), and R56.8 (convulsions) codes in picking up epilepsy (Appendix S7). R56.8 had the lowest PPV (34.4%, 95% CI 26.9–42.6, Figure 5A), and in this study it also had a poor sensitivity (32.9%, 95% CI 25.7–40.8). PPVs for algorithms containing R56.8 with or without G40, G41, or AEDs remained generally low (ranging 34.4–71.3%). Correspondingly, the highest sensitivity came from a combination of R56.8 plus one or more AEDs (88.6%, 95% CI 82.6–93.1%). Surprisingly, the study demonstrated a highest PPV for G41 (90.0%, 95% CI 55.5–99.7) as compared to PPV 69.5% (95% CI 61.2–77.0) for G40. This might suggest the dramatic nature of a status epilepticus admission in people with epilepsy tended to result in their epilepsy being coded correctly more often. However, G41 had the lowest sensitivity of all included studies: 5.7%, 95% CI 2.6–10.5%, Figure 5B. This is intuitive, as not all persons with epilepsy develop status epilepticus and therefore G41 generates many false negatives.

South Korea and Taiwan

There were two PPV algorithms from Lee 2016[163] in South Korea and one PPV algorithm from Chen 2012[151] in Taiwan, which also estimated sensitivity. Lee 2016[163] used medical records as the diagnostic gold standard, whilst Chen 2012[151] used a previous epilepsy prevalence study. Both studies had high risks of bias or applicability concern (Appendix S6B). Data in South Korea were validated from a nationwide healthcare and pharmacy database using ICD-10. The optimal South Korean algorithm consisted of ICD-10 G40–41, R56, or F80.3 (acquired aphasia with epilepsy) plus one or more AEDs. This generated a PPV of 81.0% (95% CI 80.8–81.2%, Figure 4A). This was the only study included to validate F80.3. One or more AEDs alone generated a worst PPV of 6.6% (95% CI 6.6–6.7, Figure 5A), reinforcing the potential problems of using AEDs alone. The Chen

2012[151] study validated ICD-9 345 plus one or more AEDs and this demonstrated a good sensitivity (81.4%, 95% CI 69.8–93.0) with a lower PPV trade-off of 62.5% (95% CI 49.8–75.2, Appendix S7).

Appendix S9: Meta-analysis

A - Input

Study	Country	Age	Overall quality score	Gold standard	Administrative setting	Coding system	Algorithm	TP	FP	TN	FN
Bellini 2017	Italy	All ages	UNCLEAR	Epilepsy centre patient list	Inpatients, pharmacy, procedures	ICD9, AED, Procedure	ICD9: 345 and/or EEG + ≥1AED (2P 1yr) and/or ≥1 specific AED (2P 1yr)	660	2457	2171980	96
Chen 2012	Taiwan	≥30 years	HIGH	Previous epilepsy prevalence study	Physician claims, pharmacy	ICD9, AED	ICD9: 345 + ≥1AED	35	21	12649	8
Engeland 2009	Norway	<20-40+ years	HIGH	Antiepileptic drugs	National maternal disease registry	ICD10	ICD10	314	541	107522	112
Foebel 2013	Canada	All ages	HIGH	Administrative data	Long-term care and complex continuing care facilities	ICD10, Other	ICD10: G40; RAI 2.0 pick list items	655	4026	94925	255
Foebel 2013	Canada	All ages	HIGH	Administrative data	Psychiatric inpatients	ICD10	ICD10: G40	110	296	13243	163
Fonferko-Shadrach 2017	UK	All ages	HIGH	Medical records	Primary care	Read Codes, AED	Read Code: F25, 1O30, 667B., SC200 + AEDx2 0.5 yr	122	2	141	23
Fonferko-Shadrach 2017	UK	All ages	HIGH	Medical records	Primary care	Read Codes	Read Code: F25, 1O30, 667B., SC200	125	5	138	20
Fonferko-Shadrach 2017	UK	All ages	HIGH	Medical records	Primary care	AED	AEDx2 0.5 yr	133	39	104	12
Fonferko-Shadrach 2017	UK	>16 years	HIGH	Medical records	Primary care	Read Codes, AED	Read Code: F25, 1O30, 667B., SC200 + AEDx2 0.5 yr	84	2	96	13
Fonferko-Shadrach 2017	UK	>16 years	HIGH	Medical records	Primary care	Read Codes	Read Code: F25, 1O30, 667B., SC200	85	2	96	12
Fonferko-Shadrach 2017	UK	>16 years	HIGH	Medical records	Primary care	AED	AEDx2 0.5 yr	91	38	60	6
Fonferko-Shadrach 2017	UK	≤16 years	HIGH	Medical records	Primary care	Read Codes, AED	Read Code: F25, 1O30, 667B., SC200 + AEDx2 0.5 yr	38	0	45	10
Fonferko-Shadrach 2017	UK	≤16 years	HIGH	Medical records	Primary care	Read Codes	Read Code: F25, 1O30, 667B., SC200	40	3	42	8
Fonferko-Shadrach 2017	UK	≤16 years	HIGH	Medical records	Primary care	AED	AEDx2 0.5 yr	42	1	44	6
Franchi 2013	Italy	All ages	HIGH	GP	Pharmacy	AED	≥1AED	64	405	15252	7
Franchi 2013	Italy	All ages	HIGH	GP	Pharmacy	ICD9, AED	ICD9: 345, 333.2, 780.3, 779.0, 781.0 + ≥1AED	64	405	15252	7

Franchi 2013	Italy	All ages	HIGH	GP	Pharmacy	AED, Procedure	EEG + ≥1AED	61	34	15623	10
Franchi 2013	Italy	All ages	HIGH	GP	Pharmacy	AED, Other	EXE: 017 + ≥1AED	64	405	15252	7
Franchi 2013	Italy	All ages	HIGH	GP	Pharmacy	ICD9, Other	ICD9: 345, 333.2, 780.3, 779.0, 781.0 + EXE: 017	45	52	15605	26
Franchi 2013	Italy	All ages	HIGH	GP	Pharmacy	ICD9, AED, Procedure	ICD9: 345, 333.2, 780.3, 779.0, 781.0 + EEG + ≥1AED	61	34	15623	10
Franchi 2013	Italy	All ages	HIGH	GP	Pharmacy	ICD9, AED, Other	ICD9: 345, 333.2, 780.3, 779.0, 781.0 + EXE: 017 + ≥1AED	45	52	15605	26
Franchi 2013	Italy	All ages	HIGH	GP	Pharmacy	AED, Procedure, Other	EXE: 017 + EEG + ≥1AED	61	34	15623	10
Franchi 2013	Italy	All ages	HIGH	GP	Pharmacy	Other	EXE: 017	30	7	15650	41
Franchi 2013	Italy	All ages	HIGH	GP	Pharmacy	ICD9, AED, Procedure, Other	ICD9: 345, 333.2, 780.3, 779.0, 781.0 + EXE: 017 + EEG + ≥1AED	63	40	15617	8
Franchi 2013	Italy	All ages	HIGH	GP	Pharmacy	Procedure	EEG	61	48	15609	10
Franchi 2013	Italy	All ages	HIGH	GP	Pharmacy	ICD9	ICD9: 345, 333.2, 780.3, 779.0, 781.0	26	49	15608	45
Franchi 2013	Italy	All ages	HIGH	GP	Pharmacy	ICD9, Procedure	ICD9: 345, 333.2, 780.3, 779.0, 781.0 + EEG	61	48	15609	10
Franchi 2013	Italy	All ages	HIGH	GP	Pharmacy	Procedure, Other	EXE: 017 + EEG	61	48	15609	10
Franchi 2013	Italy	All ages	HIGH	GP	Pharmacy	ICD9, Procedure, Other	ICD9: 345, 333.2, 780.3, 779.0, 781.0 + EXE: 017 + EEG	63	54	15603	8
Holden 2005	US	All ages	UNCLEAR	Medical records	Inpatients, outpatients, pharmacy	ICD9	ICD9: 345, 780.3, 780.31, 780.39, 333.2, 779, 779.1, 780.02, 780.2	274	68	837	82
Holden 2005	US	All ages	UNCLEAR	Medical records	Inpatients, outpatients, pharmacy	ICD9, AED	ICD9: 345, 780.3, 780.31, 780.39, 333.2, 779, 779.1, 780.02, 780.2 + ≥1AED	291	56	849	65
Holden 2005	US	All ages	UNCLEAR	Medical records	Inpatients, outpatients, pharmacy	ICD9, AED, Procedure	ICD9: 345, 780.3, 780.31, 780.39, 333.2, 779, 779.1, 780.02, 780.2 + EEG or VNS + ≥1AED	289	53	852	67
Reid 2012	Canada	≥18 years	LOW	Medical records	Physician claims	ICD9, ICD10	ICD9: 345; ICD10: G40-41 (1P 2yr)	287	53	369	11
Reid 2012	Canada	≥18 years	LOW	Medical records	Inpatients, physician claims	ICD9, ICD10	ICD9: 345; ICD10: G40-41 (1P or 1H 2yr)	287	53	369	11
Reid 2012	Canada	≥18 years	LOW	Medical records	Inpatients, A&E, physician claims	ICD9, ICD10	ICD9: 345; ICD10: G40-41 (1P or 1H or 1ER 2yr)	293	63	359	5
Reid 2012	Canada	≥18 years	LOW	Medical records	Physician claims	ICD9, ICD10	ICD9: 345; ICD10: G40-41 (2P 1yr)	257	31	391	41
Reid 2012	Canada	≥18 years	LOW	Medical records	Inpatients, physician claims	ICD9, ICD10	ICD9: 345; ICD10: G40-41 (2P or 1H 1yr)	260	31	391	38
Reid 2012	Canada	≥18 years	LOW	Medical records	Inpatients, A&E, physician claims	ICD9, ICD10	ICD9: 345; ICD10: G40-41 (2P or 1H or 1ER 1yr)	285	53	369	13
Reid 2012	Canada	≥18 years	LOW	Medical records	Physician claims	ICD9, ICD10	ICD9: 345; ICD10: G40-41 (2P 2yr)	262	32	390	36

Reid 2012	Canada	≥18 years	LOW	Medical records	Inpatients, physician claims	ICD9, ICD10	ICD9: 345; ICD10: G40-41 (2P or 1H 2yr)	265	32	390	33
Reid 2012	Canada	≥18 years	LOW	Medical records	Inpatients, A&E, physician claims	ICD9, ICD10	ICD9: 345; ICD10: G40-41 (2P or 1H or 1ER 2yr)	285	53	369	13
Reid 2012	Canada	≥18 years	LOW	Medical records	Physician claims	ICD9, ICD10	ICD9: 345, 780.3; ICD10: G40-41 (1P 2yr)	306	48	362	4
Reid 2012	Canada	≥18 years	LOW	Medical records	Inpatients, physician claims	ICD9, ICD10	ICD9: 345, 780.3; ICD10: G40-41 (1P or 1H 2yr)	306	50	360	4
Reid 2012	Canada	≥18 years	LOW	Medical records	Inpatients, A&E, physician claims	ICD9, ICD10	ICD9: 345, 780.3; ICD10: G40-41 (1P or 1H or 1ER 2yr)	308	59	351	2
Reid 2012	Canada	≥18 years	LOW	Medical records	Physician claims	ICD9, ICD10	ICD9: 345, 780.3; ICD10: G40-41 (2P 1yr)	282	33	377	28
Reid 2012	Canada	≥18 years	LOW	Medical records	Inpatients, physician claims	ICD9, ICD10	ICD9: 345, 780.3; ICD10: G40-41 (2P or 1H 1yr)	284	35	375	26
Reid 2012	Canada	≥18 years	LOW	Medical records	Inpatients, A&E, physician claims	ICD9, ICD10	ICD9: 345, 780.3; ICD10: G40-41 (2P or 1H or 1ER 1yr)	300	50	360	10
Reid 2012	Canada	≥18 years	LOW	Medical records	Physician claims	ICD9, ICD10	ICD9: 345, 780.3; ICD10: G40-41 (2P 2yr)	286	34	376	24
Reid 2012	Canada	≥18 years	LOW	Medical records	Inpatients, physician claims	ICD9, ICD10	ICD9: 345, 780.3; ICD10: G40-41 (2P or 1H 2yr)	288	36	374	22
Reid 2012	Canada	≥18 years	LOW	Medical records	Inpatients, A&E, physician claims	ICD9, ICD10	ICD9: 345, 780.3; ICD10: G40-41 (2P or 1H or 1ER 2yr)	300	50	360	10
Reid 2012	Canada	≥18 years	LOW	Medical records	Physician claims	ICD9, ICD10	ICD9: 345; ICD10: G40-41 (1P 2yr)	688	98	729	18
Reid 2012	Canada	≥18 years	LOW	Medical records	Inpatients, physician claims	ICD9, ICD10	ICD9: 345; ICD10: G40-41 (1P or 1H 2yr)	688	98	729	18
Reid 2012	Canada	≥18 years	LOW	Medical records	Inpatients, A&E, physician claims	ICD9, ICD10	ICD9: 345; ICD10: G40-41 (1P or 1H or 1ER 2yr)	703	146	681	3
Reid 2012	Canada	≥18 years	LOW	Medical records	Physician claims	ICD9, ICD10	ICD9: 345; ICD10: G40-41 (2P 1yr)	650	56	771	56
Reid 2012	Canada	≥18 years	LOW	Medical records	Inpatients, physician claims	ICD9, ICD10	ICD9: 345; ICD10: G40-41 (2P or 1H 1yr)	653	56	771	53
Reid 2012	Canada	≥18 years	LOW	Medical records	Inpatients, A&E, physician claims	ICD9, ICD10	ICD9: 345; ICD10: G40-41 (2P or 1H or 1ER 1yr)	700	131	696	6
Reid 2012	Canada	≥18 years	LOW	Medical records	Physician claims	ICD9, ICD10	ICD9: 345; ICD10: G40-41 (2P 2yr)	654	58	769	52
Reid 2012	Canada	≥18 years	LOW	Medical records	Inpatients, physician claims	ICD9, ICD10	ICD9: 345; ICD10: G40-41 (2P or 1H 2yr)	657	58	769	49
Reid 2012	Canada	≥18 years	LOW	Medical records	Inpatients, A&E, physician claims	ICD9, ICD10	ICD9: 345; ICD10: G40-41 (2P or 1H or 1ER 2yr)	701	131	696	5
Reid 2012	Canada	≥18 years	LOW	Medical records	Physician claims	ICD9, ICD10	ICD9: 345, 780.3; ICD10: G40-41 (1P 2yr)	701	110	711	11
Reid 2012	Canada	≥18 years	LOW	Medical records	Inpatients, physician claims	ICD9, ICD10	ICD9: 345, 780.3; ICD10: G40-41 (1P or 1H 2yr)	702	112	709	10
Reid 2012	Canada	≥18 years	LOW	Medical records	Inpatients, A&E, physician claims	ICD9, ICD10	ICD9: 345, 780.3; ICD10: G40-41 (1P or 1H or 1ER 2yr)	710	160	661	2
Reid 2012	Canada	≥18 years	LOW	Medical records	Physician claims	ICD9, ICD10	ICD9: 345, 780.3; ICD10: G40-41 (2P 1yr)	668	67	754	44
Reid 2012	Canada	≥18 years	LOW	Medical records	Inpatients, physician claims	ICD9, ICD10	ICD9: 345, 780.3; ICD10: G40-41 (2P or 1H 1yr)	672	70	751	40

Reid 2012	Canada	≥18 years	LOW	Medical records	Inpatients, A&E, physician claims	ICD9, ICD10	ICD9: 345, 780.3; ICD10: G40-41 (2P or 1H or 1ER 1yr)	707	144	677	5
Reid 2012	Canada	≥18 years	LOW	Medical records	Physician claims	ICD9, ICD10	ICD9: 345, 780.3; ICD10: G40-41 (2P 2yr)	672	68	753	40
Reid 2012	Canada	≥18 years	LOW	Medical records	Inpatients, physician claims	ICD9, ICD10	ICD9: 345, 780.3; ICD10: G40-41 (2P or 1H 2yr)	676	71	750	36
Reid 2012	Canada	≥18 years	LOW	Medical records	Inpatients, A&E, physician claims	ICD9, ICD10	ICD9: 345, 780.3; ICD10: G40-41 (2P or 1H or 1ER 2yr)	708	144	677	4
Tan 2015	Australia	≥16 years	LOW	Medical records	Inpatients, A&E, pharmacy	ICD10, AED	ICD10: G40-41 + ≥1AED	96	22	42580	62
Tu 2014	Canada	≥20 years	UNCLEAR	Medical records	Inpatients	ICD8, ICD9, ICD10	ICD8: 345; ICD9: 345 (excluding .2,.3); ICD10: G40 (H)	32	8	7397	63
Tu 2014	Canada	≥20 years	UNCLEAR	Medical records	Inpatients, physician claims	ICD8, ICD9, ICD10	ICD8: 345; ICD9: 345 (excluding .2,.3); ICD10: G40 (H or P)	86	148	7257	9
Tu 2014	Canada	≥20 years	UNCLEAR	Medical records	Physician claims (neurologist)	ICD8, ICD9, ICD10	ICD8: 345; ICD9: 345 (excluding .2,.3); ICD10: G40 (P neuro)	70	47	7358	25
Tu 2014	Canada	≥20 years	UNCLEAR	Medical records	Inpatients, physician claims	ICD8, ICD9, ICD10	ICD8: 345; ICD9: 345 (excluding .2,.3); ICD10: G40 (H or 2P 1yr)	80	52	7353	15
Tu 2014	Canada	≥20 years	UNCLEAR	Medical records	Inpatients, physician claims	ICD8, ICD9, ICD10	ICD8: 345; ICD9: 345 (excluding .2,.3); ICD10: G40 (H or 2P >30 days 1 yr)	75	34	7371	20
Tu 2014	Canada	≥20 years	UNCLEAR	Medical records	Inpatients, physician claims	ICD8, ICD9, ICD10	ICD8: 345; ICD9: 345 (excluding .2,.3); ICD10: G40 (H or 2P 2yr)	81	58	7347	14
Tu 2014	Canada	≥20 years	UNCLEAR	Medical records	Inpatients, physician claims	ICD8, ICD9, ICD10	ICD8: 345; ICD9: 345 (excluding .2,.3); ICD10: G40 (H or 2P >30 days 2 yr)	77	41	7364	18
Tu 2014	Canada	≥20 years	UNCLEAR	Medical records	Inpatients, physician claims, physician claims (neurologist)	ICD8, ICD9, ICD10	ICD8: 345; ICD9: 345 (excluding .2,.3); ICD10: G40 (H or 2P (1 neuro) >30 days 2 yr)	65	27	7378	30
Tu 2014	Canada	≥20 years	UNCLEAR	Medical records	Inpatients, physician claims (neurologist)	ICD8, ICD9, ICD10	ICD8: 345; ICD9: 345 (excluding .2,.3); ICD10: G40 (H or 2P (2 neuro) >30 days 2 yr)	61	20	7385	34
Tu 2014	Canada	≥20 years	UNCLEAR	Medical records	Inpatients, physician claims	ICD8, ICD9, ICD10	ICD8: 345; ICD9: 345 (excluding .2,.3); ICD10: G40 (H or 2P 3 yr)	82	59	7346	13
Tu 2014	Canada	≥20 years	UNCLEAR	Medical records	Inpatients, physician claims	ICD8, ICD9, ICD10	ICD8: 345; ICD9: 345 (excluding .2,.3); ICD10: G40 (H or 2P >30 days 3 yr)	78	42	7363	17
Tu 2014	Canada	≥20 years	UNCLEAR	Medical records	Inpatients, physician claims	ICD8, ICD9, ICD10	ICD8: 345; ICD9: 345 (excluding .2,.3); ICD10: G40 (H or 3P 1 yr)	71	32	7373	24
Tu 2014	Canada	≥20 years	UNCLEAR	Medical records	Inpatients, physician claims	ICD8, ICD9, ICD10	ICD8: 345; ICD9: 345 (excluding .2,.3); ICD10: G40 (H or 3P >30 days 1 yr)	67	17	7388	28
Tu 2014	Canada	≥20 years	UNCLEAR	Medical records	Inpatients, physician claims	ICD8, ICD9, ICD10	ICD8: 345; ICD9: 345 (excluding .2,.3); ICD10: G40 (H or 3P 2 yr)	73	34	7371	22
Tu 2014	Canada	≥20 years	UNCLEAR	Medical records	Inpatients, physician claims	ICD8, ICD9, ICD10	ICD8: 345; ICD9: 345 (excluding .2,.3); ICD10: G40 (H or 3P >30 days 2 yr)	70	18	7387	25
Tu 2014	Canada	≥20 years	UNCLEAR	Medical records	Inpatients, physician claims, physician claims (neurologist)	ICD8, ICD9, ICD10	ICD8: 345; ICD9: 345 (excluding .2,.3); ICD10: G40 (H or 3P (1 neuro) >30 days 2 yr)	62	16	7389	33
Tu 2014	Canada	≥20 years	UNCLEAR	Medical records	Inpatients, physician claims, physician claims (neurologist)	ICD8, ICD9, ICD10	ICD8: 345; ICD9: 345 (excluding .2,.3); ICD10: G40 (H or 3P (2 neuro) >30 days 2 yr)	59	14	7391	36
Tu 2014	Canada	≥20 years	UNCLEAR	Medical records	Inpatients, physician claims (neurologist)	ICD8, ICD9, ICD10	ICD8: 345; ICD9: 345 (excluding .2,.3); ICD10: G40 (H or 3P (3 neuro) >30 days 2 yr)	53	14	7391	42

Tu 2014	Canada	≥20 years	UNCLEAR	Medical records	Inpatients, physician claims	ICD8, ICD10	ICD9,	ICD8: 345; ICD9: 345 (excluding .2,.3); ICD10: G40 (H or 3P 2 yr)	73	34	7371	22
Tu 2014	Canada	≥20 years	UNCLEAR	Medical records	Inpatients, physician claims	ICD8, ICD10	ICD9,	ICD8: 345; ICD9: 345 (excluding .2,.3); ICD10: G40 (H or 3P >30 days 3 yr)	70	20	7385	25
Tu 2014	Canada	≥20 years	UNCLEAR	Medical records	Inpatients, physician claims	ICD8, ICD10	ICD9,	ICD8: 345; ICD9: 345 (excluding .2,.3); ICD10: G40 (H or 4P >30 days 2 yr)	65	15	7390	30
Tu 2014	Canada	≥20 years	UNCLEAR	Medical records	Inpatients, physician claims	ICD8, ICD10	ICD9,	ICD8: 345; ICD9: 345 (excluding .2,.3); ICD10: G40 (H or 4P >30 days 3 yr)	67	15	7390	28

