Expert paediatric evidence in alleged infant harm: the truth, the whole truth, and nothing but the truth?

John Beattie

A thesis presented for the degree of
Doctor of Philosophy
School of Law
The University of Edinburgh

2018
In accordance with Regulation 25 of the Assessment Regulations for Research Degrees at the University of Edinburgh, I certify that this thesis I have presented for examination for the PhD degree of the University of Edinburgh is entirely my own work, except where explicitly indicated otherwise throughout the thesis by reference or acknowledgment. This work has not been submitted for any other degree or professional qualification.

__________________________________________
John Beattie

__________________________________________
Date
Abstract

The aim of this thesis is to investigate and further understanding of the interaction between paediatricians providing expert opinion evidence in cases of alleged child abuse and the various fact-findings tribunals who require their expertise.

My original contribution to this largely unexplored area of scholarship lies in the identification and analysis of potential frailties particular to the setting and process of paediatric expert assessment of cases of alleged child abuse that need to be recognised and accounted for in order to ensure that, as far as possible, expert paediatricians provide reliable opinion evidence.

Specifically, informed by extensive practical experience, I deconstruct and examine the various elements that together constitute the sequential process of expert diagnostic decision-making and opinion generation.

First, drawing on empirical research and theoretical constructs from a range of knowledge domains, particularly behavioural psychology and cognitive neuroscience, I consider emerging research on human factors that may influence clinical decision-making and forensic judgments, and apply this evolving understanding to the process of paediatric forensic interpretation. Here I expose a variety of potential biasing factors specific to the child protection setting that may impact on expert paediatricians’ rational judgments that an injury may abusive in origin.

I then move on to provide a detailed critique of the evidence base on which such paediatric expert clinical judgments and opinions must be based. Acknowledging the absence of a gold standard test for abuse, I question some fundamental aspects of the proxy indicators of abuse currently promoted as the basis for expert opinions, before discussing the challenges for legal fact-finders tasked with interpretation of those opinions by the absence of an agreed semantic hierarchy with which the certainty of expert conclusions might be qualified.
Finally, having shown that in the context of child abuse assessment there are indeed a variety of potential threats to paediatricians’ objective forensic judgments, I use alleged “shaken baby” cases as an exemplar topic to expose those frailties within such cases, and explore the challenges posed for the courts in dealing with complex and evolving paediatric expert evidence. I conclude that neither the law nor legal institutions can resolve these threats to objective expert forensic decision-making alone, and that such frailties need to be recognised and accounted for by both individual paediatric experts and the wider child protection community.
Lay Summary

This thesis seeks to improve understanding of a particularly challenging area of forensic practice – expert paediatric evidence in cases of alleged murder or abuse of infants and young children. Such cases are complex for a number of reasons, but perhaps most notably because those who come under suspicion are usually a child’s parent or caregiver. For the families involved, even if such allegations do not lead to criminal proceedings, the effects of such investigations are likely to be life-changing.

The courts depend almost exclusively on the expertise of paediatricians and other doctors to provide opinion evidence to help them decide whether any injuries found are natural or likely to have been inflicted. If experts get it wrong it may mean, on the one hand, an innocent parent going to jail, on the other, a vulnerable child returned to be further injured or killed. The stakes are high for all concerned.

From the perspective of an experienced paediatric expert witness, I examine various core components of the process of forensic evaluation of young children with suspicious injuries. I consider whether there are frailties inherent to such evaluations that might undermine the objective judgments of expert paediatricians and thus affect the reliability of the opinions they provide to the courts.

I discuss a variety of potential problems that expert paediatricians should be aware of and take in to account in formulating their expert opinions. These include the risk from various unconscious biases and the emotive nature of such cases, weaknesses in the scientific evidence base on which such judgments must be based, and the lack of an agreed common language by which experts can explain the certainty, or otherwise, of their conclusions.

I also discuss the particular challenges faced by the courts in ensuring that, in complex cases such as alleged “shaken babies”, their decisions are firmly based on reliable expert opinion evidence.
Acknowledgements

Pursuing a PhD thesis in the afterglow of retirement from a professional career is a little unusual, and setting aside the personal demands, such an endeavour presents particular challenges for those tasked with guiding and monitoring an individual much more attuned to being a supervisor than a supervisee. All those involved with this work could not have been more helpful or supportive.