B – Analysis

```
## Descriptive summary of studies with 91 algorithms.
## The confidence level for all calculations was set to 95%
## The analysis was performed using a continuity correction of 0.5 where applicable
##
## Diagnostic accuracies
##       Sens  2.5% 97.5%  Spec  2.5% 97.5%
##  [1,] 0.873 0.847 0.894 0.999 0.999 0.999
##  [2,] 0.807 0.668 0.897 0.998 0.997 0.999
##  [3,] 0.737 0.693 0.776 0.995 0.995 0.995
##  [4,] 0.720 0.689 0.748 0.959 0.958 0.961
##  [5,] 0.403 0.347 0.462 0.978 0.975 0.980
##  [6,] 0.839 0.771 0.890 0.983 0.946 0.995
##  [7,] 0.860 0.794 0.907 0.962 0.917 0.983
##  [8,] 0.914 0.858 0.950 0.726 0.648 0.792
##  [9,] 0.862 0.780 0.917 0.975 0.922 0.992
## [10,] 0.872 0.792 0.925 0.975 0.922 0.992
## [11,] 0.934 0.866 0.968 0.611 0.513 0.701
## [12,] 0.786 0.652 0.878 0.989 0.904 0.999
## [13,] 0.827 0.698 0.908 0.924 0.811 0.972
## [14,] 0.867 0.745 0.936 0.967 0.870 0.992
## [15,] 0.896 0.804 0.947 0.974 0.971 0.976
## [16,] 0.896 0.804 0.947 0.974 0.971 0.976
## [17,] 0.854 0.755 0.918 0.998 0.997 0.998
## [18,] 0.896 0.804 0.947 0.974 0.971 0.976
## [19,] 0.632 0.517 0.734 0.997 0.996 0.997
## [20,] 0.854 0.755 0.918 0.998 0.997 0.998
## [21,] 0.632 0.517 0.734 0.997 0.996 0.997
## [22,] 0.854 0.755 0.918 0.998 0.997 0.998
## [23,] 0.424 0.316 0.539 1.000 0.999 1.000
## [24,] 0.882 0.787 0.938 0.997 0.996 0.998
## [25,] 0.854 0.755 0.918 0.997 0.996 0.998
## [26,] 0.368 0.266 0.483 0.997 0.996 0.998
## [27,] 0.854 0.755 0.918 0.997 0.996 0.998
## [28,] 0.854 0.755 0.918 0.997 0.996 0.998
## [29,] 0.882 0.787 0.938 0.997 0.995 0.997
## [30,] 0.769 0.722 0.810 0.924 0.905 0.940
## [31,] 0.817 0.773 0.853 0.938 0.920 0.952
## [32,] 0.811 0.767 0.848 0.941 0.924 0.955
## [33,] 0.962 0.933 0.978 0.874 0.838 0.902
## [34,] 0.962 0.933 0.978 0.874 0.838 0.902
## [35,] 0.982 0.959 0.992 0.850 0.813 0.881
## [36,] 0.861 0.817 0.896 0.926 0.897 0.947
## [37,] 0.871 0.829 0.905 0.926 0.897 0.947
## [38,] 0.955 0.925 0.973 0.874 0.838 0.902
## [39,] 0.878 0.836 0.910 0.923 0.894 0.945
## [40,] 0.888 0.847 0.919 0.923 0.894 0.945
## [41,] 0.955 0.925 0.973 0.874 0.838 0.902
## [42,] 0.986 0.965 0.994 0.882 0.847 0.910
## [43,] 0.986 0.965 0.994 0.877 0.842 0.905
## [44,] 0.992 0.974 0.998 0.855 0.818 0.886
## [45,] 0.908 0.871 0.936 0.918 0.888 0.941
## [46,] 0.915 0.878 0.941 0.914 0.882 0.937
## [47,] 0.966 0.940 0.981 0.877 0.842 0.905
## [48,] 0.921 0.886 0.946 0.916 0.885 0.939
## [49,] 0.928 0.893 0.952 0.911 0.880 0.935
## [50,] 0.966 0.940 0.981 0.877 0.842 0.905
## [51,] 0.974 0.959 0.983 0.881 0.857 0.901
## [52,] 0.974 0.959 0.983 0.881 0.857 0.901
## [53,] 0.995 0.987 0.998 0.823 0.796 0.848
## [54,] 0.920 0.898 0.938 0.932 0.913 0.947
## [55,] 0.924 0.902 0.942 0.932 0.913 0.947
## [56,] 0.991 0.981 0.996 0.841 0.815 0.864
## [57,] 0.926 0.904 0.943 0.929 0.910 0.945
## [58,] 0.930 0.909 0.947 0.929 0.910 0.945
## [59,] 0.992 0.983 0.997 0.841 0.815 0.864
## [60,] 0.984 0.972 0.991 0.866 0.841 0.887
## [61,] 0.985 0.973 0.992 0.863 0.838 0.885
## [62,] 0.996 0.989 0.999 0.805 0.776 0.830
## [63,] 0.938 0.917 0.953 0.918 0.897 0.935
## [64,] 0.943 0.924 0.958 0.914 0.893 0.932
## [65,] 0.992 0.983 0.997 0.824 0.797 0.849
## [66,] 0.943 0.924 0.958 0.917 0.896 0.934
## [67,] 0.949 0.930 0.963 0.913 0.892 0.930
## [68,] 0.994 0.985 0.997 0.824 0.797 0.849
## [69,] 0.607 0.529 0.679 0.999 0.999 1.000
```

```
## [70,] 0.339 0.252 0.438 0.999 0.998 0.999
## [71,] 0.901 0.825 0.946 0.980 0.976 0.983
## [72,] 0.734 0.638 0.812 0.994 0.991 0.995
## [73,] 0.839 0.752 0.899 0.993 0.991 0.995
## [74,] 0.786 0.694 0.857 0.995 0.994 0.997
## [75,] 0.849 0.764 0.907 0.992 0.990 0.994
## [76,] 0.807 0.717 0.874 0.994 0.992 0.996
## [77,] 0.682 0.584 0.767 0.996 0.995 0.997
## [78,] 0.641 0.541 0.729 0.997 0.996 0.998
## [79,] 0.859 0.776 0.915 0.992 0.990 0.994
## [80,] 0.818 0.729 0.882 0.994 0.992 0.996
## [81,] 0.745 0.649 0.821 0.996 0.994 0.997
## [82,] 0.703 0.605 0.785 0.998 0.996 0.999
## [83,] 0.766 0.672 0.839 0.995 0.994 0.997
## [84,] 0.734 0.638 0.812 0.998 0.996 0.998
## [85,] 0.651 0.552 0.739 0.998 0.996 0.999
## [86,] 0.620 0.520 0.711 0.998 0.997 0.999
## [87,] 0.557 0.458 0.653 0.998 0.997 0.999
## [88,] 0.766 0.672 0.839 0.995 0.994 0.997
## [89,] 0.734 0.638 0.812 0.997 0.996 0.998
## [90,] 0.682 0.584 0.767 0.998 0.997 0.999
## [91,] 0.703 0.605 0.785 0.998 0.997 0.999
##
## Test for equality of sensitivities:
## X-squared = 3917.026, df = 90, p-value = <2e-16
## Test for equality of specificities:
## X-squared = 120443.5, df = 90, p-value = <2e-16
##
##
## Diagnostic OR and likelihood ratios
##           DOR    2.5%    97.5%    posLR    2.5%    97.5% negLR  2.5% 97.5%
##  [1,] 6049.339 4868.175 7517.087  772.023 735.851  809.974 0.128 0.106 0.154
##  [2,] 2457.222 1040.092 5805.198  475.497 304.284  743.049 0.194 0.106 0.354
##  [3,]  555.097  440.479  699.540  146.986 132.816  162.668 0.265 0.226 0.310
##  [4,]   60.484   52.165   70.128   17.683  16.811   18.600 0.292 0.263 0.324
##  [5,]   30.187   23.104   39.442   18.416  15.340   22.110 0.610 0.553 0.672
##  [6,]  295.043   78.337 1111.233   48.329  14.114  165.482 0.164 0.113 0.237
##  [7,]  154.162   58.341  407.358   22.506   9.890   51.213 0.146 0.098 0.218
##  [8,]   28.255   14.234   56.087    3.333   2.544    4.368 0.118 0.069 0.202
##  [9,]  241.607   60.739  961.063   34.145  10.016  116.397 0.141 0.086 0.232
## [10,]  264.024   65.844 1058.697   34.549  10.137  117.749 0.131 0.078 0.220
## [11,]   22.121    9.063   53.993    2.401   1.865    3.091 0.109 0.051 0.232
## [12,]  333.667   18.931 5880.864   72.286   4.572 1142.763 0.217 0.127 0.371
## [13,]   57.857   15.501  215.955   10.863   3.936   29.979 0.188 0.101 0.348
## [14,]  193.974   31.289 1202.542   26.599   5.491  128.851 0.137 0.067 0.281
## [15,]  323.481  150.896  693.455   34.592  30.551   39.167 0.107 0.054 0.211
## [16,]  323.481  150.896  693.455   34.592  30.551   39.167 0.107 0.054 0.211
## [17,] 2652.437 1272.132 5530.418  387.668 274.084  548.322 0.146 0.084 0.256
## [18,]  323.481  150.896  693.455   34.592  30.551   39.167 0.107 0.054 0.211
## [19,]  510.369  294.375  884.843  188.476 136.521  260.202 0.369 0.273 0.500
## [20,] 2652.437 1272.132 5530.418  387.668 274.084  548.322 0.146 0.084 0.256
## [21,]  510.369  294.375  884.843  188.476 136.521  260.202 0.369 0.273 0.500
## [22,] 2652.437 1272.132 5530.418  387.668 274.084  548.322 0.146 0.084 0.256
## [23,] 1533.623  652.251 3605.974  884.387 411.718 1899.698 0.577 0.473 0.703
## [24,] 2880.788 1321.309 6280.848  340.975 247.852  469.087 0.118 0.063 0.223
## [25,] 1885.094  924.409 3844.163  275.764 204.953  371.040 0.146 0.084 0.256
## [26,]  183.650  105.503  319.681  116.425  77.184  175.616 0.634 0.531 0.756
## [27,] 1885.094  924.409 3844.163  275.764 204.953  371.040 0.146 0.084 0.256
## [28,] 1885.094  924.409 3844.163  275.764 204.953  371.040 0.146 0.084 0.256
## [29,] 2138.850  996.623 4590.182  253.385 191.854  334.650 0.118 0.063 0.223
## [30,]   40.680   28.719   57.622   10.170   8.043   12.860 0.250 0.207 0.302
## [31,]   66.913   45.763   97.838   13.093  10.124   16.934 0.196 0.157 0.244
## [32,]   68.342   46.608  100.209   13.733  10.539   17.894 0.201 0.162 0.249
## [33,]  172.664   89.649  332.548    7.602   5.912    9.776 0.044 0.025 0.078
## [34,]  172.664   89.649  332.548    7.602   5.912    9.776 0.044 0.025 0.078
## [35,]  302.114  124.722  731.807    6.539   5.210    8.207 0.022 0.009 0.050
## [36,]   77.117   47.285  125.770   11.565   8.239   16.232 0.150 0.113 0.199
## [37,]   84.095   51.196  138.133   11.699   8.338   16.417 0.139 0.103 0.187
## [38,]  146.061   78.839  270.597    7.550   5.870    9.710 0.052 0.031 0.087
## [39,]   86.412   52.521  142.172   11.427   8.190   15.942 0.132 0.097 0.179
## [40,]   95.226   57.348  158.124   11.557   8.286   16.120 0.121 0.088 0.167
## [41,]  146.061   78.839  270.597    7.550   5.870    9.710 0.052 0.031 0.087
## [42,]  509.078  191.493 1353.366    8.352   6.410   10.882 0.016 0.007 0.041
## [43,]  486.219  183.183 1290.562    8.021   6.193   10.389 0.016 0.007 0.041
## [44,]  728.993  203.791 2607.726    6.852   5.416    8.669 0.009 0.003 0.032
## [45,]  111.698   66.227  188.390   11.144   8.040   15.447 0.100 0.070 0.142
## [46,]  113.558   67.094  192.199   10.591   7.720   14.531 0.093 0.065 0.134
## [47,]  204.300  103.277  404.141    7.864   6.069   10.190 0.038 0.021 0.070
```

```
## [48,]  127.616   74.357  219.020   10.975   7.961   15.129 0.086 0.059 0.126
## [49,]  131.560   76.108  227.414   10.446   7.652   14.259 0.079 0.053 0.118
## [50,]  204.300  103.277  404.141    7.864   6.069   10.190 0.038 0.021 0.070
## [51,]  275.627  165.938  457.823    8.186   6.798    9.857 0.030 0.019 0.047
## [52,]  275.627  165.938  457.823    8.186   6.798    9.857 0.030 0.019 0.047
## [53,]  935.027  322.231 2713.196    5.624   4.855    6.514 0.006 0.002 0.017
## [54,]  157.212  107.165  230.633   13.484  10.474   17.359 0.086 0.067 0.110
## [55,]  166.794  113.139  245.893   13.546  10.522   17.438 0.081 0.063 0.105
## [56,]  570.808  257.899 1263.369    6.239   5.333    7.299 0.011 0.005 0.024
## [57,]  163.985  111.371  241.454   13.103  10.226   16.789 0.080 0.062 0.104
## [58,]  174.720  117.993  258.721   13.163  10.273   16.865 0.075 0.058 0.099
## [59,]  675.554  286.027 1595.563    6.248   5.340    7.309 0.009 0.004 0.021
## [60,]  392.774  212.099  727.355    7.319   6.152    8.708 0.019 0.010 0.033
## [61,]  421.946  222.266  801.014    7.199   6.062    8.549 0.017 0.009 0.031
## [62,] 1171.329  334.356 4103.442    5.104   4.442    5.864 0.004 0.001 0.015
## [63,]  167.918  113.404  248.638   11.418   9.078   14.361 0.068 0.051 0.090
## [64,]  177.002  118.623  264.109   10.997   8.791   13.757 0.062 0.046 0.084
## [65,]  603.122  255.734 1422.399    5.645   4.867    6.546 0.009 0.004 0.022
## [66,]  182.654  122.171  273.080   11.318   9.016   14.209 0.062 0.046 0.084
## [67,]  194.545  128.848  293.739   10.908   8.735   13.621 0.056 0.041 0.077
## [68,]  738.191  287.177 1897.529    5.653   4.874    6.555 0.008 0.003 0.019
## [69,] 2921.969 1734.349 4922.826 1149.179 746.347 1769.435 0.393 0.324 0.477
## [70,]  445.426  201.257  985.827  294.969 142.470  610.702 0.662 0.574 0.764
## [71,]  444.993  223.348  886.592   44.937  37.818   53.395 0.101 0.055 0.185
## [72,]  428.297  250.769  731.502  114.501  84.153  155.792 0.267 0.192 0.373
## [73,]  727.443  396.183 1335.678  118.290  89.092  157.058 0.163 0.103 0.257
## [74,]  786.919  435.460 1422.040  168.826 119.107  239.300 0.215 0.146 0.315
## [75,]  705.949  381.657 1305.791  107.477  82.142  140.625 0.152 0.095 0.245
## [76,]  743.404  411.323 1343.592  144.068 104.746  198.151 0.194 0.129 0.292
## [77,]  576.205  325.885 1018.801  183.747 123.509  273.365 0.319 0.238 0.428
## [78,]  642.217  351.953 1171.871  231.438 146.467  365.703 0.360 0.276 0.471
## [79,]  754.542  402.086 1415.951  106.967  82.009  139.520 0.142 0.086 0.232
## [80,]  777.190  426.760 1415.372  142.493 104.062  195.117 0.183 0.120 0.280
## [81,]  662.110  373.022 1175.239  169.721 118.117  243.870 0.256 0.182 0.361
## [82,]  999.947  526.389 1899.536  297.562 183.083  483.624 0.298 0.219 0.405
## [83,]  697.978  391.315 1244.964  164.354 115.723  233.422 0.235 0.164 0.338
## [84,] 1104.014  580.462 2099.790  293.988 183.606  470.730 0.266 0.191 0.371
## [85,]  835.538  440.618 1584.420  292.219 176.580  483.589 0.350 0.266 0.460
## [86,]  830.976  429.603 1607.347  316.564 184.930  541.895 0.381 0.295 0.492
## [87,]  641.696  333.677 1234.048  284.642 165.173  490.522 0.444 0.354 0.555
## [88,]  697.978  391.315 1244.964  164.354 115.723  233.422 0.235 0.164 0.338
## [89,]  996.036  532.111 1864.437  265.306 169.385  415.548 0.266 0.191 0.371
## [90,] 1023.961  530.300 1977.179  326.003 194.650  545.996 0.318 0.237 0.427
## [91,] 1129.278  581.708 2192.286  335.958 200.932  561.719 0.297 0.219 0.405
##
## Correlation of sensitivities and false positive rates:
##    rho  2.5 % 97.5 %
##  0.601  0.451  0.718
```

Forest Plot - Sensitivity



Forest Plot - Specificity



Bivariate Reitsma analysis

```
## Call:  reitsma.default (data = algorithms)
##
## Bivariate diagnostic random-effects meta-analysis
## Estimation method: REML
##
## Fixed-effects coefficients
##                   Estimate Std. Error       z Pr(>|z|) 95%ci.lb 95%ci.ub
## tsens.(Intercept)    2.053      0.136  15.144    0.000    1.788    2.319 ***
## tfpr.(Intercept)    -3.836      0.194 -19.763    0.000   -4.216   -3.455 ***
## sensitivity          0.886          -       -        -    0.857    0.910
## false pos. rate      0.021          -       -        -    0.015    0.031
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Variance components: between-studies Std. Dev and correlation matrix
##       Std. Dev tsens  tfpr
## tsens    1.262 1.000     .
## tfpr     1.837 0.796 1.000
##
##   logLik      AIC      BIC
##  288.557 -567.115 -551.095
##
## AUC:  0.974
## Partial AUC (restricted to observed FPRs and normalized):  0.957
##
## HSROC parameters
##       Theta      Lambda        beta sigma2theta sigma2alpha
##      -0.351       5.657       0.375       2.082       0.944
```

Appendix 3 – Supplementary material to Chapter 4

Appendix A - Scottish mortality data

Approximately 55,000 deaths are registered in Scotland annually.[59] Death certificates, completed by a patient's clinician, record cause of death in two free-text sections (Parts 1–2). Part 1 lists the immediate cause of death and works back in time by listing the diseases and conditions that lead to this (Parts 1A–D). Part 2 lists other significant diseases and conditions that contributed to the occurrence of the death (i.e. it is not simply a list of comorbidities).[195] An automated method of coding, based on a complex set of modification and selection rules, is then used alongside experienced coding staff at NRS to convert the free-text causes of death in Parts 1–2 into their corresponding ICD-10 codes within the NRS death records.[196] NRS arrange these ICD-10 codes into a single primary cause of death field and up to nine secondary cause of death fields. The primary cause of death field contains an ICD-10 code for the underlying cause of death, which is defined as the disease or condition which initiated the chain of morbid events leading directly to death. This can be taken from Part 1 or 2 of the death certificate, according to the modification and selection rules.[196] This final coding process is unique in Scotland because where there are discrepancies in the accuracy of cause of death information in the initial death certificates, NRS specialists make amendments to the cause of death through a process of medical enquiries to the certifying doctor, hospital records, and the Procurator Fiscal. The Procurator Fiscal is a statutory legal official that investigates deaths that are accidental, unexpected, unexplained, sudden or suspicious (including ordering a post mortem, where necessary).[196] This process of NRS medical enquiries helps insure accuracy within the final NRS electronic death certificate free-text and ICD-10 coding for statistical purposes and research, which are used to develop public health policies.[31] In a recent review of 6,000 Scottish death certificates (10% of all Scottish death certificates), three percent had incorrect cause of death and required either an amendment or replacement death certificate.[181]

Appendix B – Antiepileptic drug filters

B1

Broad filtering list of antiepileptic drugs

	Drug Name
1	Acetazolamide
2	Adrenocorticotropic hormone (incl. Tetracosactide)
3	Carbamazepine
4	Clobazam
5	Clonazepam
6	Diazepam
7	Eslicarbazepine
8	Ethosuximide
9	Felbamate
10	Fosphenytoin
11	Gabapentin
12	Lacosamide
13	Lamotrigine
14	Levetiracetam
15	Lorazepam
16	Mepenzolate bromide
17	Methsuximide
18	Midazolam
19	Nitrazepam
20	Oxcarbazepine
21	Perampanel
22	Phenobarbital
23	Phenytoin
24	Piracetam
25	Pregabalin
26	Primidone
27	Retigabine
28	Rufinamide
29	Sodium valproate
30	Steripentol
31	Sultiame
32	Temazepam
33	Tiagabine
34	Topiramate
35	Vigabatrin
36	Zonisamide

B2

Narrow filtering list of antiepileptic drugs

	Drug Name
1	Carbamazepine
2	Clobazam
3	Eslicarbazepine
4	Ethosuximide
5	Gabapentin
6	Lacosamide
7	Lamotrigine
8	Levetiracetam
9	Oxcarbazepine
10	Perampanel
11	Phenobarbital
12	Phenytoin
13	Piracetam
14	Primidone
15	Retigabine
16	Rufinamide
17	Sodium valproate
18	Tiagabine
19	Topiramate
20	Vigabatrin
21	Zonisamide

Appendix C – Primary care data

Primary care dataset

GP practices use Vision medical software to build electronic health records (EHRs) for patients in primary care. Within this system, GPs record symptoms, diagnoses and prescriptions using Read codes, which are a coded thesaurus of clinical terms.[126] This makes it possible to track diagnostic primary care NHS data within a data linkage study environment. In the current study, Albasoft Ltd[180] made contact with all of the GP practices in Scotland and electronically extracted an agreed list of epilepsy-related data variables (see below) from consenting GP practices and transferred them to eDRIS for linkage. Extraction was completed when 100 GP practices were recruited. The study was limited to this number of practices by monetary resource limitations. We extracted all available *F25* Read codes and their dates for each participant recruited.

Primary care recruitment diagnostic filtration codes[126,166]

Read Code	*Description*	*Event count*
F25..	Epilepsy	3151
F2516	Grand mal seizure	469
F25z.	Epilepsy NOS	494
F2510	Grand mal (major) epilepsy	238
F253.	Grand mal status	86
F2514	Epileptic seizures - tonic	67
F2512	Epileptic seizures - clonic	114
F2540	Temporal lobe epilepsy	57
F2515	Tonic-clonic epilepsy	62
F2545	Complex partial epileptic seizure	60
F2500	Petit mal (minor) epilepsy	110
F25X.	Status epilepticus, unspecified	43
F2550	Jacksonian, focal or motor epilepsy	52
F2556	Simple partial epileptic seizure	28
F258.	Post-ictal state	27
F251z	Generalised convulsive epilepsy NOS	26
F251.	Generalised convulsive epilepsy	21
F254z	Partial epilepsy with impairment of consciousness NOS	14
F2513	Epileptic seizures - myoclonic	14
F25B.	Alcohol-induced epilepsy	10
F2503	Epileptic seizures - akinetic	8

F252.	Petit mal status	7
F254.	Partial epilepsy with impairment of consciousness	≤5
F25y.	Other forms of epilepsy	6
F251y	Other specified generalised convulsive epilepsy	11
F25H.	Generalised seizure	≤5
F25y2	Locationally-related (focal) (partial) idiopathic epilepsy and epileptic syndrome with seizures of local onset	≤5
F2502	Epileptic seizures - atonic	≤5
F255.	Partial epilepsy without impairment of consciousness	9
F25y3	Complex partial status epilepticus	6
F2544	Epileptic automatism	≤5
F25yz	Other forms of epilepsy NOS	≤5
F255z	Partial epilepsy without impairment of consciousness NOS	≤5
F256.	Infantile spasms	≤5
F2560	Hypsarrhythmia	≤5
F259.	Early infant epileptic encephalopathy with suppression bursts	≤5
F25F.	Photosensitive epilepsy	≤5
Total		**5218**

Appendix 4 – Supplementary material to Chapter 5

Appendix S1 – Primary care recruitment

A trusted third party (Albasoft Ltd)[180] made contact with all of the GP practices in Scotland and electronically extracted an agreed list of epilepsy-related data variables (see below) from consenting GP practices and transferred them to eDRIS for linkage. Extraction was completed when the first 100 GP practices were recruited. The study was limited to this number of practices by monetary resource limitations. We extracted all available *F25* Read codes and their dates for each participant recruited (S2). The burden of comorbid mental health problems (including depression, psychiatric illness and suicidality), learning disabilities and congenital abnormalities, and substance misuse (including alcohol and drug abuse) was identified by extracting codes for these conditions within the comorbidity and other primary care activity codes captured (S3).

Appendix S2 – Epilepsy Read codes used

Diagnostic filtration codes for capturing epilepsy in primary care

Read Code	*Description*	*Event count*
F25..	Epilepsy	3151
F2516	Grand mal seizure	469
F25z.	Epilepsy NOS	494
F2510	Grand mal (major) epilepsy	238
F253.	Grand mal status	86
F2514	Epileptic seizures - tonic	67
F2512	Epileptic seizures - clonic	114
F2540	Temporal lobe epilepsy	57
F2515	Tonic-clonic epilepsy	62
F2545	Complex partial epileptic seizure	60
F2500	Petit mal (minor) epilepsy	110
F25X.	Status epilepticus, unspecified	43
F2550	Jacksonian, focal or motor epilepsy	52
F2556	Simple partial epileptic seizure	28
F258.	Post-ictal state	27
F251z	Generalised convulsive epilepsy NOS	26
F251.	Generalised convulsive epilepsy	21
F254z	Partial epilepsy with impairment of consciousness NOS	14
F2513	Epileptic seizures - myoclonic	14
F25B.	Alcohol-induced epilepsy	10
F2503	Epileptic seizures - akinetic	8
F252.	Petit mal status	7
F254.	Partial epilepsy with impairment of consciousness	≤5
F25y.	Other forms of epilepsy	6
F251y	Other specified generalised convulsive epilepsy	11

F25H.	Generalised seizure	≤5
F25y2	Locationally-related (focal) (partial) idiopathic epilepsy and epileptic syndrome with seizures of local onset	≤5
F2502	Epileptic seizures - atonic	≤5
F255.	Partial epilepsy without impairment of consciousness	9
F25y3	Complex partial status epilepticus	6
F2544	Epileptic automatism	≤5
F25yz	Other forms of epilepsy NOS	≤5
F255z	Partial epilepsy without impairment of consciousness NOS	≤5
F256.	Infantile spasms	≤5
F2560	Hypsarrhythmia	≤5
F259.	Early infant epileptic encephalopathy with suppression bursts	≤5
F25F.	Photosensitive epilepsy	≤5
Total		**5218**

Appendix S3 – Comorbidity Read codes used

Comorbidity and other primary care activity codes captured

Read code	Description	Event count	*Relevant comorbidity group indicator*
2469.	O/E - Systolic BP reading	25692	
246A.	O/E - Diastolic BP reading	25686	
9N31.	Telephone encounter	18990	
9N11.	Seen in GP's surgery	17337	
8B314	Medication review	16449	
246..	O/E - blood pressure reading	16268	
9N1C.	Seen in own home	16267	
44J3.	Serum creatinine	15622	
9NZ..	Patient encounter data NOS	13101	
22A..	O/E - weight	12348	
44I5.	Serum sodium	10021	
44I4.	Serum potassium	9657	
22K..	Body Mass Index	8573	
9Na..	Consultation	8364	
9N36.	Letter from specialist	7881	
44F..	Serum alkaline phosphatase	7568	
44M4.	Serum albumin	7124	
8H...	Referral for further care	6789	
42QE.	International normalised ratio	6682	
8BIF.	Epilepsy medication review	6339	
44JB.	Urea and electrolytes	5143	
424..	Full blood count - FBC	4988	
9c0C.	Result	4511	
9Z...	Administration NOS	4485	
44J9.	Serum urea level	4157	
1371.	Never smoked tobacco	4154	
44E..	Serum bilirubin level	4109	
8B3V.	Medication review done	4091	
9....	Administration	3793	
44D6.	Liver function test	3710	
137S.	Ex smoker	3378	
44G4.	Gamma - G.T. level	3261	
667F.	Seizure free >12 months	3241	
8B311	Medication given	2900	
6A...	Patient reviewed	2900	
9N33.	Letter encounter from patient	2886	

8B3H.	Medication requested	2827	
679..	Health education - subject	2340	
413..	Laboratory test requested	2321	
137R.	Current smoker	2306	
44GB.	Serum alanine aminotransferase level	2294	
44EC.	Serum total bilirubin level	2202	
9344.	Notes summary on computer	1937	
6896.	Depression screening using questions	1915	
8B3S.	Medication review	1841	
9b0w.	OOH report	1824	
41...	Laboratory procedures -general	1763	
9N32.	Third party encounter	1754	
44G9.	Serum gamma-glutamyl transferase level	1691	
41D0.	Blood sample taken	1679	
667..	Epilepsy monitoring	1654	
44HB.	AST serum level	1563	
8CB..	Had a chat to patient	1506	
6675.	Fit frequency	1422	
136L.	Alcohol intake within recommended sensible limits	1372	
136..	Alcohol consumption	1368	
667Q.	1 to 12 seizures a year	1254	
8C1B.	Nursing care blood sample taken	1224	
6676.	Last fit	1209	
9N42.	Did not attend - no reason	1199	
H06z0	Chest infection NOS	1133	
9N3D.	Letter received	1069	
4145.	Blood sample -> Lab NOS	1025	
9124.	Pat. GP7B/GP8B card from HB	1021	
9NC3.	Letter sent to patient	1006	
137P.	Cigarette smoker	886	
8BL3.	Patient on maximal tolerated anticonvulsant therapy	873	
93A..	Discharge summary	837	
Z4A..	Discussion	835	
8HE..	Discharged from hospital	830	
1Z...	History/symptom NOS	827	
414..	Sample sent to lab. for test	793	
9RZ..	Patient admin. data NOS	777	
9N4..	Failed encounter	769	
9Of0.	Epilepsy screen invite 1	761	
C10F.	Type 2 diabetes mellitus	752	
R003z	[D]Convulsion NOS	707	
G66..	Stroke and cerebrovascular accident unspecified	695	
9125.	Patient MRE received from HB	674	
692B.	Unfit to drive	643	
663H.	Inhaler technique - good	638	
6792.	Health ed. - alcohol	634	Alcohol abuse
8B313	Medication commenced	619	
ZG52.	Advice about investigation results	612	
1362.	Trivial drinker - <1u/day	587	
41C..	Patient informed - test result	576	
8H2..	Emergency hospital admission	556	
9ND5.	Consultant letter from:	534	
4....	Laboratory procedures	528	
9Of5.	Epilepsy monitoring call first letter	527	
6A5..	Ongoing review	516	
9N35.	Letter encounter to patient	510	
22J..	O/E - dead	509	
9N2R.	Seen by co-operative doctor	508	
4143.	Blood sample -> Biochem Lab	494	
67I..	Advice	486	
8B3x.	Medication review with patient	479	

9h61.	Excepted from epilepsy quality indicators: Patient unsuitabl	474	
H06z1	Lower resp tract infection	444	
9123.	Patient reg. form sent to HB	438	
8CMM.	Has anticipatory care plan	434	
ZV68z	[V]Unspecified administrative encounter	431	
6A6..	Mental health review	411	Mental health problems
9b0g.	Referral letter	410	
8A...	Monitoring of patient	408	
E23..	Alcohol dependence syndrome	392	Alcohol abuse
136M.	Current non drinker	392	
TC...	Accidental falls	376	
E2B..	Depressive disorder NEC	372	Mental health problems
9N38.	Message from patient	372	
924..	Patient de-reg. MRE to HB	371	
9b0c.	Out of hours non practice note	369	
44W3.	Serum phenytoin level	360	
9N...	Patient encounter admin. data	350	
8CM3.	Palliative care plan review	344	
9122.	Patient signed reg. form	334	
9N1y.	Seen in other clinic	332	
22K2.	Body Mass Index high K/M2	319	
5675.	CAT scan - brain	318	
67H..	Lifestyle counselling	317	
412..	Laboratory procedure performed	315	
H33..	Asthma	313	
6671.	Initial epilepsy assessment	306	
1z...	Read Code Administration	303	
663Q.	Asthma not limiting activities	300	
44W4.	Serum sodium valproate level	300	
9OL..	Diabetes monitoring admin.	294	
9D1..	MED3 - doctor's statement	294	
9234.	GP22-death	291	
C10..	Diabetes mellitus	286	
44W5.	Serum carbamazepine level	286	
G65..	Transient cerebral ischaemia	284	
388P.	HAD scale: depression score	282	
9N0l.	Seen in out of hours centre	273	
13CA.	Housebound	270	
9Of1.	Epilepsy screen invite 2	263	
8CAM.	Patient advised about alcohol	263	Alcohol abuse
H33zz	Asthma NOS	262	
66AS.	Diabetic annual review	262	
9h62.	Excepted from epilepsy quality indicators: Informed dissent	262	
388N.	HAD scale: anxiety score	260	
1374.	Moderate smoker - 10-19 cigs/d	259	
9k16.	Alcohol screen - fast alcohol screening test completed	258	
6672.	Follow-up epilepsy assessment	258	
14Z1.	No relevant past medical hist.	253	
4142.	Blood sample -> Haematol Lab	252	
9N1G.	Seen in nursing home	251	
8CS..	Agreement of care plan	249	
9b0Z.	OOH attendance note	249	
E200.	Anxiety states	248	Mental health problems
137L.	Current non-smoker	248	
136V.	Alcohol units per week	243	
667R.	2 to 4 seizures a month	242	
8BM0.	Mental health medication review	242	Mental health problems
667X.	No epilepsy drug side effects	239	
136Z.	Alcohol consumption NOS	237	

9D11.	MED3 issued to patient	237	
137..	Tobacco consumption	237	
1B17.	Depressed	234	Mental health problems
8B318	Medication reconciliation	233	
677B.	Advice about treatment given	226	
6781.	Health education offered	226	
9R8..	Date records held from	225	
9e3..	Out of hours patient call detail received	225	
9Of6.	Epilepsy monitoring call second letter	222	
9hC0.	Excepted from depression quality indicators: Patient unsuita	220	Mental health problems
8CAK.	Patient given telephone advice out of hours	218	
1367.	Stopped drinking alcohol	217	
136K.	Alcohol intake above recommended sensible limits	216	Alcohol abuse
1BT..	Depressed mood	216	Mental health problems
9N4F.	Failed encounter - message left on answer machine	216	
8B316	Medication changed	215	
388u.	Fast alcohol screening test	214	
667S.	1 to 7 seizures a week	212	
8B23.	Drug addiction therapy	210	Mental health problems
13G1.	District nurse attends	209	
9b0Q.	Laboratory result	209	
94B..	Cause of death	208	
934..	Computer record	205	
8Hd..	Admission to hospital	204	
22K5.	Body mass index 30+ - obesity	204	
1375.	Heavy smoker - 20-39 cigs/day	198	
1363.	Light drinker - 1-2u/day	198	
8BIH.	Medication review done by doctor	196	
44f..	Serum glucose level	195	
9N9..	Ambulance request for patient	194	
E24..	Drug dependence	193	Drug abuse
388f.	Patient health questionnaire (PHQ-9) score	187	
22K1.	Body Mass Index normal K/M2	185	
918e.	On learning disability register	183	Learning disabilities and congenital abnormalities
6A1..	Patient reviewed at hospital	182	
8B3A.	Drug dosage altered	179	
9N3A.	Telephone triage encounter	179	
H3z..	Chronic obstructive airways disease NOS	178	
G6...	Cerebrovascular disease	175	
8B31B	Polypharmacy medication review	174	
22K8.	Body mass index 20-24 - normal	173	
662M.	Stroke monitoring	172	
Eu01.	[X]Vascular dementia	171	
8B36.	Patient on numerous drugs	171	Drug abuse
2126.	Patient's condition improved	171	
C109.	Non-insulin dependent diabetes mellitus	170	
9N19.	Seen in hospital casualty	170	
8CM2.	Psychiatry care plan	168	Mental health problems
16D..	Falls	167	
8BMC.	Prescription collected by pharmacy	167	
Eu02z	[X] Unspecified dementia	163	
8BIQ.	Equivalent quantities for all medication checked	163	
G30..	Acute myocardial infarction	162	
663..	Respiratory disease monitoring	161	
9NF4.	Visit out of hours	159	
Eu32.	[X]Depressive episode	158	Mental health problems
9N4H.	DNA hospital appointment	153	
66AI.	Diabetic - good control	152	
22K4.	Body mass index index 25-29 - overweight	151	
Eu32z	[X]Depressive episode, unspecified	150	Mental health problems

388m.	Mini-mental state examination	148	
41G..	Laboratory administration procedures	147	
8B3A1	Medication increased	146	
667M.	Epilepsy management plan given	144	
9c0B.	Referral	144	
G64..	Cerebral arterial occlusion	143	
G580.	Congestive heart failure	142	
9N4C.	Failed encounter - no answer when rang back	142	
9Of2.	Epilepsy screen invite 3	134	
692A.	Fit to drive	134	
918F.	Has a carer	131	
1Y...	Patient feels well	131	
8B3h.	Medication review without patient	130	
1364.	Moderate drinker - 3-6u/day	130	
9Of7.	Epilepsy monitoring call third letter	129	
9NC1.	Letter sent to consultant	128	
9b0x.	GP to GP generated transaction	127	
8B3S0	Post hospital discharge medication reconciliation with pt	126	
8BI..	Other medication review	124	
8CA..	Patient given advice	121	
8H7t.	Referral to pharmacist	121	
J12..	Duodenal ulcer - (DU)	120	
G30z.	Acute myocardial infarction NOS	119	
9k1..	Alcohol misuse - enhanced services administration	119	Alcohol abuse
5C00.	CT scan brain - normal	119	
E204.	Neurotic depression reactive type	119	Mental health problems
R003.	[D]Convulsions	117	
8CM..	Care plan	115	
667Z.	Epilepsy monitoring NOS	115	
9b8D.	Accident & emergency	114	
9180.	Carer's details	112	
G65zz	Transient cerebral ischaemia NOS	111	
1373.	Light smoker - 1-9 cigs/day	110	
567C.	CAT scan brain - abnormal	109	
66c..	Medication monitoring	109	
8BK0.	Depression management programme	107	Mental health problems
TK...	Suicide and selfinflicted injury	107	Mental health problems
E23z.	Alcohol dependence syndrome NOS	106	Alcohol abuse
E2003	Anxiety with depression	106	Mental health problems
G581.	Left ventricular failure	105	
B33..	Other malignant neoplasm of skin	105	
1B13.	Anxiousness	104	Mental health problems
923D.	GP22-practice changed-moved	104	
ZV681	[V]Issue of repeat prescription	103	
66A5.	Diabetic on insulin	103	
8HE2.	Discharged from inpatient care	103	
2128.	Patient's condition the same	102	
137K.	Stopped smoking	102	
663P.	Asthma limiting activities	101	
21260	Epilepsy resolved	100	
8H46.	Neurological referral	100	
667P.	No seizures on treatment	100	
8B3R.	Drug therapy discontinued	99	
8B3A3	New medication commenced	98	
282..	O/E - fit/convulsion	98	
46QB.	Urine methadone	98	Drug abuse
66AJ.	Diabetic - poor control	97	
E25z.	Misuse of drugs NOS	97	Drug abuse
14Q..	H/O: regular medication	96	
8H3Z.	Other hospital admission NOS	96	
9Ol0.	Mental health monitoring first letter	95	Mental health problems

9N21.	Seen by general practitioner	94	
8H62.	Referral to G.P.	93	
1463.	H/O: drug dependency	92	Drug abuse
G61..	Intracerebral haemorrhage	91	
46QA.	Urine cocaine	91	Drug abuse
31130	EEG normal	90	
8B3E.	Drug compliance good	89	
926..	Patient deregistration - health record sent to Health Board	88	
8IB2.	Contraceptiv advice for patients with epilepsy not indicated	87	
2127.	Patient's condition worsened	87	
8BC1.	Treatment plan given	87	
31140	EEG abnormal	84	
9e...	GP out of hours service administration	84	
B22z.	Malignant neoplasm of bronchus or lung NOS	84	
9N0i.	Seen in minor injuries department	83	
137X.	Cigarette consumption	82	
44W2.	Serum phenobarbitone level	82	
66AF.	Attends out-patients	81	
388i.	Case management risk assessment score	81	
G58..	Heart failure	81	
8BIC.	Medication review done by pharmacist	81	
9h91.	Except from mental health quality indicators: Patient unsuit	79	Mental health problems
137d.	Not interested in stopping smoking	79	
E00..	Senile and presenile organic psychotic conditions	78	Mental health problems
9k11.	Alcohol consumption counselling	77	Alcohol abuse
E200z	Anxiety state NOS	77	Mental health problems
663U.	Asthma management plan given	77	
1379.	Ex-moderate smoker (10-19/day)	77	
94...	Death administration	76	
67IM.	Advice to GP to change patient medication	75	
8IAG.	Pandemic influenza vaccination declined	75	
8H49.	Psychiatric referral	74	Mental health problems
C10E.	Type 1 diabetes mellitus	74	
8B35.	Drug Rx stopped-medical advice	73	
13Z4E	Learning difficulties	72	Learning disabilities and congenital abnormalities
9b0d.	Out of hours, practice note	72	
2BBP.	O/E - right eye background diabetic retinopathy	71	
9495.	Patient died in hospital	71	
22K6.	Body mass index less than 20	70	
663N.	Asthma disturbing sleep	69	
8BA8.	Alcohol detoxification	68	Alcohol abuse
E01y0	Alcohol withdrawal syndrome	68	Alcohol abuse
J612.	Alcoholic cirrhosis of liver	68	Alcohol abuse
9hA1.	Excepted from asthma quality indicators: Patient unsuitable	68	
9EV2.	Medical report received from hospital	68	
H333.	Acute exacerbation of asthma	67	
663q.	Asthma daytime symptoms	67	
TJ...	Drugs and other substances-adverse effects in theraputic use	67	
388V.	Mini mental state score	67	
137G.	Trying to give up smoking	67	
2BBQ.	O/E - left eye background diabetic retinopathy	66	
E240z	Opioid drug dependence NOS	66	Drug abuse
667T.	Daily seizures	65	
912C.	Patient registered by HB	65	
B34..	Malignant neoplasm of female breast	62	
923..	Patient de-reg.- GP22 from HB	62	

1B63.	Had a fit	61	
Eu108	[X]Alcohol withdrawal-induced seizure	60	Alcohol abuse
U20..	[X]Intentional self poisoning/exposure to noxious substances	60	Mental health problems
70650	Electroencephalography	60	
137F.	Ex-smoker - amount unknown	60	
9N4Z.	Failed encounter NOS	60	
8B3A2	Medication decreased	60	
J6160	Primary biliary cirrhosis	60	
9N57.	Admin reason for encounter	59	
8B2P.	Drug addiction maintenance therapy - methadone	59	Drug abuse
9NF1.	Home visit request by patient	59	
B34z.	Malignant neoplasm of female breast NOS	59	
3119.	Electroencephalogram	58	
419..	Lab. test result abnormal	58	
41B..	Laboratory test due	58	
8B3y.	Medication review of medical notes	58	
9N3F.	Nurse telephone triage	58	
8IB3.	Pre-conception advic fr patients with epilepsy not indicated	58	
Eu00.	[X]Dementia in Alzheimer's disease	57	
F22..	Hemiplegia	57	
E240.	Opioid type drug dependence	57	Drug abuse
G60..	Subarachnoid haemorrhage	57	
H33z.	Asthma unspecified	56	
9N58.	Emergency appointment	56	
E250.	Nondependent alcohol abuse	56	Alcohol abuse
E2500	Nondependent alcohol abuse, unspecified	56	Alcohol abuse
8IB4.	Pregnancy advice for patients with epilepsy not indicated	56	
9D15.	eMED3 (2010) new statement issued, not fit for work	55	
1465.	H/O: depression	55	Mental health problems
8B41.	Repeated prescription	55	
9N3G.	SMS text message sent to patient	55	
66AV.	Diabetic on insulin and oral treatment	54	
66C..	Obesity monitoring	54	
8H21.	Admit medical emergency unsp.	53	
9491.	Patient died at home	52	
9OJA.	Asthma monitoring check done	51	
5C12.	Computerised tomography brain scan abnormal	51	
8CR7.	Mental health personal health plan	51	Mental health problems
9H8..	On severe mental illness register	51	Mental health problems
8BL2.	Patient on maximal tolerated therapy for diabetes	51	
388G.	Cardiovasc Limits and Symptoms Profile ankle swelling score	50	
388F.	Cardiovascular Limitations and Symptoms Profile angina score	50	
9Ni..	Did not attend	50	
8B2C.	Prescription by nurse practitioner	50	
8B3N.	Prescription not collected	50	
8HHl.	Referral to practice phlebotomist	50	
C1001	Diabetes mellitus, adult onset, no mention of complication	49	
8B39.	Drugs not taken/completed	49	
B46..	Malignant neoplasm of prostate	49	
9k1A.	Brief intervention for excessive alcohol consumptn completed	48	Alcohol abuse
9hC1.	Excepted from depression quality indicators: Informed dissen	48	Mental health problems
8H41.	General medical referral	48	
68R..	New patient screen	48	