I would like to thank Professor Sheila McLean, Professor Emerita of Law and Ethics in Medicine, University of Glasgow, for her encouragement and advice when I first raised the possibility of pursuing this project. She warned me that a PhD “was a huge step up from a Masters” – and she was right! Professor James Chalmers, then Senior Lecturer in Law at the University of Edinburgh, was very encouraging in response to my initial enquiry and subsequent PhD application. I thank him for that, and for his excellent supervision during my first year at Edinburgh Law School, before he moved to the Regius Professorship at Glasgow University.

For the last six years I have been supported and guided in this project by Professor Burkhard Schafer and Professor Graeme Laurie. I have benefitted greatly from their combined wisdom, breadth of knowledge and guidance, in particular the skill they quietly applied in steering me in the right direction when my focus drifted towards “interesting” side issues. Their differing backgrounds and approaches combined extremely well to provide the blend of supervision I needed, and I cannot thank them enough for their help in bringing this project to a conclusion. Any insightful analysis is likely to be down to them; any errors are mine.

Many individuals have also provided me with practical advice or other support at various times. However, I would particularly like in to thank the following:

Dr Geoff Debelle, Officer for Child Protection, RCPCH, for information on trends in NAHI convictions in England, and in relation the experience, or otherwise, of child protection paediatricians with pre-trial expert meetings in the English criminal prosecution system.

Professor Gary Edmond, School of Law, The University of New South Wales, Sydney, for providing pre-prints of his papers on expert evidence.

Mr Mike Hally, producer, SquareDog Radio, for supplying me with a digital copy of the Radio 4 broadcast “Peer Review in the Dock”, presented by Mark Whitaker in 2008.
Mr Jason Harper, Content Strategy Manager, Charles Seale-Hayne Library, University of Plymouth, for his time and expertise in resolving coding errors in my Refworks reference manager software application.

Dr Patrick Kelly, Director, Te Puaruruhau (Child Protection, Shaken Baby Prevention and Family Violence Intervention), Starship Children's Hospital, Auckland, for providing information on secular trends in infant abusive head injury prosecutions in New Zealand.

Ms Gill Moreton, Cognitive Behavioural Therapist, Rivers Centre for Traumatic Stress, NHS Lothian, for advice on the Post Traumatic Stress Disorder literature in relation to child protection workers.

Dr Catherine Skellern, Consultant Paediatrician, Child Protection and Forensic Medicine, Lady Cilento Children's Hospital, Brisbane, Australia, for providing pre-publication copy of her paper on minimising bias in the forensic evaluation of suspicious paediatric injury.

Dr Amanda Stephens, Emergency Department, Westmead Hospital, New South Wales, Australia for providing me with personal electronic copies of selected chapters from her 2011 PhD thesis deposited at the University of Sydney: “Legal Outcomes in Non-accidental Head Injury ('Shaken Baby Syndrome') Cases: Inevitable Inconsistencies?”

I acknowledge the generous permission of Lord Woolman and the anonymous agreement of a Scottish Appeal Court judge in allowing me access to Lord Woolman's confidential report in relation to Ms Kimberley Hainey’s appeal against her murder conviction (Chapter 2).

Finally, I wish to thank Ago (Dr Agomoni Ganguli-Mitra, Co-Director, Mason Institute), and the disparate attendees at the Mason Institute writing retreats for their fellowship, knowledge, gossip and humour during our boot camp sessions. I will almost miss those long days.
Dedication

As anyone who has embarked on a PhD knows, pursuing such a goal demands many things, including a measure of self-centred focus and a dogged determination to see such a marathon through to the end. However, one of the most important requirements is a loyal friend to accompany you on the journey, to sympathise and counter self-doubt and boredom, to spot procrastination, and to claim to see the light at the end of the tunnel, even at a stage when you both know any perceived glimmer reflects hope rather than reality.

This thesis is dedicated with unbounded gratitude and love to my loyal friend on this venture, my wife Marie, who agreed from the start that even at my stage in life “this was a really worthwhile thing to do”, and who in every conceivable way has encouraged and supported me, despite the many restrictions this selfish pursuit has had on her own plans for some fun after retirement.