9O4..	Patient "recall" admin.	48	
137c.	Thinking about stopping smoking	48	
13Zu.	At risk of emergency hospital admission	47	
8O7..	Carer support	46	
6110.	Contraceptive advice for patients with epilepsy	46	
1365.	Heavy drinker - 7-9u/day	46	Alcohol abuse
9N22.	Seen by practice nurse	46	
H30..	Bronchitis unspecified	45	
9N37.	Message given to patient	45	
1366.	Very heavy drinker - >9u/day	45	Alcohol abuse
R002.	[D]Syncope and collapse	44	
6892.	Alcohol consumption screen	44	
1473.	H/O: epilepsy	44	
8B3U.	Medication review due	44	
22K3.	Body Mass Index low K/M2	43	
F420.	Diabetic retinopathy	43	
8HG..	Died in hospital	43	
E13z.	Nonorganic psychosis NOS	43	Mental health problems
9N2N.	Seen by Rota Doctor	43	
136X.	Alcohol units consumed on heaviest drinking day	42	
9k4..	Depression - enhanced services administration	42	Mental health problems
9Ol1.	Mental health monitoring second letter	42	Mental health problems
H30z.	Bronchitis NOS	40	
J11..	Gastric ulcer - (GU)	40	
9O41.	Initial patient "recall"	40	
8H2Z.	Admit hospital emergency NOS	39	
663s.	Asthma never causes daytime symptoms	39	
9Na1.	Consultation for minor injury	39	
67IJ0	Pre-conception advice for patients with epilepsy	39	
67AF.	Pregnancy advice for patients with epilepsy	39	
8BME.	Prescription sent to pharmacy	39	
9N1R.	Seen in neurology clinic	39	
663n.	Asthma treatment compliance satisfactory	38	
9H92.	Depression interim review	38	Mental health problems
66Ai.	Diabetic 6 month review	38	
14A7.	H/O: CVA/stroke	38	
B51..	Malignant neoplasm of brain	38	
E10..	Schizophrenic disorders	38	Mental health problems
R0020	[D]Blackout	37	
9Oj1.	Hypothyroidism monitoring second letter	37	
9N1A.	Seen in hospital out-pat.	37	
173C.	Short of breath on exertion	37	
13J7.	Unemployed	37	
9N47.	Appointment cancelled by patient	36	
13JJ.	Unfit for work	36	
2828.	Absence seizure	35	
68S..	Alcohol consumption screen	35	
8H47.	Geriatric referral	35	
ZRaA.	Mini-mental state examination	35	
68R2.	New patient screen done	35	
9N34.	Encounter by computer link	34	
137A.	Ex-heavy smoker (20-39/day)	34	
6654.	Follow-up psych. assessment	34	Mental health problems
9e0..	GP out of hours service notified	34	
5694.	Magnetic resonance imaging of brain abnormal	34	
8AZ..	Monitoring of patient NOS	34	
G3071	Acute non-ST segment elevation myocardial infarction	33	
B13..	Malignant neoplasm of colon	33	
8CA9.	Patient Advised to inform DVLA	33	
8H4D.	Referral to psychogeriatrician	33	Mental health problems
R1402	[D]Electroencephalogram (EEG) abnormal	32	

9O4Z.	Patient "recall" admin. NOS	32	
9Nj1.	Patient failed to respond to appointment opt-in letter	32	
Eu102	[X]Mental and behav dis due to use alcohol: dependence syndr	31	Alcohol abuse
9883.	GP/RF - new reg.check to HB	31	
Eu72.	[X]Severe mental retardation	30	Learning disabilities and congenital abnormalities
ZV40.	[V]Mental and behavioural problems	30	Mental health problems
663v.	Asthma causes daytime symptoms most days	30	
137g.	Cigarette pack-years	30	
671E.	Discussed with carer	30	
9h6..	Exception reporting: epilepsy quality indicators	30	
1S3..	General mental state	30	
6774.	Medication counselling	30	
9493.	Patient died in nursing home	30	
8H4..	Referral to physician	30	
E3...	Mental retardation	29	Learning disabilities and congenital abnormalities
H34..	Bronchiectasis	29	
8Hu..	Admission to accident and emergency department	28	
1B6..	Disturbance of consciousness	28	
137M.	Rolls own cigarettes	28	
ZL91.	Seen by accident and emergency doctor	28	
9N2L.	Seen by deputising doctor	28	
E2000	Anxiety state unspecified	27	Mental health problems
E004.	Arteriosclerotic dementia	27	
8BA9.	Detoxification dependence drug	27	Drug abuse
E103.	Paranoid schizophrenia	27	Mental health problems
8H7B.	Refer to community psych.nurse	27	Mental health problems
8CMG.	Review of care plan	27	
R0003	[D]Unconsciousness	26	
44qA.	Drugs of abuse urine screening test	26	
9Of..	Epilepsy screen administration	26	
9E...	Medical examinations/reports	26	
8CAN.	Patient given telephone advice during surgery hours	26	
8H7p.	Referral to community alcohol team	26	Alcohol abuse
9N41.	Did not attend - reason given	25	
1296.	FH: Epilepsy	25	
388W.	Joint British Societies cardiac risk score	25	
T1...	Motor vehicle traffic accidents (MVTA)	25	
E12z.	Paranoid psychosis NOS	25	Mental health problems
66R..	Repeat prescription monitoring	25	
E10z.	Schizophrenia NOS	25	Mental health problems
1372.	Trivial smoker - < 1 cig/day	25	
1376.	Very heavy smoker - 40+cigs/d	25	
Eu300	[X]Hypomania	24	Mental health problems
ZLF2.	Discharge from hospital	24	
8BIq.	Drug compliance checked	24	
67H0.	Lifestyle advice regarding alcohol	24	Alcohol abuse
9P2..	Prescription event monitoring	24	
9N0G.	Seen in primary care centre	24	
13G4.	Social worker involved	24	
9k40.	Depression - enhanced service completed	23	Mental health problems
8B21.	Drug prescription	23	
H32..	Emphysema	23	
6678.	Epilepsy treatment changed	23	
1378.	Ex-light smoker (1-9/day)	23	
13C1.	Fully mobile	23	
8CAn.	Pt given written advice on benefits of physical activity	23	
ZLB5.	Seen by mental health counsellor	23	Mental health problems

Eu815	[X]Severe learning disability	22	Learning disabilities and congenital abnormalities
Eu025	[X]Lewy body dementia	22	
9OL1.	Attends diabetes monitoring	22	
8HE8.	Discharged from accident and emergency	22	
13F8.	Hospital patient	22	
C10FJ	Insulin treated Type 2 diabetes mellitus	22	
B440.	Malignant neoplasm of ovary	22	
9EX1.	Medical report sent	22	
9e02.	Notif to primary care OOHS of palliative care plan in place	22	
92...	Patient de-registration	22	
8B3S1	Post hospital dischrge med reconciliation with medical notes	22	
9N1T.	Seen in psychiatry clinic	22	Mental health problems
E112.	Single major depressive episode	22	Mental health problems
ZV68.	[V]Administrative encounters	21	
ZV6D6	[V]Alcohol abuse counselling and surveillance	21	Alcohol abuse
663t.	Asthma causes daytime symptoms 1 to 2 times per month	21	
9Ov0.	Depression monitoring first letter	21	Mental health problems
6891.	Depression screen	21	
8BMX.	Medication review done by medicines management technician	21	
8CAZ.	Patient given advice NOS	21	
9N7..	Provider initiated encounter	21	
8H71.	Refer to practice nurse	21	
9N1U.	Seen in elderly assessment clinic	21	
1BD1.	Suicidal ideation	21	Mental health problems
8H90.	Telephone contact by consultant	21	
9HB3.	Learning disabilities health assessment	20	Learning disabilities and congenital abnormalities
9k12.	Alcohol misuse - enhanced service completed	20	Alcohol abuse
J613.	Alcoholic liver damage unspecified	20	Alcohol abuse
663u.	Asthma causes daytime symptoms 1 to 2 times per week	20	
B590.	Disseminated malignancy NOS	20	
1B64.	Had a convulsion	20	
9N3..	Indirect encounter	20	
13C..	Mobility - social functioning	20	
9N79.	New patient consultation	20	
9N1v.	Seen in diabetic eye clinic	20	
1B1U.	Symptoms of depression	20	Mental health problems
9b0o.	Telephone consultation	20	
69DB.	Learning disability health examination	19	Learning disabilities and congenital abnormalities
E2F2.	Other specific learning difficulty	19	Learning disabilities and congenital abnormalities
ZV790	[V]Screening for depression	19	
9N4E.	Failed encounter - message left with household member	19	
G667.	Left sided CVA	19	
569F.	Magnetic resonance imaging of brain normal	19	
936..	Patient held record	19	
9Nf..	Patient information status	19	
136J.	Social drinker	19	
Eu04.	[X]Delirium, not induced by alcohol+other psychoactive subs	18	
U0...	[X]Transport accidents	18	
E2B1.	Chronic depression	18	Mental health problems
923B.	GP22-moved out of area	18	
9NJ1.	In-house counselling	18	

569K.	Magnetic resonance imaging of head	18	
667V.	Many seizures a day	18	
13C2.	Mobile outside with aid	18	
8H3..	Non-urgent hospital admission	18	
8H75.	Refer to social worker	18	
9N2G.	Seen by consultant	18	
E1000	Unspecified schizophrenia	18	Mental health problems
44W40	Valproate level therapeutic	18	
9HB2.	Learning disabilities health action plan reviewed	17	Learning disabilities and congenital abnormalities
Eu31.	[X]Bipolar affective disorder	17	Mental health problems
9OF..	Adult immunisation admin.	17	
J61..	Cirrhosis and chronic liver disease	17	
9G2..	Drug addiction notification	17	Drug abuse
212J.	Epilepsy resolved	17	
9DG1.	IB113 DLS Incapacity for work form completed	17	
3891.	Initial patient assessment	17	
9OI2.	Mental health monitoring third letter	17	Mental health problems
13CD.	Mobility very poor	17	
8CAJ.	Patient advised about driving	17	
925..	Patient de-reg. MRE kept	17	
ZL95.	Seen by general practitioner	17	
9N1c.	Seen in general medical clinic	17	
9N12.	Seen in Health Centre	17	
E000.	Uncomplicated senile dementia	17	
9k13.	Alcohol questionnaire completed	16	
663W.	Asthma prophylactic medication used	16	
9OR1.	Attends well adult monitoring	16	
13ZD.	Eye witness to epileptic seizure	16	
9p...	Medication monitoring administration	16	
J615z	Non-alcoholic cirrhosis NOS	16	
212..	Patient examined	16	
91Z..	Patient registration NOS	16	
663S.	Peak flow meter at home	16	
8B2F.	Prescription by GP	16	
665..	Psych. disorder monitoring	16	Mental health problems
9N2a.	Seen by community psychiatric nurse	16	Mental health problems
9N0d.	Seen in open access surgery	16	
1B62.	Syncope/vasovagal faint	16	
ZGB66	Advice to stop drug treatment	15	
8IAF.	Brief intervention for excessive alcohol consumptn declined	15	Alcohol abuse
137J.	Cigar smoker	15	
1BH..	Delusions	15	Mental health problems
F4204	Diabetic maculopathy	15	
6677.	Epilepsy drug side effects	15	
1462.	H/O: alcoholism	15	Alcohol abuse
146B.	H/O: deliberate self harm	15	Mental health problems
C108.	Insulin dependent diabetes mellitus	15	
B141.	Malignant neoplasm of rectum	15	
8BI9.	Medication stopped - side effect	15	
9HA0.	On depression register	15	Mental health problems
9H7..	Removed from severe mental illness register	15	
13O..	Sickness/invalidity benefit	15	
Eu70.	[X]Mild mental retardation	14	Learning disabilities and congenital abnormalities
Eu71.	[X]Moderate mental retardation	14	Learning disabilities and congenital abnormalities
Eu73.	[X]Profound mental retardation	14	Learning disabilities and congenital abnormalities
U10..	[X]Falls	14	
Eu2z.	[X]Unspecified nonorganic psychosis	14	Mental health problems

T8...	Accidental poisoning by drugs, medicines and biologicals	14	
66e..	Alcohol disorder monitoring	14	Alcohol abuse
22K7.	Body mass index 40+ - severely obese	14	
B5832	Cerebral metastasis	14	
9N4o.	Did not attend blood test	14	
671F.	Discussed with patient	14	
14B4.	H/O: asthma	14	
1B1E.	Hallucinations	14	Mental health problems
13G61	Home help attends	14	
8BIw.	Inappropriate medication stopped	14	
13c0.	Injecting drug user	14	Drug abuse
F222.	Left hemiplegia	14	
B060.	Malignant neoplasm of tonsil	14	
9D21.	MED5 - issued to patient	14	
8CY..	Mental Health Care Programme Approach	14	Mental health problems
6A60.	Mental health review follow-up	14	Mental health problems
9OW..	New patient screen admin.	14	
G85..	Other varicose veins	14	
911B.	Patient registered - GPR	14	
137H.	Pipe smoker	14	
13JM.	Problems at work	14	
9O42.	Second patient "recall"	14	
9N0r.	Seen in epilepsy clinic	14	
9No5.	Seen in substance misuse clinic	14	
TK00.	Suicide + selfinflicted poisoning by analgesic/antipyretic	14	Mental health problems
R0030	[D]Convulsions, febrile	13	
8CS7.	Agreeing on mental health care plan	13	Mental health problems
388k.	Christo inventory for substance-misuse services score	13	
E231.	Chronic alcoholism	13	Alcohol abuse
9451.	Death notif. from hospital	13	
9H91.	Depression medication review	13	Mental health problems
8Hg..	Discharge from care	13	
8BIc.	Drug monitoring done	13	
1B6B.	Febrile convulsion	13	
136T.	Harmful alcohol use	13	Alcohol abuse
662W.	Heart failure annual review	13	
T1z..	Motor vehicle traffic accident NOS	13	
6AE..	Multidisciplinary review	13	
E250z	Nondependent alcohol abuse NOS	13	Alcohol abuse
T3...	Other road vehicle accidents	13	
J623.	Portal hypertension	13	
8HC..	Refer to hospital casualty	13	
66R9.	Repeat prescription reviewed by pharmacist	13	
ZL9D4	Seen by psychogeriatrician	13	Mental health problems
700..	Brain tissue operations	12	
9N4V.	DNA - Did not attend epilepsy clinic	12	
8B2R.	Drug addiction detoxification therapy - buprenorphine	12	Drug abuse
8H2P.	Emergency admission, asthma	12	
9Of4.	Epilepsy monitoring telephone invite	12	
9h92.	Except from mental health quality indicators: Informed diss	12	Mental health problems
9k1B.	Extended intervention for excessive alcohol consumptn complt	12	Alcohol abuse
ZRL6.	Geriatric depression scale	12	
8BIJ.	Has shown no side effects from medication	12	
13F3.	Lives alone -no help available	12	
B10z.	Malignant neoplasm of oesophagus NOS	12	
B22..	Malignant neoplasm of trachea, bronchus and lung	12	

8HB8.	Mental therapy follow-up	12	Mental health problems
F372.	Polyneuropathy in diabetes	12	
E001.	Presenile dementia	12	
E1137	Recurrent depression	12	Mental health problems
E1132	Recurrent major depressive episodes, moderate	12	Mental health problems
ZL5AO	Referral to neurologist	12	
ZLA31	Seen by community psychiatric nurse	12	Mental health problems
1B19.	Suicidal	12	Mental health problems
TK05.	Suicide + selfinflicted poisoning by drug or medicine NOS	12	Mental health problems
388z.	Wells deep vein thrombosis clinical probability score	12	
Eu411	[X]Generalized anxiety disorder	11	Mental health problems
Eu192	[X]Mental and behav dis mlti/oth psych sbs: dependence syndr	11	Drug abuse
Eu33.	[X]Recurrent depressive disorder	11	Mental health problems
G30X0	Acute ST segment elevation myocardial infarction	11	
TJ632	Adverse reaction to carbamazepine	11	
8IA7.	Alcohol consumption screening test declined	11	
J153.	Alcoholic gastritis	11	
66E..	B12 deficiency monitoring	11	
E2004	Chronic anxiety	11	Mental health problems
8BIu.	Compliance issues discussed with patient	11	
1B6Z.	Consciousness disturbance NOS	11	
C106.	Diabetes mellitus with neurological manifestation	11	
8B3i.	Drug compliance poor	11	
6679.	Epilepsy treatment started	11	
136D.	Ex-heavy drinker - (7-9u/day)	11	Alcohol abuse
988..	GP/RF-new patient check admin.	11	
1B66.	Had a blackout	11	
13F2.	Lives alone - help available	11	
1BD7.	Low suicide risk	11	
B11..	Malignant neoplasm of stomach	11	
B060z	Malignant neoplasm tonsil NOS	11	
671G0	Medication discussed with pharmacist	11	
9j0..	Medication error	11	
8BI7.	Medication stopped - ineffective	11	
6652.	Mental state	11	
B627.	Non - Hodgkin's lymphoma	11	
G858.	Oesophageal varices NOS	11	
8CMD.	Personal care plan completed	11	
6A8..	Psychological review	11	Mental health problems
8He1.	Referral to intermediate care - community rehabilitation	11	
9Ee1.	Summary report	11	
9NN5.	Under care of psychiatrist	11	Mental health problems
9HB5.	Learning disabilities annual health assessment	10	Learning disabilities and congenital abnormalities
Eu410	[X]Panic disorder [episodic paroxysmal anxiety]	10	Mental health problems
8H35.	Admitted to alcohol detoxification centre	10	Alcohol abuse
67Is.	Advice to continue with drug treatment	10	
1368.	Alcohol consumption unknown	10	
22...	General examination of patient	10	
E2002	Generalised anxiety disorder	10	Mental health problems
B32..	Malignant melanoma of skin	10	
B4302	Malignant neoplasm of endometrium of corpus uteri	10	
B49..	Malignant neoplasm of urinary bladder	10	
8CQ..	Mental health crisis plan	10	Mental health problems
8A3Z.	Neurological monitoring NOS	10	
9187.	New patient question.completed	10	
212A.	O/E patient well	10	
1BH3.	Paranoid ideation	10	Mental health problems

677A.	Psychosexual counselling	10	Mental health problems
E227.	Psychosexual dysfunction	10	
8H4h.	Referral to neurologist	10	
B58..	Secondary malignant neoplasm of other specified sites	10	
9N2c.	Seen by triage nurse	10	
E1122	Single major depressive episode, moderate	10	Mental health problems
8GE..	Social care	10	
TKz..	Suicide and selfinflicted injury NOS	10	Mental health problems
13KB.	Telephone in house	10	
Eu00z	[X]Dementia in Alzheimer's disease, unspecified	9	
Eu321	[X]Moderate depressive episode	9	Mental health problems
9b00.	A&E report	9	
8Hd0.	Admission to community hospital	9	
E11..	Affective psychoses	9	Mental health problems
E010.	Alcohol withdrawal delirium	9	Alcohol abuse
J617.	Alcoholic hepatitis	9	Alcohol abuse
S62..	Cerebral haemorrhage following injury	9	
8BMT.	Current medication as reported by patient	9	
212S.	Depression resolved	9	
667C.	Epilepsy control good	9	
ZRLr.	Hospital anxiety and depression scale	9	
B49z.	Malignant neoplasm of urinary bladder NOS	9	
667B.	Nocturnal epilepsy	9	
8GEZ.	Other social care NOS	9	
9NF5.	Out of hours, not night visit	9	
949..	Patient died - to record place	9	
91...	Patient registration	9	
4Q1F.	Phenytoin level	9	
G668.	Right sided CVA	9	
B577.	Secondary malignant neoplasm of liver	9	
9N1u.	Seen in vascular clinic	9	
137Z.	Tobacco consumption NOS	9	
Ry18.	[D]Self neglect	8	
Eu41z	[X]Anxiety disorder, unspecified	8	Mental health problems
Eu112	[X]Mental and behav dis due to use opioids: dependence syndr	8	Drug abuse
Eu011	[X]Multi-infarct dementia	8	
TGyz3	Accidental injury NOS	8	
E230.	Acute alcoholic intoxication in alcoholism	8	Alcohol abuse
8H23.	Admit psychiatric emergency	8	Mental health problems
67IL.	Advice to GP to stop patient medication	8	
J610.	Alcoholic fatty liver	8	Alcohol abuse
136R.	Binge drinker	8	Alcohol abuse
137T.	Date ceased smoking	8	
66Av.	Diabetic assessment of erectile dysfunction	8	
8B351	Dr stopped drugs - side effect	8	
E24z.	Drug dependence NOS	8	Drug abuse
68U..	Drugs of abuse screening	8	
1461.	H/O: dementia	8	
G58z.	Heart failure NOS	8	
B10..	Malignant neoplasm of oesophagus	8	
B01..	Malignant neoplasm of tongue	8	
8BC2.	Medication given	8	
8B315	Medication recommenced	8	
388B.	Pack years	8	
9O3..	Patient "call" administration	8	
9NA1.	Patient made appointment	8	
9R5..	Patient on waiting list	8	
137b.	Ready to stop smoking	8	
9N6F.	Referred by consultant	8	
8CMG1	Review of mental health care plan	8	Mental health problems

9N1g.	Seen in gastroenterology clinic	8	
44Wg.	Serum lamotrigine level	8	
1JA0.	Suspected epilepsy	8	
SC200	Traumatic epilepsy	8	
9NN2.	Under care of community alcohol team	8	Alcohol abuse
R010.	[D]Abnormal involuntary movements	7	
ZV4KC	[V] Alcohol use	7	Alcohol abuse
Eu107	[X]Men & behav dis due alcoh: resid & late-onset psychot dis	7	Alcohol abuse
B640.	Acute lymphoid leukaemia	7	
8CE1.	Alcohol leaflet given	7	Alcohol abuse
663N0	Asthma causing night waking	7	
9OW1.	Attended new patient screen	7	
E0120	Chronic alcoholic brain syndrome	7	Alcohol abuse
B641.	Chronic lymphoid leukaemia	7	
946..	Death notif.- non.hosp source	7	
8B397	Drug not taken - patient lost tablets	7	
667D.	Epilepsy control poor	7	
137j.	Ex-cigarette smoker	7	
137B.	Ex-very heavy smoker (40+/day)	7	
1282.	FH: Alcoholism	7	
146F.	H/O: drug abuse	7	Drug abuse
14K0.	H/O: repeated overdose	7	Mental health problems
9N3D1	Hospital discharge letter received	7	
6AG..	Housebound patient review	7	
13JV.	In employment	7	
918f.	Is no longer a carer	7	
E0110	Korsakov's alcoholic psychosis	7	Alcohol abuse
8O3..	Long term social support	7	
B4A..	Malig neop of kidney and other unspecified urinary organs	7	
B08..	Malignant neoplasm of hypopharynx	7	
8B3E1	Medication satisfactory	7	
B232.	Mesothelioma of pleura	7	
663V1	Mild asthma	7	
13cB.	Misuses drugs orally	7	Drug abuse
68A1.	Neurological screen	7	
13V6.	No benefits received	7	
F4206	Non proliferative diabetic retinopathy	7	
8BM..	Other medication management	7	
912A.	Patient allocated by HB	7	
8C90.	Patient reassured	7	
665Z.	Psych. disorder monitoring NOS	7	Mental health problems
9HA1.	Removed from depression register	7	
F223.	Right hemiplegia	7	
1J03.	Suspected brain tumour	7	
1JK..	Suspected transient ischaemic attack	7	
9O43.	Third patient "recall"	7	
1B1W.	Transient epileptic amnesia	7	
6664.	Mental handicap problem	6	Learning disabilities and congenital abnormalities
E3z..	Mental retardation NOS	6	Learning disabilities and congenital abnormalities
R21..	[D]Sudden death, cause unknown	6	
Eu320	[X]Mild depressive episode	6	Mental health problems
G307.	Acute subendocardial infarction	6	
Z1Q2.	Administration of medication	6	
8H24.	Admit geriatric emergency	6	
F375.	Alcoholic polyneuropathy	6	
66RB.	Batch prescription cancelled	6	
38G5.	Biopsychosocial assessment	6	
H34z.	Bronchiectasis NOS	6	

G6z..	Cerebrovascular disease NOS	6	
9OEA.	Chronic disease management annual review completed	6	
13o..	Communication skills	6	
56F0.	CT scan brain declined	6	
8H33.	Day hospital care	6	
9N4t.	DNA - Did not attend mental health review	6	
9b0C.	Emergency consultation note	6	
1O30.	Epilepsy confirmed	6	
1284.	FH: Schizophrenia	6	
1471.	H/O: meningitis	6	
14AB.	H/O: TIA	6	
C3500	Haemochromatosis	6	
136S.	Hazardous alcohol use	6	Alcohol abuse
C1089	Insulin dependent diabetes maturity onset	6	
137C.	Keeps trying to stop smoking	6	
B41..	Malignant neoplasm of cervix uteri	6	
8BMH.	Medication review done by pharmacy technician	6	
8BI8.	Medication stopped - contra-indication	6	
8H55.	Neurosurgical referral	6	
68R1.	New patient screen requested	6	
8H38.	Non-urgent psychiatric admisn.	6	Mental health problems
22J4.	O/E - dead - sudden death	6	
9N1t.	Out of hours consultation at surgery	6	
44W31	Phenytoin level therapeutic	6	
SL610	Phenytoin poisoning	6	
E1121	Single major depressive episode, mild	6	Mental health problems
G60z.	Subarachnoid haemorrhage NOS	6	
G65z.	Transient cerebral ischaemia NOS	6	
46QF.	Urine opiates screen	6	Drug abuse
R034E	[D]Developmental delay	≤5	Learning disabilities and congenital abnormalities
Eu7..	[X]Mental retardation	≤5	Learning disabilities and congenital abnormalities
Eu816	[X]Mild learning disability	≤5	Learning disabilities and congenital abnormalities
Eu814	[X]Moderate learning disability	≤5	Learning disabilities and congenital abnormalities
Eu817	[X]Profound learning disability	≤5	Learning disabilities and congenital abnormalities
Eu7z.	[X]Unspecified mental retardation	≤5	Learning disabilities and congenital abnormalities
G669.	Cerebral palsy, not congenital or infantile, acute	≤5	Learning disabilities and congenital abnormalities
E2Fz.	Developmental disorder NOS	≤5	Learning disabilities and congenital abnormalities
9HB4.	Learning disabilities health action plan completed	≤5	Learning disabilities and congenital abnormalities
9HB1.	Learning disabilities health action plan offered	≤5	Learning disabilities and congenital abnormalities
9mA..	Learning disability annual health check invitation	≤5	Learning disabilities and congenital abnormalities
9mA20	Learning disability annual health check invtation 1st letter	≤5	Learning disabilities and congenital abnormalities
E30..	Mild mental retardation, IQ in range 50-70	≤5	Learning disabilities and congenital abnormalities
E310.	Moderate mental retardation, IQ in range 35-49	≤5	Learning disabilities and congenital abnormalities
E31..	Other specified mental retardation	≤5	Learning disabilities and congenital abnormalities
E312.	Profound mental retardation with IQ less than 20	≤5	Learning disabilities and congenital abnormalities

8HHP.	Referral to learning disability team	≤5	Learning disabilities and congenital abnormalities
E311.	Severe mental retardation, IQ in range 20-34	≤5	Learning disabilities and congenital abnormalities
R103.	[D]Alcohol blood level excessive	≤5	Alcohol abuse
R0620	[D]Cough syncope	≤5	
R0033	[D]Reflex anoxic seizure	≤5	
R002z	[D]Syncope and collapse NOS	≤5	
ZV57A	[V]Alcohol rehabilitation	≤5	Alcohol abuse
ZV6..	[V]Other reasons for encounter	≤5	
ZV113	[V]Personal history of alcoholism	≤5	Alcohol abuse
ZV110	[V]Personal history of schizophrenia	≤5	Mental health problems
ZV792	[V]Screening for mental retardation	≤5	Mental health problems
ByuC2	[X]2ndry+unspcf malignant neoplasm lymph nodes/multi regions	≤5	
U1A..	[X]Accidental poisoning by + exposure to noxious substances	≤5	
Gyu6.	[X]Cerebrovascular diseases	≤5	
Eu002	[X]Dementia in Alzheimer's dis, atypical or mixed type	≤5	
Eu000	[X]Dementia in Alzheimer's disease with early onset	≤5	
Eu001	[X]Dementia in Alzheimer's disease with late onset	≤5	
Eu02.	[X]Dementia in other diseases classified elsewhere	≤5	
Eu445	[X]Dissociative convulsions	≤5	
U205.	[X]Intent self poison/exposure to narcotic drug	≤5	Mental health problems
U200.	[X]Intent self poison/exposure to nonopioid analgesic	≤5	Mental health problems
U202.	[X]Intent self poison/exposure to sedative hypnotic	≤5	Mental health problems
Eu326	[X]Major depression, moderately severe	≤5	Mental health problems
ByuA.	[X]Malignant neoplasm of eye, brain and other parts of cent	≤5	
Eu30.	[X]Manic episode	≤5	Mental health problems
Eu100	[X]Mental & behav dis due to use alcohol: acute intoxication	≤5	Alcohol abuse
Eu105	[X]Mental & behav dis due to use alcohol: psychotic disorder	≤5	Alcohol abuse
Eu412	[X]Mixed anxiety and depressive disorder	≤5	Mental health problems
Eu013	[X]Mixed cortical and subcortical vascular dementia	≤5	
Eu41.	[X]Other anxiety disorders	≤5	Mental health problems
Eu200	[X]Paranoid schizophrenia	≤5	Mental health problems
Eu330	[X]Recurrent depressive disorder, current episode mild	≤5	Mental health problems
Eu331	[X]Recurrent depressive disorder, current episode moderate	≤5	Mental health problems
Eu25.	[X]Schizoaffective disorders	≤5	Mental health problems
Eu322	[X]Severe depressive episode without psychotic symptoms	≤5	Mental health problems
Eu01z	[X]Vascular dementia, unspecified	≤5	
2JJ..	Accident and emergency department attendance avoided	≤5	
TCz..	Accidental falls NOS	≤5	
T8z..	Accidental poisoning by drugs NOS	≤5	
A788z	Acquired human immunodeficiency virus infection syndrome NOS	≤5	
A788.	Acquired immune deficiency syndrome	≤5	
J611.	Acute alcoholic hepatitis	≤5	Alcohol abuse
G300.	Acute anterolateral infarction	≤5	
G3011	Acute anteroseptal infarction	≤5	
J120.	Acute duodenal ulcer	≤5	
J120z	Acute duodenal ulcer NOS	≤5	
J1202	Acute duodenal ulcer with perforation	≤5	
J110.	Acute gastric ulcer	≤5	

J1101	Acute gastric ulcer with haemorrhage	≤5	
G302.	Acute inferolateral infarction	≤5	
8Hd3.	Admission by out of hours service doctor	≤5	
9OF1.	Adult imm.- 1st call	≤5	
9OF5.	Adult imm.- 1st recall	≤5	
9OF6.	Adult imm.- 2nd recall	≤5	
8CMe.	Advance care planning	≤5	
TJ90.	Adverse reaction to antidepressants	≤5	
TJ633	Adverse reaction to sodium valproate	≤5	
8CAm.	Advice aim for 30 min of mod act on 5 days or more each week	≤5	
8CAz.	Advice given about breast screening programme	≤5	
ZG231	Advice on alcohol consumption	≤5	Alcohol abuse
ZGB65	Advice to continue with drug treatment	≤5	
67Ie.	Advice to GP to start patient medication	≤5	
ZG122	Advice to return to work	≤5	
E135.	Agitated depression	≤5	Mental health problems
Z923.	Agreeing on care plan	≤5	
66e0.	Alcohol abuse monitoring	≤5	Alcohol abuse
9k19.	Alcohol assesment declined - enhanced services admin	≤5	
SM0..	Alcohol causing toxic effect	≤5	Alcohol abuse
1B1c.	Alcohol induced hallucinations	≤5	Alcohol abuse
136W.	Alcohol misuse	≤5	Alcohol abuse
9k15.	Alcohol screen - AUDIT completed	≤5	
J6710	Alcohol-induced chronic pancreatitis	≤5	Alcohol abuse
13Y8.	Alcoholics anonymous	≤5	Alcohol abuse
8B3I.	Analgesics requested	≤5	
G661.	Anterior cerebral artery syndrome	≤5	
44WF0	Anticonvulsant level therapeutic	≤5	
8B66.	Anticonvulsant therapy	≤5	
8G94.	Anxiety management training	≤5	
E004z	Arteriosclerotic dementia NOS	≤5	
H41..	Asbestosis	≤5	
68910	Assessment using Whooley depression screen	≤5	
663h.	Asthma - currently dormant	≤5	
H33z1	Asthma attack	≤5	
1O2..	Asthma confirmed	≤5	
663N1	Asthma disturbs sleep weekly	≤5	
663w.	Asthma limits walking up hills or stairs	≤5	
663f.	Asthma never restricts exercise	≤5	
66YP.	Asthma night-time symptoms	≤5	
663e.	Asthma restricts exercise	≤5	
66Yp.	Asthma review using Roy Colleg of Physicians three questions	≤5	
663V.	Asthma severity	≤5	
663p.	Asthma treatment compliance unsatisfactory	≤5	
1BD8.	At risk of DSH - deliberate self harm	≤5	Mental health problems
9NM4.	Attending day hospital	≤5	
9OJ1.	Attends asthma monitoring	≤5	
66RA.	Batch prescription issued	≤5	
388g.	Beck depression inventory second edition score	≤5	
136G.	Beer drinker	≤5	
J616.	Biliary cirrhosis	≤5	
8IAH.	Body weight measurement declined	≤5	
9Ow..	Bowel cancer screening programme administration	≤5	
G663.	Brain stem stroke syndrome	≤5	
E243.	Cannabis type drug dependence	≤5	Drug abuse
44W52	Carbamazepine level high	≤5	
44W50	Carbamazepine level therapeutic	≤5	
G613.	Cerebellar haemorrhage	≤5	
G664.	Cerebellar stroke syndrome	≤5	

7A23.	Cerebral artery and circle of Willis aneurysm operations	≤5	
9127.	Chaser sent to HB for MRE	≤5	
J6141	Chronic active hepatitis	≤5	
E231z	Chronic alcoholism NOS	≤5	Alcohol abuse
H3120	Chronic asthmatic bronchitis	≤5	
H31..	Chronic bronchitis	≤5	
H31z.	Chronic bronchitis NOS	≤5	
H320.	Chronic bullous emphysema	≤5	
G5801	Chronic congestive heart failure	≤5	
J614.	Chronic hepatitis	≤5	
B651.	Chronic myeloid leukaemia	≤5	
E04z.	Chronic organic psychosis NOS	≤5	
J615.	Cirrhosis - non alcoholic	≤5	
9OV..	Clinic "D" monitoring admin.	≤5	
661N.	Clinical management plan review	≤5	
9P...	Clinical trial administration	≤5	
76270	Closure of perforated duodenal ulcer	≤5	
13J1.	Commencement of employment	≤5	
8IAg.	Contraceptive advice for patients with epilepsy declined	≤5	
Q480.	Convulsions in newborn	≤5	
941..	Death certificate - form 11	≤5	
13CB.	Depends on public transport	≤5	
9Ov1.	Depression monitoring second letter	≤5	Mental health problems
9Ov2.	Depression monitoring third letter	≤5	Mental health problems
E2112	Depressive personality disorder	≤5	Mental health problems
C106z	Diabetes mellitus NOS with neurological manifestation	≤5	
C100z	Diabetes mellitus NOS with no mention of complication	≤5	
C10zz	Diabetes mellitus NOS with unspecified complication	≤5	
C102.	Diabetes mellitus with hyperosmolar coma	≤5	
C101.	Diabetes mellitus with ketoacidosis	≤5	
C100.	Diabetes mellitus with no mention of complication	≤5	
C104.	Diabetes mellitus with renal manifestation	≤5	
C10z.	Diabetes mellitus with unspecified complication	≤5	
C1000	Diabetes mellitus, juvenile type, no mention of complication	≤5	
F4640	Diabetic cataract	≤5	
F420z	Diabetic retinopathy NOS	≤5	
382..	Diagnostic psychology	≤5	
ZLD81	Discharge by community psychiatric nurse	≤5	Mental health problems
8HE1.	Discharged from outpatients	≤5	
9N44.	DNA - appointment mix-up	≤5	
136I.	Drinks wine	≤5	
8B2Q.	Drug addiction maintenance therapy - buprenorphine	≤5	Drug abuse
8BIr.	Drug changed to cost effective alternative	≤5	
8BI0.	Drug compliance aid requested	≤5	
8B3O.	Drug declined by patient	≤5	
8B3O1	Drug declined by patient - side effects	≤5	
9k50.	Drug misuse - enhanced service completed	≤5	Drug abuse
9k5..	Drug misuse - enhanced services administration	≤5	Drug abuse
1V...	Drug misuse behaviour	≤5	
8B380	Drug not collected - no reason	≤5	
8B392	Drug not taken - side-effects	≤5	
E02..	Drug psychoses	≤5	Drug abuse
E02z.	Drug psychosis NOS	≤5	Drug abuse
8BIY.	Drug treatment no longer needed	≤5	
13c..	Drug user	≤5	Drug abuse

E020.	Drug withdrawal syndrome	≤5	Drug abuse
8B38.	Drugs not issued	≤5	
68U0.	Drugs of abuse urine screening	≤5	Drug abuse
9N3B.	E-mail received from patient	≤5	
3885.	Edinburgh postnatal depression scale	≤5	
31C..	EEG observations	≤5	
7064.	Electroconvulsive therapy	≤5	Mental health problems
7A252	Embolisation of cerebral artery NEC	≤5	
8H230	Emerg psychiatric admiss MHA	≤5	Mental health problems
8B1..	Emergency treatment	≤5	
8B1Z.	Emergency treatment NOS	≤5	
H32z.	Emphysema NOS	≤5	
6674.	Epilepsy associated problems	≤5	
667L.	Epilepsy does not limit activities	≤5	
667K.	Epilepsy limits activities	≤5	
667H.	Epilepsy prevents employment	≤5	
667G.	Epilepsy restricts employment	≤5	
667N.	Epilepsy severity	≤5	
667A.	Epilepsy treatment stopped	≤5	
388n.	Epworth score	≤5	
137N.	Ex pipe smoker	≤5	
136B.	Ex-light drinker - (1-2u/day)	≤5	
136A.	Ex-trivial drinker (<1u/day)	≤5	
1377.	Ex-trivial smoker (<1/day)	≤5	
136E.	Ex-very heavy drinker-(>9u/d)	≤5	Alcohol abuse
70011	Excision of lesion of temporal lobe of brain	≤5	
H330.	Extrinsic (atopic) asthma	≤5	
H330z	Extrinsic asthma NOS	≤5	
42Qe.	Factor V Leiden genotypy	≤5	
137m.	Failed attempt to stop smoking	≤5	
9N4D.	Failed encounter - message left with work place	≤5	
9N4G.	Failed encounter - phone number unobtainable	≤5	
16D6.	Fall	≤5	
136b.	Feels should cut down drinking	≤5	Alcohol abuse
1285.	FH: Depression	≤5	
128..	FH: Mental disorder	≤5	
13JW.	Fit for work	≤5	
9EE1.	Fitness to drive rep sent off	≤5	
B627C	Follicular non-Hodgkin's lymphoma	≤5	
8HBA.	Follow up in outpatient clinic	≤5	
13JA.	Full-time employment	≤5	
J11z.	Gastric ulcer NOS	≤5	
38QN.	Generalised anxiety disorder 2 scale	≤5	
388w.	Generalised anxiety disorder 7 item score	≤5	
388K.	Geriatric depression scale	≤5	
9b1F.	GP practice	≤5	
923K.	GP22 - practice advises outside of their area	≤5	
923E.	GP22-moved-c/a in area	≤5	
9232.	GP22-removed at Dr's request	≤5	
9238.	GP22-removed-patients request	≤5	
9511.	GP24 signed by patient	≤5	
9829.	GP81 - higher rate-sent to HB	≤5	
1T8..	H/O cannabis misuse	≤5	Drug abuse
1T0..	H/O heroin misuse	≤5	Drug abuse
14P4.	H/O: anticonvulsant therapy	≤5	
1466.	H/O: anxiety state	≤5	Mental health problems
146A.	H/O: attempted suicide	≤5	Mental health problems
147D.	H/O: blackout	≤5	
1434.	H/O: diabetes mellitus	≤5	
1479.	H/O: disturbance of consciousness	≤5	
146..	H/O: psychiatric disorder	≤5	Mental health problems
146Z.	H/O: psychiatric disorder NOS	≤5	Mental health problems

146H.	H/O: psychosis	≤5	Mental health problems
146E.	H/O: recreational drug use	≤5	Drug abuse
1464.	H/O: schizophrenia	≤5	Mental health problems
1BD..	Harmful thoughts	≤5	Mental health problems
8CME.	Has end of life advanced care plan	≤5	
8BIE.	Has shown side effects from medication	≤5	
679A.	Health ed.- drugs of addiction	≤5	
8CL3.	Heart failure care plan discussed with patient	≤5	
1O1..	Heart failure confirmed	≤5	
136P.	Heavy drinker	≤5	Alcohol abuse
F22z.	Hemiplegia NOS	≤5	
J622.	Hepatic coma	≤5	
J633.	Hepatitis unspecified	≤5	
B1503	Hepatocellular carcinoma	≤5	
1V65.	Heroin misuse	≤5	Drug abuse
1BD5.	High suicide risk	≤5	Mental health problems
B61..	Hodgkin's disease	≤5	
B61z.	Hodgkin's disease NOS	≤5	
9NF..	Home visit admin	≤5	
388J.	Hospital anxiety and depression scale	≤5	
A789.	Human immunodef virus resulting in other disease	≤5	
C109K	Hyperosmolar non-ketotic state in type 2 diabetes mellitus	≤5	
E2015	Hysterical seizures	≤5	
9ND8.	I/patient disch summary rec'vd	≤5	
13CP.	Impaired mobility	≤5	
9NJ8.	In-house chiropody - domiciliary visit	≤5	
9N3Z.	Indirect encounter NOS	≤5	
9O31.	Initial patient "call"	≤5	
6653.	Initial psych. assessment	≤5	Mental health problems
C109J	Insulin treated Type 2 diabetes mellitus	≤5	
14K1.	Intentional overdose of prescription only medication	≤5	
G61z.	Intracerebral haemorrhage NOS	≤5	
G617.	Intracerebral haemorrhage, intraventricular	≤5	
G62z.	Intracranial haemorrhage NOS	≤5	
13c1.	Intravenous drug user	≤5	Drug abuse
H331.	Intrinsic asthma	≤5	
H331z	Intrinsic asthma NOS	≤5	
H3310	Intrinsic asthma without status asthmaticus	≤5	
5742.	Isotope brain angiography	≤5	
5731.	Isotope static scan brain	≤5	
662m.	JBS cardiovascular disease risk >20% up to 30% ov next 10 yr	≤5	
B33z0	Kaposi's sarcoma of skin	≤5	
B6303	Lambda light chain myeloma	≤5	
H33z2	Late-onset asthma	≤5	
67H3.	Lifestyle advice regarding drug misuse	≤5	Drug abuse
136N.	Light drinker	≤5	
ZU31.	Lives alone	≤5	
13F81	Long stay hospital inpatient	≤5	
13FS.	Long stay hospital inpatient	≤5	
8B3X.	Lost prescription	≤5	
13U..	Low income benefits	≤5	
569K0	Magnetic resonance imaging of brain	≤5	
B5762	Malignant ascites	≤5	
B62y.	Malignant lymphoma NOS	≤5	
B3275	Malignant melanoma of ankle	≤5	
B327.	Malignant melanoma of lower limb and hip	≤5	
B32z.	Malignant melanoma of skin NOS	≤5	
B162.	Malignant neoplasm of ampulla of Vater	≤5	
B142.	Malignant neoplasm of anal canal	≤5	
B136.	Malignant neoplasm of ascending colon	≤5	