Marie, I promise to make it up to you.
**Selective Glossary of Terms and Abbreviations**

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>AFSP</td>
<td>Association of Forensic Science Providers</td>
</tr>
<tr>
<td>AHT</td>
<td>Abusive head trauma</td>
</tr>
<tr>
<td>ALTE</td>
<td>Apparent Life Threatening Event. This is usually defined as an episode that is frightening to the observer and that is characterized by some combination of apnoea (see below), colour change, marked change in muscle tone (usually limpness), choking, or gagging.</td>
</tr>
<tr>
<td>Apnoea</td>
<td>Cessation of breathing - transient or more prolonged – can arise from a number of reasons in young babies, including infection, low blood sugar, and trauma.</td>
</tr>
<tr>
<td>CACD</td>
<td>English Court of Appeal (Criminal Division)</td>
</tr>
<tr>
<td>CCRC</td>
<td>Criminal Cases Review Commission</td>
</tr>
<tr>
<td>CEE</td>
<td>Concurrent expert evidence</td>
</tr>
<tr>
<td>CI</td>
<td>Confidence interval</td>
</tr>
<tr>
<td>CLA</td>
<td>Case linkage analysis: “a process that aims to identify crimes that are likely to have been committed by the same suspect because of the behavioural similarity evident across the crimes”(^2)</td>
</tr>
<tr>
<td>CONSORT</td>
<td>Consolidated Standards of Reporting Trials initiative - a standardised statement that those reporting RCTs (see below).</td>
</tr>
<tr>
<td>COPFS</td>
<td>Crown Office and Procurator Fiscal Service</td>
</tr>
<tr>
<td>CPCC</td>
<td>Child protection case conference</td>
</tr>
<tr>
<td>CPG</td>
<td>Clinical practice guideline</td>
</tr>
<tr>
<td>CT</td>
<td>Computerised tomography</td>
</tr>
<tr>
<td>Dura</td>
<td>The outermost, toughest, and most fibrous of the three membranes covering the brain and the spinal cord</td>
</tr>
<tr>
<td>EBM</td>
<td>Evidence-based medicine</td>
</tr>
<tr>
<td>ELC</td>
<td>Law Commission (England and Wales)</td>
</tr>
<tr>
<td>Encephalopathy</td>
<td>Damage to the brain affecting function</td>
</tr>
<tr>
<td>ENT</td>
<td>Ear, Nose, and Throat specialist</td>
</tr>
</tbody>
</table>