B495.	Malignant neoplasm of bladder neck	≤5
B43z.	Malignant neoplasm of body of uterus NOS	≤5
B51z.	Malignant neoplasm of brain NOS	≤5
B134.	Malignant neoplasm of caecum	≤5
B3310	Malignant neoplasm of canthus	≤5
B516.	Malignant neoplasm of cerebellum	≤5
B41z.	Malignant neoplasm of cervix uteri NOS	≤5
B13z.	Malignant neoplasm of colon NOS	≤5
B3114	Malignant neoplasm of connective and soft tissue of finger	≤5
B5421	Malignant neoplasm of craniopharyngeal duct	≤5
B132.	Malignant neoplasm of descending colon	≤5
B04..	Malignant neoplasm of floor of mouth	≤5
B04z.	Malignant neoplasm of floor of mouth NOS	≤5
B511.	Malignant neoplasm of frontal lobe	≤5
B160.	Malignant neoplasm of gallbladder	≤5
B052.	Malignant neoplasm of hard palate	≤5
B170.	Malignant neoplasm of head of pancreas	≤5
B130.	Malignant neoplasm of hepatic flexure of colon	≤5
B4Az.	Malignant neoplasm of kidney or urinary organs NOS	≤5
B4A0.	Malignant neoplasm of kidney parenchyma	≤5
B21..	Malignant neoplasm of larynx	≤5
B21z.	Malignant neoplasm of larynx NOS	≤5
B0zz.	Malignant neoplasm of lip, oral cavity and pharynx NOS	≤5
B152.	Malignant neoplasm of liver unspecified	≤5
B2241	Malignant neoplasm of lower lobe of lung	≤5
B221.	Malignant neoplasm of main bronchus	≤5
B301.	Malignant neoplasm of mandible	≤5
B104.	Malignant neoplasm of middle third of oesophagus	≤5
B514.	Malignant neoplasm of occipital lobe	≤5
B06..	Malignant neoplasm of oropharynx	≤5
B5...	Malignant neoplasm of other and unspecified sites	≤5
B17..	Malignant neoplasm of pancreas	≤5
B17z.	Malignant neoplasm of pancreas NOS	≤5
B483.	Malignant neoplasm of penis, part unspecified	≤5
B543.	Malignant neoplasm of pineal gland	≤5
B111.	Malignant neoplasm of pylorus of stomach	≤5
B140.	Malignant neoplasm of rectosigmoid junction	≤5
B2zz.	Malignant neoplasm of respiratory tract NOS	≤5
B133.	Malignant neoplasm of sigmoid colon	≤5
B33z.	Malignant neoplasm of skin NOS	≤5
B3333	Malignant neoplasm of skin of forehead	≤5
B3334	Malignant neoplasm of skin of nose (external)	≤5
B11z.	Malignant neoplasm of stomach NOS	≤5
B172.	Malignant neoplasm of tail of pancreas	≤5
B47z.	Malignant neoplasm of testis NOS	≤5
B53..	Malignant neoplasm of thyroid gland	≤5
B3311	Malignant neoplasm of upper eyelid	≤5
B2221	Malignant neoplasm of upper lobe of lung	≤5
B222.	Malignant neoplasm of upper lobe, bronchus or lung	≤5
B4A2.	Malignant neoplasm of ureter	≤5
B496.	Malignant neoplasm of ureteric orifice	≤5
B40..	Malignant neoplasm of uterus, part unspecified	≤5
B054.	Malignant neoplasm of uvula	≤5
B333.	Malignant neoplasm skin of other and unspecified parts face	≤5
8B3o.	Medication change to generic	≤5
8BT..	Medication review - additional	≤5

8BMY.	Medication review done by medicines management pharmacist	≤5	
8BI6.	Medication stopped - interaction	≤5	
E....	Mental disorders	≤5	Mental health problems
69F..	Mental Health Act examination	≤5	Mental health problems
ZQ3E.	Mental health assessment	≤5	Mental health problems
9OI..	Mental health monitoring administration	≤5	Mental health problems
B226.	Mesothelioma	≤5	
137h.	Minutes from waking to first tobacco consumption	≤5	
1BD6.	Moderate suicide risk	≤5	Mental health problems
B630.	Multiple myeloma	≤5	
B6y0.	Myeloproliferative disorder	≤5	
8BML.	Needs domiciliary care worker to administer medication	≤5	
B....	Neoplasms	≤5	
Bz...	Neoplasms NOS	≤5	
68A..	Neurolog./special sense screen	≤5	
1O3..	Neurological disorder confirmed	≤5	
7P10.	Neuropsychology tests	≤5	
13c2.	Never injecting drug user	≤5	Drug abuse
6613.	Next hospital appointment	≤5	
6AH..	NHS Health Check annual review	≤5	
9b0W.	Night visit, deputising service note	≤5	
8B37.	No drug therapy prescribed	≤5	
663T.	No peak flow meter at home	≤5	
E040.	Non-alcoholic amnestic syndrome	≤5	
J61y1	Non-alcoholic fatty liver	≤5	
C1097	Non-insulin dependent diabetes mellitus - poor control	≤5	
E1z..	Non-organic psychosis NOS	≤5	Mental health problems
E2502	Nondependent alcohol abuse, episodic	≤5	Alcohol abuse
E257.	Nondependent amphetamine or other psychostimulant abuse	≤5	Drug abuse
137U.	Not a passive smoker	≤5	
2258.	O/E - anxious	≤5	Mental health problems
2577.	O/E - breath - alcohol smell	≤5	Alcohol abuse
2257.	O/E - depressed	≤5	Mental health problems
2BBL.	O/E - diabetic maculopathy present both eyes	≤5	
2231.	O/E - fully conscious	≤5	
2822.	O/E - grand mal fit	≤5	
2BBS.	O/E - left eye preproliferative diabetic retinopathy	≤5	
2B3..	O/E - meningism	≤5	
2BBR.	O/E - right eye preproliferative diabetic retinopathy	≤5	
2236.	O/E - unconscious/comatose	≤5	
G8523	Oesophageal varices in alcoholic cirrhosis of the liver	≤5	
G852z	Oesophageal varices in diseases EC NOS	≤5	
G850.	Oesophageal varices with bleeding	≤5	
G851.	Oesophageal varices without bleeding	≤5	
663F.	Oral steroids started	≤5	
E0...	Organic psychotic conditions	≤5	Mental health problems
G30yz	Other acute myocardial infarction NOS	≤5	
E012.	Other alcoholic dementia	≤5	Alcohol abuse
B62..	Other malignant neoplasm of lymphoid and histiocytic tissue	≤5	
G301.	Other specified anterior myocardial infarction	≤5	
G65y.	Other transient cerebral ischaemia	≤5	
6612.	Out-patients last attended	≤5	
9N3D2	Outpatient clinic letter received	≤5	
8B3Y.	Over the counter nicotine replacement therapy	≤5	
F241.	Paraplegia	≤5	
70002	Partial lobectomy of brain	≤5	

137I.	Passive smoker	≤5	
9R52.	Patient awaiting OPD appt.	≤5	
9N85.	Patient comes yearly	≤5	
9Nd8.	Patient consented to randomised clinical trial	≤5	
6721.	Patient counselled	≤5	
66Af.	Patient diabetes education review	≤5	
949Z.	Patient died in place NOS	≤5	
9131.	Patient has moved away	≤5	
8BL..	Patient on maximum tolerated dose	≤5	
912B.	Patient re-registered	≤5	
915..	Patient registration data	≤5	
9N5A.	Patient review of medical records	≤5	
8I4..	Patient self discharge	≤5	
9N43.	Patient too late	≤5	
9RT..	Patient transfer	≤5	
9N45.	Patient walked out	≤5	
1956.	Peptic ulcer symptoms	≤5	
7A238	Percutaneous coil embolisation of cerebral artery aneurysm	≤5	
66RD.	Pharmacy managed repeat prescription	≤5	
44W32	Phenytoin level high	≤5	
44W30	Phenytoin level low	≤5	
137a.	Pipe tobacco consumption	≤5	
ZQ3F.	Planned mental health assessment	≤5	Mental health problems
44Ws.	Plasma carbamazepine level	≤5	
44AN.	Plasma pro-brain natriuretic peptide level	≤5	
H410.	Pleural plaque disease due to asbestosis	≤5	
SL6x1	Poisoning by carbamazepine	≤5	
8B3e.	Poor compliance with inhaler	≤5	
H341.	Post-infective bronchiectasis	≤5	
8IAh.	Pre-conception advice for patients with epilepsy declined	≤5	
8Ce1.	Preferred place of care - hospice	≤5	
8IAi.	Pregnancy advice for patients with epilepsy declined	≤5	
F4202	Preproliferative diabetic retinopathy	≤5	
9bG5.	Prescription by another organisation	≤5	
8B2E.	Prescription by nurse	≤5	
8B2J.	Prescription by supplementary prescriber	≤5	
E001z	Presenile dementia NOS	≤5	
13F82	Previous multiple hospital admissions	≤5	
13cJ.	Previously injecting drug user	≤5	Drug abuse
F1321	Progressive myoclonic epilepsy	≤5	
F4201	Proliferative diabetic retinopathy	≤5	
9N7Z.	Provider initiated encount NOS	≤5	
6659.	Psych.treatment started	≤5	Mental health problems
665A.	Psych.treatment stopped	≤5	Mental health problems
8A2..	Psychiatric monitoring	≤5	Mental health problems
384..	Psychological analysis	≤5	
6779.	Psychological counselling	≤5	Mental health problems
8GA..	Psychological nursing	≤5	Mental health problems
212T.	Psychosis, schizophrenia + bipolar affective disord resolved	≤5	Mental health problems
G666.	Pure sensory lacunar syndrome	≤5	
E130.	Reactive depressive psychosis	≤5	Mental health problems
E2005	Recurrent anxiety	≤5	Mental health problems
E113.	Recurrent major depressive episode	≤5	Mental health problems
E1131	Recurrent major depressive episodes, mild	≤5	Mental health problems
E1133	Recurrent major depressive episodes, severe, no psychosis	≤5	Mental health problems
8B25.	Reducing dose prescribed	≤5	
8HR5.	Refer for EEG	≤5	

66aF.	Refer for falls assessment	≤5	
8H7b.	Refer to day hospital	≤5	
8H7a.	Refer to hospital	≤5	
8HD..	Refer to hospital OPD	≤5	
8H7A.	Refer to mental health worker	≤5	Mental health problems
8H7T.	Refer to psychologist	≤5	Mental health problems
8HHq.	Referral for guided self-help for depression	≤5	Mental health problems
8HZ0.	Referral needed	≤5	
ZL51.	Referral to accident and emergency doctor	≤5	
8HHR.	Referral to child and adolescent psychiatry service	≤5	
8H4P.	Referral to child psychiatrist	≤5	
8HHe.	Referral to community drug and alcohol team	≤5	Drug abuse
8Hc0.	Referral to community mental health team	≤5	Mental health problems
8H4F.	Referral to diabetologist	≤5	
8Hj..	Referral to education service	≤5	
8Hlp.	Referral to epilepsy clinic	≤5	
8He..	Referral to intermediate care	≤5	
8Hc..	Referral to mental health team	≤5	Mental health problems
8H4Y.	Referral to neurology special interest general practitioner	≤5	
ZL5GL	Referral to neurosurgeon	≤5	
8HHo.	Referral to older age community mental health team	≤5	Mental health problems
8Hc2.	Referral to primary care mental health team	≤5	Mental health problems
ZL5B.	Referral to psychiatrist	≤5	Mental health problems
ZL78.	Referral to psychologist	≤5	Mental health problems
ZL79.	Referral to social worker	≤5	
8HkG.	Referral to specialist alcohol treatment service	≤5	Alcohol abuse
137k.	Refusal to give smoking status	≤5	
9OL2.	Refuses diabetes monitoring	≤5	
9134.	Registration ghost - deceased	≤5	
13L3.	Relative - psychotic illness	≤5	
70121	Removal of cerebroventricular shunt	≤5	
9N73.	Repeat medication check	≤5	
66R8.	Repeat Rx reviewed by hospital	≤5	
9Ee09	Report of clinical encounter	≤5	
13JP.	Returned to work	≤5	
8CMG5	Review of anticipatory care plan	≤5	
T3z..	Road vehicle accident NOS	≤5	
G600.	Ruptured berry aneurysm	≤5	
B5630	Secondary and unspec malig neop axillary lymph nodes	≤5	
B56..	Secondary and unspecified malignant neoplasm of lymph nodes	≤5	
B57..	Secondary malig neop of respiratory and digestive systems	≤5	
B585.	Secondary malignant neoplasm of bone and bone marrow	≤5	
B583.	Secondary malignant neoplasm of brain and spinal cord	≤5	
B58z.	Secondary malignant neoplasm of other specified site NOS	≤5	
9N2C.	Seen by dentist	≤5	
ZL9..	Seen by doctor	≤5	
ZL956	Seen by GP locum	≤5	
9N2n.	Seen by herbalist	≤5	
9N2D.	Seen by locum doctor	≤5	
ZLBA.	Seen by pharmacist	≤5	
ZL9D.	Seen by psychiatrist	≤5	Mental health problems
9N2W.	Seen by psychologist	≤5	Mental health problems
9N2g.	Seen by respiratory physician	≤5	
9N04.	Seen in emergency clinic	≤5	
9N1Y.	Seen in gastroscopy clinic	≤5	

9N1J.	Seen in gynaecology clinic	≤5	
9N0I.	Seen in lipid clinic	≤5	
9N1m.	Seen in nephrology clinic	≤5	
9N1j.	Seen in neurosurgery clinic	≤5	
9N1V.	Seen in paediatric clinic	≤5	
9N1a.	Seen in plastic surgery clinic	≤5	
9N1M.	Seen in psychology clinic	≤5	Mental health problems
1B27.	Seizures in response to acute event	≤5	
TK601	Self inflicted lacerations to wrist	≤5	Mental health problems
ZX...	Self-harm	≤5	Mental health problems
8HJJ.	Self-referral to accident and emergency department	≤5	
8HJ..	Self-referral to hospital	≤5	
E00z.	Senile or presenile psychoses NOS	≤5	Mental health problems
663V3	Severe asthma	≤5	
1B1C.	Sexual symptom	≤5	
E1123	Single major depressive episode, severe, without psychosis	≤5	Mental health problems
9N1Z.	Site of encounter NOS	≤5	
H3101	Smokers' cough	≤5	
137e.	Smoking restarted	≤5	
137Q.	Smoking started	≤5	
13HL.	Social isolation	≤5	
13Z8.	Social problem	≤5	
B6301	Solitary myeloma	≤5	
F221.	Spastic hemiplegia	≤5	
136F.	Spirit drinker	≤5	
H33z0	Status asthmaticus NOS	≤5	
70086	Stereotactic radiosurgery on tissue of brain	≤5	
G602.	Subarachnoid haemorrhage from middle cerebral artery	≤5	
SK07.	Subcutaneous emphysema	≤5	
G623.	Subdural haemorrhage NOS	≤5	
13Hl.	Subject of multi-agency public protection arrangements	≤5	
13cM.	Substance misuse	≤5	Drug abuse
1BD3.	Suicidal plans	≤5	Mental health problems
TK01.	Suicide + selfinflicted poisoning by barbiturates	≤5	Mental health problems
TK02.	Suicide + selfinflicted poisoning by oth sedatives/hypnotics	≤5	Mental health problems
TK04.	Suicide + selfinflicted poisoning by other drugs/medicines	≤5	Mental health problems
TK0..	Suicide + selfinflicted poisoning by solid/liquid substances	≤5	Mental health problems
TK03.	Suicide + selfinflicted poisoning tranquilliser/psychotropic	≤5	Mental health problems
1BD4.	Suicide risk	≤5	Mental health problems
1369.	Suspect alcohol abuse - denied	≤5	Alcohol abuse
1JA10	Suspected cerebrovascular accident	≤5	
1JA1.	Suspected cerebrovascular disease	≤5	
1JJ..	Suspected depression	≤5	Mental health problems
1J1..	Suspected drug abuse	≤5	Drug abuse
8BM2.	Synchronisation of repeat medication	≤5	
9b0n.	Telephone call to a patient	≤5	
8B2..	Therapeutic prescription	≤5	
8G91.	Therapeutic psychology	≤5	Mental health problems
9b0q.	Third party consultation	≤5	
137E.	Tobacco consumption unknown	≤5	
137n.	Total time smoked	≤5	
1BZ2.	Transient neurological symptoms	≤5	
S627.	Traumatic subarachnoid haemorrhage	≤5	
S628.	Traumatic subdural haemorrhage	≤5	
1B26.	Trigger factor for seizure	≤5	

C10EE	Type 1 diabetes mellitus with hypoglycaemic coma	≤5	
C10EM	Type 1 diabetes mellitus with ketoacidosis	≤5	
C10FM	Type 2 diabetes mellitus with persistent microalbuminuria	≤5	
C10FL	Type 2 diabetes mellitus with persistent proteinuria	≤5	
C10F4	Type 2 diabetes mellitus with ulcer	≤5	
C10F9	Type 2 diabetes mellitus without complication	≤5	
E0010	Uncomplicated presenile dementia	≤5	
9NNM.	Under care of community psychiatric nurse	≤5	Mental health problems
13R..	Unemployment benefits	≤5	
E117.	Unspecified bipolar affective disorder	≤5	Mental health problems
9Nj0.	Unsuccessful attempt to contact patient by telephone	≤5	
46Qf.	Urine methadone metabolite level	≤5	
44W42	Valproate level high	≤5	
44W41	Valproate level low	≤5	
136Q.	Very heavy drinker	≤5	Alcohol abuse
70656	Video EEG	≤5	
8IAf.	Waist measurement declined	≤5	
Total		441101	

Appendix S4 – Regression analyses

S4a - Annual trend in epilepsy-related death mortality rate (MR) overall

```
Regression analysis

lm(formula = Rate ~ Year, data = mydata2)

Residuals:
      1       2       3       4       5       6       7
 0.04659  0.67079 -0.03030 -0.64001 -0.45295 -0.68639  1.09226

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) -607.9299   281.0622  -2.163   0.0829 .
Year           0.3055     0.1397   2.187   0.0804 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.7392 on 5 degrees of freedom
Multiple R-squared:  0.4889,    Adjusted R-squared:  0.3867
F-statistic: 4.783 on 1 and 5 DF,  p-value: 0.08041

> confint.lm(lm(Rate~Year,data=mydata2))
                  2.5 %    97.5 %
(Intercept) -1.330423e+03 114.56339
Year        -5.359385e-02   0.66459
```

S4b - Annual trend in epilepsy-related death mortality rate (MR) in men

```
Call:
lm(formula = Rate ~ Year, data = mydata3)

Residuals:
     1      2      3      4      5      6      7
-0.1786  0.7143  0.4071 -0.6000 -0.5071 -1.3143  1.4786

Coefficients:
                          Estimate Std. Error t value Pr(>|t|)
(Intercept)               -610.2714   389.1980  -1.568    0.178
Change in annual MR          0.3071     0.1934   1.588    0.173

Residual standard error: 1.024 on 5 degrees of freedom
Multiple R-squared:  0.3352,  Adjusted R-squared:  0.2022
F-statistic: 2.521 on 1 and 5 DF,  p-value: 0.1732

> confint.lm(lm(Rate~Year,data=mydata3))
                                    2.5 %     97.5 %
(Intercept)                 -1610.7367568  390.1938996
95% CI for change in annual MR -0.1901061  0.8043918
```

S4c - Annual trend in epilepsy-related death mortality rate (MR) in women

Call:
lm(formula = Rate ~ Year, data = mydata4)

Residuals:
```
     1      2      3      4      5      6      7
0.1893 0.6214 -0.3464 -0.6143 -0.4821 -0.0500 0.6821
```

Coefficients:
```
                    Estimate Std. Error t value Pr(>|t|)
(Intercept)          -533.1143   216.4082  -2.463   0.0570 .
Change in annual MR     0.2679     0.1076   2.490   0.0551 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Residual standard error: 0.5691 on 5 degrees of freedom
Multiple R-squared: 0.5536, Adjusted R-squared: 0.4644
F-statistic: 6.202 on 1 and 5 DF, p-value: 0.05514

```
> confint.lm(lm(Rate~Year,data=mydata4))
                                  2.5 %     97.5 %
                (Intercept) -1.089409e+03 23.1808181
95% CI for change in annual MR -8.631341e-03  0.5443456
```

S4d - Trend in age-specific mortality ratios overall

lm(formula = SMR ~ Year, data = subset(mydata, **Age_Group == "16-24 years"**))

Residuals:
```
     1      2      3      4      5      6      7
-3.5105 3.4319 -0.6763 3.5064 -0.9236 -2.0637 0.2357
```

Coefficients:
```
                                Estimate Std. Error t value Pr(>|t|)
(Intercept)                     -333.4808  1102.2877  -0.303    0.774
Δannual age-specific mortality ratio  0.1686     0.5479   0.308    0.771
```

Residual standard error: 2.899 on 5 degrees of freedom
Multiple R-squared: 0.01859, Adjusted R-squared: -0.1777
F-statistic: 0.09473 on 1 and 5 DF, p-value: 0.7707

```
> confint.lm(lm(SMR~Year,data=subset(mydata, Age_Group=='16-24 years')))
                                          2.5 %      97.5 %
(Intercept)                          -3167.001632 2500.040064
95% CI for Δannual age-specific mortality ratio  -1.239694    1.576926
```

lm(formula = SMR ~ Year, data = subset(mydata, **Age_Group == "25-34 years"**))

Residuals:
```
      1       2       3       4       5       6       7
-0.49818  0.08109  0.08611  1.16981 -0.67120  0.07220 -0.23982
```

Coefficients:
```
                                Estimate Std. Error t value Pr(>|t|)
(Intercept)                     175.24289  248.97980   0.704    0.513
```
∆annual age-specific mortality ratio -0.08527 0.12375 -0.689 0.521

Residual standard error: 0.6548 on 5 degrees of freedom
Multiple R-squared: 0.08672, Adjusted R-squared: -0.09593
F-statistic: 0.4748 on 1 and 5 DF, p-value: 0.5215

```
> confint.lm(lm(SMR~Year,data=subset(mydata, Age_Group=='25-34 years')))
                                        2.5 %      97.5 %
(Intercept)                         -464.7800629 815.2658331
```
95% CI for ∆annual age-specific mortality ratio -0.4033703 0.2328351

>

lm(formula = SMR ~ Year, data = subset(mydata, **Age_Group == "35-44 years"))**

Residuals:
```
      1       2       3       4       5       6       7
 0.17665  0.05169  0.17088 -0.38167 -0.41090 -0.03504  0.42840
```

Coefficients:
```
                                Estimate Std. Error t value Pr(>|t|)
(Intercept)                     -120.80347  127.51566  -0.947    0.387
```
∆annual age-specific mortality ratio 0.06112 0.06338 0.964 0.379

Residual standard error: 0.3354 on 5 degrees of freedom
Multiple R-squared: 0.1568, Adjusted R-squared: -0.01181
F-statistic: 0.93 on 1 and 5 DF, p-value: 0.3792

```
> confint.lm(lm(SMR~Year,data=subset(mydata, Age_Group=='35-44 years')))
                                        2.5 %      97.5 %
(Intercept)                         -448.5929193 206.9859723
```
95% CI for ∆annual age-specific mortality ratio -0.1017991 0.2240352

lm(formula = SMR ~ Year, data = subset(mydata, **Age_Group == "45-54 years"**))

Residuals:
```
      1       2       3       4       5       6       7
-0.07108 -0.20022  0.39534 -0.02252  0.05507 -0.19637  0.03978
```

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-0.871199	84.188408	-0.010	0.992
∆annual age-specific mortality ratio	**0.001219**	0.041843	0.029	0.978

Residual standard error: 0.2214 on 5 degrees of freedom
Multiple R-squared: 0.0001698, Adjusted R-squared: -0.1998
F-statistic: 0.0008493 on 1 and 5 DF, p-value: 0.9779

> confint.lm(lm(SMR~Year,data=subset(mydata, Age_Group=='45-54 years')))

	2.5 %	97.5 %
(Intercept)	-217.2843910	215.5419938
95% CI for ∆annual age-specific mortality ratio	**-0.1063417**	**0.1087806**

lm(formula = SMR ~ Year, data = subset(mydata, **Age_Group == "55-64 years"**))

Residuals:

1	2	3	4	5	6	7
0.07007	-0.01511	0.14619	-0.23147	-0.22964	0.22403	0.03593

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-187.90436	73.00056	-2.574	0.0498 *
∆annual age-specific mortality ratio	**0.09391**	0.03628	2.588	0.0489 *

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.192 on 5 degrees of freedom
Multiple R-squared: 0.5726, Adjusted R-squared: 0.4872
F-statistic: 6.7 on 1 and 5 DF, p-value: 0.04893

> confint.lm(lm(SMR~Year,data=subset(mydata, Age_Group=='55-64 years')))

	2.5 %	97.5 %
(Intercept)	-3.755583e+02	-0.2504523
95% CI for ∆annual age-specific mortality ratio	**6.466026e-04**	**0.1871812**

lm(formula = SMR ~ Year, data = subset(mydata, **Age_Group == "65-74 years"**))

Residuals:

1	2	3	4	5	6	7
0.058396	-0.001005	-0.082181	0.083712	-0.047471	-0.172821	0.161370

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-102.49588	46.67214	-2.196	0.0795 .

∆annual age-specific mortality ratio 0.05128 0.02320 2.211 0.0780 .

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.1227 on 5 degrees of freedom
Multiple R-squared: 0.4943, Adjusted R-squared: 0.3931
F-statistic: 4.887 on 1 and 5 DF, p-value: 0.07804

> confint.lm(lm(SMR~Year,data=subset(mydata, Age_Group=='65-74 years')))
                                        2.5 %   97.5 %
(Intercept)   -2.224704e+02 17.4786651
95% CI for ∆annual age-specific mortality ratio -8.348717e-03 0.1109102

lm(formula = SMR ~ Year, data = subset(mydata, Age_Group == "75+ years"))

Residuals:
      1   2   3   4   5   6   7
-0.082683 0.095025 0.038945 0.008933 -0.050567 -0.060475 0.050821

Coefficients:
                                Estimate Std. Error t value Pr(>|t|)
(Intercept)   8.152945 27.555558   0.296   0.779
∆annual age-specific mortality ratio -0.003842 0.013696 -0.281   0.790

Residual standard error: 0.07247 on 5 degrees of freedom
Multiple R-squared: 0.0155, Adjusted R-squared: -0.1814
F-statistic: 0.07872 on 1 and 5 DF, p-value: 0.7903

> confint.lm(lm(SMR~Year,data=subset(mydata, Age_Group=='75+ years')))
                                        2.5 %   97.5 %
(Intercept)   -62.68087226 78.98676302
95% CI for ∆annual age-specific mortality ratio -0.03904813 0.03136318

S4e - Trend in age-specific mortality ratios in men

Call:
lm(formula = SMR ~ Year, data = subset(mydata5, Age_Group ==
"16-24 years"))

Residuals:
      1   2   3   4   5   6   7
-2.74286 3.24286 0.12857 0.01429 0.80000 -1.91429 0.47143

Coefficients:
                                Estimate Std. Error t value Pr(>|t|)
                    (Intercept) -225.9571   808.0786 -0.280   0.791
∆annual age-specific mortality ratio 0.1143 0.4016 0.285   0.787

```
Residual standard error: 2.125 on 5 degrees of freedom
Multiple R-squared:  0.01594,    Adjusted R-squared:  -0.1809
F-statistic: 0.08097 on 1 and 5 DF,  p-value: 0.7874

>
> #in this, year shows the lower and upper unit CI for the above, meaning that if they cross 0,
> #there is no significant linear increase or decrease (in SMR here) of the period studies
>
> confint.lm(lm(SMR~Year,data=subset(mydata5, Age_Group=='16-24 years')))
                                          2.5 %      97.5 %
                              (Intercept) -2303.1893666 1851.275081
95% CI for ∆annual age-specific mortality ratio        -0.9181354    1.146707

Call:
lm(formula = SMR ~ Year, data = subset(mydata5, Age_Group ==
    "25-34 years"))

Residuals:
     1      2      3      4      5      6      7
-1.3750  0.4357  0.7464  1.6571 -0.6321 -0.6214 -0.2107

Coefficients:
                                Estimate Std. Error t value Pr(>|t|)
                    (Intercept) 427.7000   423.8843   1.009    0.359
∆annual age-specific mortality ratio      -0.2107     0.2107  -1.000    0.363

Residual standard error: 1.115 on 5 degrees of freedom
Multiple R-squared:  0.1667,      Adjusted R-squared:  5.747e-05
F-statistic:     1 on 1 and 5 DF,  p-value: 0.3631

> confint.lm(lm(SMR~Year,data=subset(mydata5, Age_Group=='25-34 years')))
                                          2.5 %       97.5 %
                              (Intercept) -661.9292072 1517.3292072
95% CI for ∆annual age-specific mortality ratio       -0.7522792    0.3308507
>
> #35-44 years
> summary(lm(SMR~Year,data=subset(mydata5,Age_Group =='35-44 years')))

Call:
lm(formula = SMR ~ Year, data = subset(mydata5, Age_Group ==
    "35-44 years"))

Residuals:
      1       2       3       4       5       6       7
 0.08929  0.13571  0.08214 -0.17143 -0.52500  0.02143  0.36786

Coefficients:
                                Estimate Std. Error t value Pr(>|t|)
                    (Intercept) -306.71429  117.06326  -2.620   0.0471 *
∆annual age-specific mortality ratio    0.15357    0.05818   2.639   0.0460 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.3079 on 5 degrees of freedom
Multiple R-squared:  0.5822,      Adjusted R-squared:  0.4986
F-statistic: 6.967 on 1 and 5 DF,  p-value: 0.04601

> confint.lm(lm(SMR~Year,data=subset(mydata5, Age_Group=='35-44 years')))
                                          2.5 %      97.5 %
                              (Intercept) -6.076350e+02 -5.7935927
95% CI for ∆annual age-specific mortality ratio      4.008534e-03  0.3031343
```

```
>
> #45-54 years
> summary(lm(SMR~Year,data=subset(mydata5,Age_Group =='45-54 years')))

Call:
lm(formula = SMR ~ Year, data = subset(mydata5, Age_Group ==
    "45-54 years"))

Residuals:
      1       2       3       4       5       6       7
 0.13929  0.03571  0.03214 -0.17143 -0.07500 -0.47857  0.51786

Coefficients:
                                  Estimate Std. Error t value Pr(>|t|)
                       (Intercept)  -5.614286 126.558424  -0.044    0.966
∆annual age-specific mortality ratio       0.003571   0.062902   0.057    0.957

Residual standard error: 0.3328 on 5 degrees of freedom
Multiple R-squared:  0.0006443, Adjusted R-squared:  -0.1992
F-statistic: 0.003224 on 1 and 5 DF,  p-value: 0.9569

> confint.lm(lm(SMR~Year,data=subset(mydata5, Age_Group=='45-54 years')))
                                              2.5 %      97.5 %
                                   (Intercept) -330.9430703 319.7144989
95% CI for ∆annual age-specific mortality ratio         -0.1581227   0.1652656
>
> #55-64 years
> summary(lm(SMR~Year,data=subset(mydata5,Age_Group =='55-64 years')))

Call:
lm(formula = SMR ~ Year, data = subset(mydata5, Age_Group ==
    "55-64 years"))

Residuals:
      1       2       3       4       5       6       7
-0.01786 -0.05000  0.21786 -0.21429 -0.04643  0.22143 -0.11071

Coefficients:
                                  Estimate Std. Error t value Pr(>|t|)
                       (Intercept) -63.65714   67.94209  -0.937    0.392
∆annual age-specific mortality ratio       0.03214    0.03377   0.952    0.385

Residual standard error: 0.1787 on 5 degrees of freedom
Multiple R-squared:  0.1534,       Adjusted R-squared:  -0.01591
F-statistic: 0.906 on 1 and 5 DF,  p-value: 0.3849

> confint.lm(lm(SMR~Year,data=subset(mydata5, Age_Group=='55-64 years')))
                                              2.5 %      97.5 %
                                   (Intercept) -238.30785201 110.9935663
95% CI for ∆annual age-specific mortality ratio      -0.05466163   0.1189473
>
> #65-74 years
> summary(lm(SMR~Year,data=subset(mydata5,Age_Group =='65-74 years')))

Call:
lm(formula = SMR ~ Year, data = subset(mydata5, Age_Group ==
    "65-74 years"))

Residuals:
      1       2       3       4       5       6       7
-0.02857 -0.14286  0.14286  0.12857  0.11429 -0.30000  0.08571
```

```
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
                (Intercept) -229.2714    69.8162  -3.284   0.0219 *
∆annual age-specific mortality ratio     0.1143     0.0347   3.294   0.0216 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.1836 on 5 degrees of freedom
Multiple R-squared:  0.6845,      Adjusted R-squared:  0.6214
F-statistic: 10.85 on 1 and 5 DF,  p-value: 0.02163

> confint.lm(lm(SMR~Year,data=subset(mydata5, Age_Group=='65-74 years')))
                                        2.5 %      97.5 %
                            (Intercept) -408.73966333 -49.8031938
95% CI for ∆annual age-specific mortality ratio     0.02508683   0.2034846
>
> #75+ years
> summary(lm(SMR~Year,data=subset(mydata5,Age_Group =='75+ years')))

Call:
lm(formula = SMR ~ Year, data = subset(mydata5, Age_Group ==
    "75+ years"))

Residuals:
      1       2       3       4       5       6       7
-0.07857  0.12857  0.03571 -0.05714 -0.05000 -0.04286  0.06429

Coefficients:
                                Estimate Std. Error t value Pr(>|t|)
                    (Intercept) 14.728571  32.135507   0.458    0.666
∆annual age-specific mortality ratio    -0.007143   0.015972  -0.447    0.673

Residual standard error: 0.08452 on 5 degrees of freedom
Multiple R-squared:  0.03846,     Adjusted R-squared:  -0.1538
F-statistic:   0.2 on 1 and 5 DF,  p-value: 0.6734

> confint.lm(lm(SMR~Year,data=subset(mydata5, Age_Group=='75+ years')))
                                        2.5 %      97.5 %
                            (Intercept) -67.87837939 97.33552225
95% CI for ∆annual age-specific mortality ratio       -0.04819997  0.03391426
```

S4f - Trend in age-specific mortality ratios in women

```
Call:
lm(formula = SMR ~ Year, data = subset(mydata6, Age_Group ==
   "16-24 years"))

Residuals:
    1     2     3     4     5     6     7
-4.864  3.671 -2.193 10.743 -4.821 -2.986  0.450

Coefficients:
                                Estimate Std. Error t value Pr(>|t|)
                    (Intercept) -723.2857  2342.1913  -0.309    0.770
∆annual age-specific mortality ratio      0.3643     1.1641   0.313    0.767

Residual standard error: 6.16 on 5 degrees of freedom
Multiple R-squared:  0.01921,     Adjusted R-squared:  -0.1769
F-statistic: 0.09793 on 1 and 5 DF,  p-value: 0.767
```

```
> confint.lm(lm(SMR~Year,data=subset(mydata6, Age_Group=='16-24 years')))
                                    2.5 %      97.5 %
                (Intercept)  -6744.080163  5297.508735
95% CI for ∆annual age-specific mortality ratio      -2.628155    3.356727

> #25-34 years
> summary(lm(SMR~Year,data=subset(mydata6,Age_Group =='25-34 years')))

Call:
lm(formula = SMR ~ Year, data = subset(mydata6, Age_Group ==
    "25-34 years"))

Residuals:
     1       2       3       4       5       6       7
 0.5321 -0.7929 -0.1179  0.6571 -0.9679  1.2071 -0.5179

Coefficients:
                                Estimate Std. Error t value Pr(>|t|)
                (Intercept) -449.0571   340.9243  -1.317    0.245
∆annual age-specific mortality ratio       0.2250     0.1694   1.328    0.242

Residual standard error: 0.8966 on 5 degrees of freedom
Multiple R-squared:  0.2607,      Adjusted R-squared:  0.1128
F-statistic: 1.763 on 1 and 5 DF,  p-value: 0.2416

> confint.lm(lm(SMR~Year,data=subset(mydata6, Age_Group=='25-34 years')))
                                    2.5 %      97.5 %
                (Intercept) -1325.4308542 427.3165684
95% CI for ∆annual age-specific mortality ratio       -0.2105732   0.6605732
>


> #35-44 years
> summary(lm(SMR~Year,data=subset(mydata6,Age_Group =='35-44 years')))

Call:
lm(formula = SMR ~ Year, data = subset(mydata6, Age_Group ==
    "35-44 years"))

Residuals:
     1       2       3       4       5       6       7
 0.4321 -0.2071  0.3536 -0.7857 -0.2250 -0.1643  0.5964

Coefficients:
                                Estimate Std. Error t value Pr(>|t|)
                (Intercept) 124.04286  201.58468   0.615    0.565
∆annual age-specific mortality ratio      -0.06071    0.10019  -0.606    0.571

Residual standard error: 0.5302 on 5 degrees of freedom
Multiple R-squared:  0.06842,     Adjusted R-squared:  -0.1179
F-statistic: 0.3672 on 1 and 5 DF,  p-value: 0.571

> confint.lm(lm(SMR~Year,data=subset(mydata6, Age_Group=='35-44 years')))
                                    2.5 %      97.5 %
                (Intercept) -394.1470622 642.2327765
95% CI for ∆annual age-specific mortality ratio       -0.3182638   0.1968352
>

> #45-54 years
> summary(lm(SMR~Year,data=subset(mydata6,Age_Group =='45-54 years')))

Call:
```

```
lm(formula = SMR ~ Year, data = subset(mydata6, Age_Group ==
   "45-54 years"))

Residuals:
     1      2      3      4      5      6      7
-0.3929 -0.5000  0.8929  0.1857  0.2786  0.1714 -0.6357

Coefficients:
                                  Estimate Std. Error t value Pr(>|t|)
                      (Intercept) -12.757143 224.764843  -0.057    0.957
∆annual age-specific mortality ratio      0.007143   0.111712   0.064    0.951

Residual standard error: 0.5911 on 5 degrees of freedom
Multiple R-squared:  0.000817,   Adjusted R-squared:  -0.199
F-statistic: 0.004088 on 1 and 5 DF,  p-value: 0.9515

> confint.lm(lm(SMR~Year,data=subset(mydata6, Age_Group=='45-54 years')))
                                        2.5 %      97.5 %
                          (   Intercept) -590.5335661 565.0192803
95% CI for ∆annual age-specific mortality ratio     -0.2800222   0.2943079
>
> #55-64 years
> summary(lm(SMR~Year,data=subset(mydata6,Age_Group =='55-64 years')))

Call:
lm(formula = SMR ~ Year, data = subset(mydata6, Age_Group ==
   "55-64 years"))

Residuals:
      1       2       3       4       5       6       7
 0.22143 -0.05714  0.16429 -0.31429 -0.49286  0.22857  0.25000

Coefficients:
                                  Estimate Std. Error t value Pr(>|t|)
                      (Intercept) -358.17143  124.46028  -2.878   0.0347 *
∆annual age-specific mortality ratio   0.17857    0.06186   2.887   0.0343 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.3273 on 5 degrees of freedom
Multiple R-squared:  0.625,        Adjusted R-squared:   0.55
F-statistic: 8.333 on 1 and 5 DF,  p-value: 0.03432

> confint.lm(lm(SMR~Year,data=subset(mydata6, Age_Group=='55-64 years')))
                                        2.5 %      97.5 %
                          (Intercept) -678.10677338 -38.2360838
95% CI for ∆annual age-specific mortality ratio      0.01955792   0.3375849


> #65-74 years
> summary(lm(SMR~Year,data=subset(mydata6,Age_Group =='65-74 years')))

Call:
lm(formula = SMR ~ Year, data = subset(mydata6, Age_Group ==
   "65-74 years"))

Residuals:
      1       2       3       4       5       6       7
 0.22143  0.08571 -0.35000  0.01429 -0.22143  0.04286  0.20714

Coefficients:
                                  Estimate Std. Error t value Pr(>|t|)
```

```
                                  (Intercept) 130.02857   88.82431   1.464    0.203
∆annual age-specific mortality ratio    -0.06429    0.04415  -1.456    0.205

Residual standard error: 0.2336 on 5 degrees of freedom
Multiple R-squared:  0.2978,       Adjusted R-squared:  0.1574
F-statistic:  2.12 on 1 and 5 DF,  p-value: 0.2051

> confint.lm(lm(SMR~Year,data=subset(mydata6, Age_Group=='65-74 years')))
                                              2.5 %      97.5 %
                                  (Intercept) -98.3015834 358.3587262
95% CI for ∆annual age-specific mortality ratio      -0.1777698   0.0491984


> #75+ years
> summary(lm(SMR~Year,data=subset(mydata6,Age_Group =='75+ years')))

Call:
lm(formula = SMR ~ Year, data = subset(mydata6, Age_Group ==
    "75+ years"))

Residuals:
       1        2        3        4        5        6        7
-0.107143  0.100000  0.007143  0.114286 -0.078571 -0.071429  0.035714

Coefficients:
                                       Estimate Std. Error t value Pr(>|t|)
                                  (Intercept) 14.857143  36.920887   0.402    0.704
∆annual age-specific mortality ratio     -0.007143   0.018350  -0.389    0.713

Residual standard error: 0.0971 on 5 degrees of freedom
Multiple R-squared:  0.02941,     Adjusted R-squared:  -0.1647
F-statistic: 0.1515 on 1 and 5 DF,  p-value: 0.7131

> confint.lm(lm(SMR~Year,data=subset(mydata6, Age_Group=='75+ years')))
                                              2.5 %       97.5 %
                                  (Intercept) -80.05101795 109.76530366
95% CI for ∆annual age-specific mortality ratio     -0.05431389   0.04002817
```

Appendix 5 – Supplementary material to Chapter 6

Appendix S1 - Scottish Epilepsy-related Deaths Study (SEDS) data collection pro formas

Patient's study ID……..……..

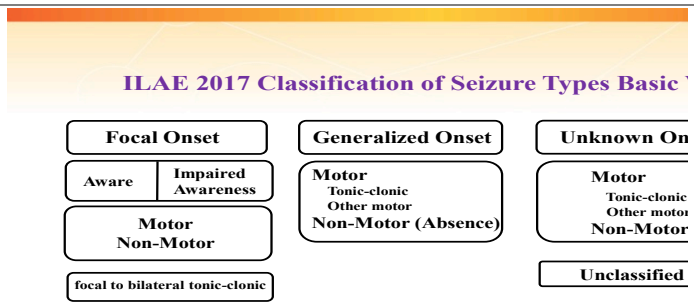
Date pro forma completed……..

TO THE BEST OF YOUR KNOWLEDGE AND USING THE PATIENT'S MEDICAL RECORDS, PLEASE ANSWER THE QUESTIONS BELOW, CIRCLING YES/ NO / NOT KNOWN WHERE RELEVANT, AND GIVE AS MUCH RELEVANT ADDITIONAL INFORMATION AS YOU CAN THROUGHOUT.

CASES

Did the patient have epilepsy? YES/NO – if not, go on to next patient

How did death happen? – describe the sequence of events leading up to death	
Include details on the following: Where did death occur?	**In hospital // out of hospital**
What were the illnesses leading up to death?	
What caused death?	
When was the last seizure?	
Any evidence that seizure lead to death? E.g. lead to aspiration pneumonia? Fall? Drowning? Injury? Suicide? Depression? Cardiac arrhythmias? Or that antiepileptic drugs/their side effects contributed to death?	**Yes // no // not known**
Any evidence of increased seizure frequency in the month leading up to death?	**Yes // no // not known**
Any changes to their antiepileptic drug script in the month leading up to death?	**Yes // no // not known**
Was the death witnessed?	Yes // no // not known
Any family history of sudden unexpected death?	Yes // no // not known
Date last seen in hospital? Type of contact that was? –eg clinic/A&E/scan etc	
What was their state of employment when last seen? (please circle)	Employed // Unemployed // Retired // Not known
Age at first seizure? (In years)	
Age when epilepsy diagnosed?	*(In years)*
Who made the epilepsy diagnosis? –eg GP/neurologist/geriatrician etc	
Any past medical history of risk factors for epilepsy? (Please circle all that apply)	Prolonged febrile convulsions Bacterial meningitis Encephalitis Significant head injury 1st degree relative with epilepsy (mum/dad/siblings) Brain tumour Cerebrovascular disease (ie CVA/TIA) Congenital abnormality eg CP/metabolic infancy syndrome ADHD Autistic spectrum disorder Developmental/intellectual delay Paraneoplastic or autoimmune/limbic encephalitis Genetic syndrome Perinatal and or Birth difficulties including Hypoxic Ischaemic Encephalopathy
Aside from the above, any comorbidities other than	Yes // no // not known

epilepsy known about when last seen?	
Any evidence of alcohol abuse in the 2 years prior to death? (Including documented Etoh excess history plus minus deranged LFTs/GGTor use of alcohol abuse services)	Yes // no // not known
Any evidence of illicit drug abuse/addiction in the 2 years prior to death? (including documented medical history of/taking methadone)	Yes // no // not known
Any history of depression in the 2 years prior to death? (including documented history/NIDDE/HADS/on an antidepressant)	Yes // no // not recorded
Any history of anxiety in the 2 years prior to death? (including documented history/NIDDE/HADS)	Yes // no // not recorded
Any history of deliberate self harm?	Yes // no // not known
Any suicide attempts in the 2 years prior to death? If so, how many/when?	Yes // no // not known
Any evidence of contact with psychiatric services in the 2 years prior to death?	Yes // no // not known
What type of seizures did they have? (circle all that apply)	 From *Fisher et al. Instruction manual for the ILAE 2017 operationa classification of seizure types. Epilepsia doi: 10.1111/epi.13671*
What type of epilepsy did they have?	Focal Generalised Combined generalised and focal Not known
Were they known to have an epilepsy syndrome? If so, which one?	
What was the aetiology of their epilepsy?	Structural Genetic Infectious Metabolic Immune Unknown
What was their seizure frequency when last seen? (or when last documented)	1 per week or more 1 per month or more but less than 1 per week 1 per year or more, but less than 1 per month Other (please state):
Did seizure frequency stay the same/get better/get worse over the 2 years prior to death.	
When did they have seizures mainly?	Daytime only Night time only Mixture of daytime and nighttime
How many antiepileptic drugs were they taking when last seen?	0 1 2 >2

What drugs were they? Include doses if possible. *Were any drugs changed/discontinued in the 2 years prior to death? If so, circle reasons:*	Side effects Drug ineffective Both Reason unclear		
Has the participant ever had brain surgery to treat or control seizures?	Yes No Was considered for surgery but did not get it (include reason where possible) Not known		
If the following surgical procedures were used to treat seizures, select the status, outcome and the age of surgery. (Select all that apply	Procedure		Outcome
If the following surgical procedures were used to treat seizures, select the status, outcome and the age of surgery. (Select all that apply *Does the patient have an seizure alarm?*	Corpus callosotomy Gamma knife radiosurgery Hemispherectomy Lobectomy Resection surgery Subpial transection Other (please state): Not known		Seizure free Greatly reduced frequency of seizures Slightly reduced frequency of seizures Did not change frequency of seizures Not known
If the following surgical procedures were used to treat seizures, select the status, outcome and the age of surgery. (Select all that apply *Does the patient have an seizure alarm?* *Number of A&E attendances to for epilepsy/seizure-related conditions in the 1 years prior to death? (If known)*	Yes No Not known	Age at surgery	
	1 2 ≥3	0 – 1 2-5 5-15 16-20 20-30 30-50 >50 Not known	
Number of inpatient admissions to hospital for epilepsy/seizure-related conditions in the previous 1 years	1 2 ≥3		
Total number of admissions to hospital in the previous 2 years	1 2 ≥3		
Patient seen by an epilepsy specialist in the previous 1 years?.	Yes // no // not known		
Patient seen by a epilepsy nurse specialist in the previous 1 years?	Yes // no // not known		
Documented clear treatment plan in notes after last encounter in epilepsy clinic/GP?	Yes // no // not known		
Did patient get sent a copy of hospital letter after last seen in hospital?	Yes // no // not known		
Any history of seizure clustering in the previous 2 years? (ie 3 or more seizures in a 24 hour period)	Yes // no // not known		
Any history of prolonged GTC seizure in the previous 2 years? (ie seizure > 3 min)	Yes // no // not known		
Any history of status epilepticus in the previous 2 years?	Yes // no // not known		
Any evidence of use of buccal midazolam out of hospital for those with history of seizure clustering/prolonged GTCS/Status in previous 2 years?	Yes // no // not known		

Any evidence of training in the use of buccal midazolam for those with history of seizure clustering/prolonged GTCS/Status in previous 2 years?	Yes // no // not known

Scottish Epilepsy-related Deaths Study (SEDS) data collection pro forma

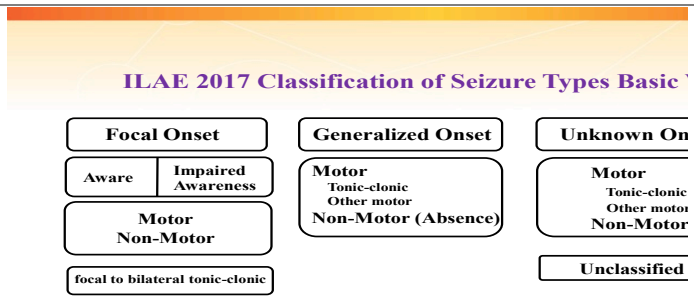
Patient's STUDY NUMBER……..

Patient's hospital……..

CONTROLS

Did the patient have epilepsy? YES/NO – if not, go on to next patient

Patient's age	
How long since they were last seen 1) as a routine and 2) as an emergency	
What and when was their latest epilepsy illnesses?	
When was the last seizure?	
Any evidence of increased seizure frequency in the month prior to last seen?	
Any changes to their antiepileptic drug script in the month prior to last seen?	
If so, what	
Any family history of sudden unexpected death?	Yes // no // not known
What was their state of employment when last seen? (please circle)	Employed // Unemployed // Retired // Not known
Age at first seizure? (In years)	
Age when epilepsy diagnosed?	*(In years)*
Who made the epilepsy diagnosis? –eg GP/neurologist/geriatrician etc	
Any past medical history of risk factors for epilepsy? (Please circle all that apply)	Prolonged febrile convulsions Bacterial meningitis Encephalitis Significant head injury 1st degree relative with epilepsy (mum/dad/siblings) Brain tumour Cerebrovascular disease (ie CVA/TIA) Congenital abnormality eg CP/metabolic infancy syndrome ADHD Autistic spectrum disorder Developmental/intellectual delay Paraneoplastic or autoimmune/limbic encephalitis Genetic syndrome Perinatal and or Birth difficulties including Hypoxic Ischaemic Encephalopathy
Aside from the above, any comorbidities other than epilepsy known about when last seen?	Yes // no // not known
Any evidence of alcohol abuse in the 2 years prior to death? (Including documented Etoh excess history plus minus deranged LFTs/GGTor use of alcohol abuse services)	Yes // no // not known

Any evidence of illicit drug abuse/addiction in the 2 years prior to death? (including documented medical history of/taking methadone)	Yes // no // not known
Any history of depression in the 2 years prior to death? (including documented history/NIDDE/HADS/on an antidepressant)	Yes // no // not recorded
Any history of anxiety in the 2 years prior to death? (including documented history/NIDDE/HADS)	Yes // no // not recorded
Any history of deliberate self harm?	Yes // no // not known
Any suicide attempts in the 2 years prior to death? If so, how many/when?	Yes // no // not known
Any evidence of contact with psychiatric services in the 2 years prior to death?	Yes // no // not known
What type of seizures did they have? (circle all that apply)	 From *Fisher et al. Instruction manual for the ILAE 2017 operationa classification of seizure types. Epilepsia doi: 10.1111/epi.13671*
What type of epilepsy did they have?	Focal Generalised Combined generalised and focal Not known
Were they known to have an epilepsy syndrome? If so, which one?	
What was the aetiology of their epilepsy?	Structural Genetic Infectious Metabolic Immune Unknown
What was their seizure frequency when last seen? (or when last documented)	1 per week or more 1 per month or more but less than 1 per week 1 per year or more, but less than 1 per month Other (please state):
Did seizure frequency stay the same/get better/get worse over the past 2 years?	
When did they have seizures mainly?	Daytime only Night time only Mixture of daytime and nighttime
How many antiepileptic drugs were they taking when last seen? *What drugs were they? Include doses if possible.*	0 1 2 >2
Were any drugs changed/discontinued in the 2 years prior to death? If so, circle reasons:	Side effects Drug ineffective Both Reason unclear

Has the participant ever had brain surgery to treat or control seizures?	Yes No Was considered for surgery but did not get it (include reason where possible) Not known		
If the following surgical procedures were used to treat seizures, select the status, outcome and the age of surgery. (Select all that apply	Procedure Corpus callosotomy Gamma knife radiosurgery Hemispherectomy Lobectomy Resection surgery Subpial transection Other (please state): Not known	Age at surgery 0 – 1 2-5 5-15 16-20 20-30 30-50 >50 Not known	Outcome Seizure free Greatly reduced frequency of seizures Slightly reduced frequency of seizures Did not change frequency of seizures Not known
Does the patient have an seizure alarm?	Yes //No // Not known		
Number of A&E attendances to for epilepsy/seizure-related conditions in the past 1 year?			
Number of inpatient admissions to hospital for epilepsy/seizure-related conditions in the previous 1 year			
Total number of admissions to hospital in the previous 2 years			
Patient seen by an epilepsy specialist in the previous 1 year?.	Yes // no // not known		
Patient seen by a epilepsy nurse specialist in the previous 1 year?	Yes // no // not known		
Documented clear treatment plan in notes after last encounter in epilepsy clinic/GP?	Yes // no // not known		
Did patient get sent a copy of hospital letter after last seen in hospital?	Yes // no // not known		
Any history of seizure clustering in the previous 2 years? (ie 3 or more seizures in a 24 hour period)	Yes // no // not known		
Any history of prolonged GTC seizure in the previous 2 years? (ie seizure > 3 min)	Yes // no // not known		
Any history of status epilepticus in the previous 2 years?	Yes // no // not known		
Any evidence of use of buccal midazolam out of hospital for those with history of seizure clustering/prolonged GTCS/Status in previous 2 years?	Yes // no // not known		
Any evidence of training in the use of buccal midazolam for those with history of seizure clustering/prolonged GTCS/Status in previous 2 years?	Yes // no // not known		
SUDEP ever mentioned?	Yes // no // not known		
Drowning/bath safety ever mentioned?	Yes // no // not known		
Driving risks ever mentioned?	Yes // no // not known		
Any other epilepsy mortality risks ever mentioned? If so, which			

Appendix S2 – Discrimination estimates

A - *Multivariable model*

Iteration	Statistic	Value	95% CI
Original data	Sensitivity	68.31%	61.03% to 74.97%
	Specificity	80.11%	73.64% to 85.59%
	PPV	77.16%	71.35% to 82.09%
	NPV	71.98%	67.24% to 76.28%
	Accuracy	74.25%	69.47% to 78.64%
Imputation 1			
	Sensitivity	69.64%	63.17% to 75.59%
	Specificity	76.34%	70.22% to 81.75%
	PPV	74.64%	69.61% to 79.09%
	NPV	71.55%	67.06% to 75.65%
	Accuracy	72.99%	68.62% to 77.05%
Imputation 2			
	Sensitivity	67.86%	61.31% to 73.92%
	Specificity	76.79%	70.70% to 82.15%
	PPV	74.51%	69.38% to 79.04%
	NPV	70.49%	66.09% to 74.54%
	Accuracy	72.32%	67.93% to 76.42%
Imputation 3			
	Sensitivity	67.86%	61.31% to 73.92%
	Specificity	76.34%	70.22% to 81.75%
	PPV	74.15%	69.03% to 78.68%
	NPV	70.37%	65.95% to 74.44%
	Accuracy	72.10%	67.70% to 76.21%
Imputation 4			
	Sensitivity	70.09%	63.63% to 76.01%
	Specificity	75.45%	69.27% to 80.94%
	PPV	74.06%	69.08% to 78.48%
	NPV	71.61%	67.07% to 75.75%
	Accuracy	72.77%	68.39% to 76.84%
Imputation 5			
	Sensitivity	67.41%	60.85% to 73.50%
	Specificity	76.34%	70.22% to 81.75%
	PPV	74.02%	68.89% to 78.57%
	NPV	70.08%	65.68% to 74.14%
	Accuracy	71.88%	67.46% to 75.99%
Imputation 6			
	Sensitivity	68.75%	62.24% to 74.76%
	Specificity	75.89%	69.75% to 81.34%
	PPV	74.04%	68.98% to 78.53%

	NPV	70.83%	66.36% to 74.93%
	Accuracy	72.32%	67.93% to 76.42%
Imputation 7			
	Sensitivity	68.30%	61.77% to 74.34%
	Specificity	76.79%	70.70% to 82.15%
	PPV	74.63%	69.53% to 79.14%
	NPV	70.78%	66.36% to 74.84%
	Accuracy	72.54%	68.16% to 76.63%
Imputation 8			
	Sensitivity	68.75%	62.24% to 74.76%
	Specificity	75.89%	69.75% to 81.34%
	PPV	74.04%	68.98% to 78.53%
	NPV	70.83%	66.36% to 74.93%
	Accuracy	72.32%	67.93% to 76.42%
Imputation 9			
	Sensitivity	67.41%	60.85% to 73.50%
	Specificity	76.34%	70.22% to 81.75%
	PPV	74.02%	68.89% to 78.57%
	NPV	70.08%	65.68% to 74.14%
	Accuracy	71.88%	67.46% to 75.99%
Imputation 10			
	Sensitivity	70.09%	63.63% to 76.01%
	Specificity	75.89%	69.75% to 81.34%
	PPV	74.41%	69.42% to 78.83%
	NPV	71.73%	67.21% to 75.85%
	Accuracy	72.99%	68.62% to 77.05%
Imputation 11			
	Sensitivity	66.96%	60.39% to 73.09%
	Specificity	76.34%	70.22% to 81.75%
	PPV	73.89%	68.74% to 78.46%
	NPV	69.80%	65.42% to 73.84%
	Accuracy	71.65%	67.23% to 75.78%
Imputation 12			
	Sensitivity	66.96%	60.39% to 73.09%
	Specificity	76.79%	70.70% to 82.15%
	PPV	74.26%	69.08% to 78.83%
	NPV	69.92%	65.56% to 73.95%
	Accuracy	71.88%	67.46% to 75.99%
Imputation 13			
	Sensitivity	69.64%	63.17% to 75.59%
	Specificity	75.89%	69.75% to 81.34%
	PPV	74.29%	69.27% to 78.73%
	NPV	71.43%	66.92% to 75.55%

	Accuracy	72.77%	68.39% to 76.84%
Imputation 14			
	Sensitivity	69.20%	62.70% to 75.17%
	Specificity	75.89%	69.75% to 81.34%
	PPV	74.16%	69.13% to 78.63%
	NPV	71.13%	66.64% to 75.24%
	Accuracy	72.54%	68.16% to 76.63%
Imputation 15			
	Sensitivity	70.09%	63.63% to 76.01%
	Specificity	75.45%	69.27% to 80.94%
	PPV	74.06%	69.08% to 78.48%
	NPV	71.61%	67.07% to 75.75%
	Accuracy	72.77%	68.39% to 76.84%
Imputation 16			
	Sensitivity	70.54%	64.10% to 76.42%
	Specificity	75.89%	69.75% to 81.34%
	PPV	74.53%	69.56% to 78.93%
	NPV	72.03%	67.49% to 76.17%
	Accuracy	73.21%	68.86% to 77.26%
Imputation 17			
	Sensitivity	68.30%	61.77% to 74.34%
	Specificity	75.89%	69.75% to 81.34%
	PPV	73.91%	68.84% to 78.42%
	NPV	70.54%	66.09% to 74.63%
	Accuracy	72.10%	67.70% to 76.21%
Imputation 18			
	Sensitivity	70.09%	63.63% to 76.01%
	Specificity	75.89%	69.75% to 81.34%
	PPV	74.41%	69.42% to 78.83%
	NPV	71.73%	67.21% to 75.85%
	Accuracy	72.99%	68.62% to 77.05%
Imputation 19			
	Sensitivity	70.09%	63.63% to 76.01%
	Specificity	77.23%	71.18% to 82.55%
	PPV	75.48%	70.44% to 79.90%
	NPV	72.08%	67.61% to 76.16%
	Accuracy	73.66%	69.32% to 77.68%
Imputation 20			
	Sensitivity	69.64%	63.17% to 75.59%
	Specificity	76.79%	70.70% to 82.15%
	PPV	75.00%	69.96% to 79.45%
	NPV	71.67%	67.19% to 75.75%
	Accuracy	73.21%	68.86% to 77.26%

B – *SEDS Score model*

Original data	Statistic	Value	95% CI
	Sensitivity	70.83%	64.28% to 76.80%
	Specificity	65.75%	59.06% to 72.01%
	PPV	67.11%	62.49% to 71.41%
	NPV	69.57%	64.52% to 74.18%
	Accuracy	68.28%	63.67% to 72.63%
Imputation 1			
	Sensitivity	72.77%	66.44% to 78.48%
	Specificity	64.29%	57.63% to 70.56%
	PPV	67.08%	62.68% to 71.19%
	NPV	70.24%	65.11% to 74.92%
	Accuracy	68.53%	64.00% to 72.80
Imputation 2			
	Sensitivity	72.77%	66.44% to 78.48%
	Specificity	64.29%	57.63% to 70.56%
	PPV	67.08%	62.68% to 71.19%
	NPV	70.24%	65.11% to 74.92%
	Accuracy	68.53%	64.00% to 72.80
Imputation 3			
	Sensitivity	72.32%	65.97% to 78.07%
	Specificity	63.39%	56.72% to 69.71%
	PPV	66.39%	62.02% to 70.50%
	NPV	69.61%	64.45% to 74.32%
	Accuracy	67.86%	63.31% to 72.16%
Imputation 4			
	Sensitivity	72.32%	65.97% to 78.07%
	Specificity	63.39%	56.72% to 69.71%
	PPV	66.39%	62.02% to 70.50%
	NPV	69.61%	64.45% to 74.32%
	Accuracy	67.86%	63.31% to 72.16%
Imputation 5			
	Sensitivity	72.32%	65.97% to 78.07%
	Specificity	63.39%	56.72% to 69.71%
	PPV	66.39%	62.02% to 70.50%
	NPV	69.61%	64.45% to 74.32%
	Accuracy	67.86%	63.31% to 72.16%
Imputation 6			
	Sensitivity	72.32%	65.97% to 78.07%
	Specificity	63.39%	56.72% to 69.71%
	PPV	66.39%	62.02% to 70.50%
	NPV	69.61%	64.45% to 74.32%

	Accuracy	67.86%	63.31% to 72.16%
Imputation 7			
	Sensitivity	72.32%	65.97% to 78.07%
	Specificity	63.39%	56.72% to 69.71%
	PPV	66.39%	62.02% to 70.50%
	NPV	69.61%	64.45% to 74.32%
	Accuracy	67.86%	63.31% to 72.16%
Imputation 8			
	Sensitivity	72.32%	65.97% to 78.07%
	Specificity	63.39%	56.72% to 69.71%
	PPV	66.39%	62.02% to 70.50%
	NPV	69.61%	64.45% to 74.32%
	Accuracy	67.86%	63.31% to 72.16%
Imputation 9			
	Sensitivity	72.32%	65.97% to 78.07%
	Specificity	63.39%	56.72% to 69.71%
	PPV	66.39%	62.02% to 70.50%
	NPV	69.61%	64.45% to 74.32%
	Accuracy	67.86%	63.31% to 72.16%
Imputation 10			
	Sensitivity	72.32%	65.97% to 78.07%
	Specificity	63.39%	56.72% to 69.71%
	PPV	66.39%	62.02% to 70.50%
	NPV	69.61%	64.45% to 74.32%
	Accuracy	67.86%	63.31% to 72.16%
imputation 11			
	Sensitivity	72.32%	65.97% to 78.07%
	Specificity	63.39%	56.72% to 69.71%
	PPV	66.39%	62.02% to 70.50%
	NPV	69.61%	64.45% to 74.32%
	Accuracy	67.86%	63.31% to 72.16%
Imputation 12			
	Sensitivity	72.32%	65.97% to 78.07%
	Specificity	63.39%	56.72% to 69.71%
	PPV	66.39%	62.02% to 70.50%
	NPV	69.61%	64.45% to 74.32%
	Accuracy	67.86%	63.31% to 72.16%
Imputation 13			
	Sensitivity	72.32%	65.97% to 78.07%
	Specificity	63.39%	56.72% to 69.71%
	PPV	66.39%	62.02% to 70.50%
	NPV	69.61%	64.45% to 74.32%
	Accuracy	67.86%	63.31% to 72.16%

Imputation 14			
	Sensitivity	71.88%	65.50% to 77.66%
	Specificity	64.29%	57.63% to 70.56%
	PPV	66.80%	62.38% to 70.96%
	NPV	69.57%	64.47% to 74.22%
	Accuracy	68.08%	63.54% to 72.38%
Imputation 15			
	Sensitivity	73.66%	67.38% to 79.30%
	Specificity	63.39%	56.72% to 69.71%
	PPV	66.80%	62.48% to 70.86%
	NPV	70.65%	65.42% to 75.38%
	Accuracy	68.53%	64.00% to 72.80%
imputation 16			
	Sensitivity	73.66%	67.38% to 79.30%
	Specificity	64.29%	57.63% to 70.56%
	PPV	67.35%	62.98% to 71.43%
	NPV	70.94%	65.76% to 75.62%
	Accuracy	68.97%	64.46% to 73.23%
Imputation 17			
	Sensitivity	74.55%	68.32% to 80.12%
	Specificity	63.39%	56.72% to 69.71%
	PPV	67.07%	62.78% to 71.09%
	NPV	71.36%	66.09% to 76.10%
	Accuracy	68.97%	64.46% to 73.23%
Imputation 18			
	Sensitivity	72.32%	65.97% to 78.07%
	Specificity	64.73%	58.09% to 70.98%
	PPV	67.22%	62.79% to 71.37%
	NPV	70.05%	64.95% to 74.69%
	Accuracy	68.53%	64.00% to 72.80%
Imputation 19			
	Sensitivity	72.32%	65.97% to 78.07%
	Specificity	64.29%	57.63% to 70.56%
	PPV	66.94%	62.53% to 71.08%
	NPV	69.90%	64.78% to 74.57%
	Accuracy	68.30%	63.77% to 72.59%
Imputation 20			
	Sensitivity	71.88%	65.50% to 77.66%
	Specificity	64.73%	58.09% to 70.98%
	PPV	67.08%	62.63% to 71.25%
	NPV	69.71%	64.63% to 74.35%
	Accuracy	68.30%	63.77% to 72.59%