---

1 These terms and definitions have been derived or modified from a number of relevant professional sources.

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>FAI</td>
<td>Fatal Accident Inquiry</td>
</tr>
<tr>
<td>GMC</td>
<td>General Medical Council</td>
</tr>
<tr>
<td>GOR</td>
<td>Gastro-oesophageal reflux is a very common and usually benign feature in babies in which stomach contents are regurgitated into the throat, often associated with minor choking.</td>
</tr>
<tr>
<td>GP</td>
<td>General Medical Practitioner</td>
</tr>
<tr>
<td>GRADE</td>
<td>Grading of Recommendations Assessment, Development and Evaluation</td>
</tr>
<tr>
<td>HB</td>
<td>Heuristics and biases</td>
</tr>
<tr>
<td>HCJ</td>
<td>The Scottish High Court of Judiciary</td>
</tr>
<tr>
<td>HV</td>
<td>Health Visitor</td>
</tr>
<tr>
<td>Hypoxia</td>
<td>A pathological condition in which the body as a whole (generalised hypoxia) or a region of the body (tissue hypoxia) is deprived of adequate oxygen supply.</td>
</tr>
<tr>
<td>Hypoxic ischemic encephalopathy</td>
<td>Damage to cells in the central nervous system (the brain and spinal cord) from inadequate oxygen.</td>
</tr>
<tr>
<td>ICI</td>
<td>Intracranial injury</td>
</tr>
<tr>
<td>Intracranial</td>
<td>Within the cranium, the bony skull that houses and protects the brain.</td>
</tr>
<tr>
<td>Intradural</td>
<td>Part of the dura</td>
</tr>
<tr>
<td>JDM</td>
<td>Judgment and decision-making</td>
</tr>
<tr>
<td>Laryngomalacia</td>
<td>Here congenitally soft, immature cartilage of the upper larynx (voice box) collapses inward during inhalation, and is the most common cause of noisy breathing in infancy. It is usually benign and resolves as infancy progresses.</td>
</tr>
<tr>
<td>LR</td>
<td>Likelihood ratio</td>
</tr>
<tr>
<td>LTDNA</td>
<td>Low Template DNA</td>
</tr>
<tr>
<td>Maltreatment</td>
<td>All types of maltreatment – physical abuse, sexual abuse, emotional abuse and all forms of neglect – perpetrated by a parent or carer towards a child.</td>
</tr>
<tr>
<td>MRI</td>
<td>Magnetic resonance imaging.</td>
</tr>
<tr>
<td>NAHI</td>
<td>Non-accidental head injury; the preferred alternative to SBS.</td>
</tr>
<tr>
<td>NAHT</td>
<td>Non-abusive head trauma</td>
</tr>
<tr>
<td>NAI</td>
<td>Non-accidental injury</td>
</tr>
<tr>
<td>NDM</td>
<td>Naturalistic Decision Making</td>
</tr>
<tr>
<td>------------</td>
<td>------------------------------</td>
</tr>
<tr>
<td>Neonatal</td>
<td>Newly or recently born baby</td>
</tr>
<tr>
<td>NICE</td>
<td>National Institute for Health and Care Excellence.</td>
</tr>
<tr>
<td>OR</td>
<td>Odds ratio</td>
</tr>
<tr>
<td>Perinatal</td>
<td>Pertaining to the period immediately before and after birth. The perinatal period is defined in diverse ways. Depending on the definition, it starts at the 20th to 28th week of gestation and ends 1 to 4 weeks after birth.</td>
</tr>
<tr>
<td>Physical abuse</td>
<td>At least one act of severe violence from a parent or carer towards a child. Severe violence is usually defined as a kick, a bite, a scald/burn, ‘beating up’, hitting with an object, shaking a young child, or threatening to use a weapon.</td>
</tr>
<tr>
<td>PICU</td>
<td>Paediatric Intensive Care Unit</td>
</tr>
<tr>
<td>PPV</td>
<td>Positive predictive value</td>
</tr>
<tr>
<td>RAD</td>
<td>The Reflex Anal Dilatation test, wrongly purported to be a secure sign of sexual abuse.</td>
</tr>
<tr>
<td>RCPCH</td>
<td>Royal College of Paediatrics and Child Health.</td>
</tr>
<tr>
<td>RCT</td>
<td>Randomised controlled trial.</td>
</tr>
<tr>
<td>RH/s</td>
<td>Retinal haemorrhage/Retinal haemorrhages - bleeding into the linings of the eyes.</td>
</tr>
</tbody>
</table>
| Safeguarding | The term safeguarding (Child Protection in Scotland) is focussed on:  
  - protecting children from maltreatment;  
  - preventing impairment of children’s health or development;  
  - ensuring that children are growing up in circumstances consistent with the provision of safe and effective care;  
  - taking action to enable all children to have the best outcomes. |
<table>
<thead>
<tr>
<th>Acronym</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>SBS</td>
<td>Shaken Baby Syndrome. The term used to describe the constellation of injuries resulting from violent shaking of an infant by an adult or adolescent. Now superseded by the alternative acronym NAHI (see above).</td>
</tr>
<tr>
<td>SC</td>
<td>Supreme Court</td>
</tr>
<tr>
<td>SCR</td>
<td>Serious case review</td>
</tr>
<tr>
<td>SCRA</td>
<td>Scottish Children’s Reporter Administration</td>
</tr>
<tr>
<td>SDH</td>
<td>Subdural haemorrhage. Bleeding beneath the dural membrane of the brain.</td>
</tr>
<tr>
<td>SIDS</td>
<td>Sudden Infant Death Syndrome. Also known as cot death. It is the sudden unexpected death of an apparently well infant, for which there is no explanation.</td>
</tr>
<tr>
<td>SIGN</td>
<td>Scottish Intercollegiate Guidelines Network</td>
</tr>
<tr>
<td>STROBE</td>
<td>Strengthening the Reporting of Observational Studies in Epidemiology checklist, used to expose potential bias in diagnostic and observational clinical studies.</td>
</tr>
<tr>
<td>Substantiated reports</td>
<td>An American term for a report of maltreatment that has been verified by the child protection agency in accordance with state law. If an investigator finds that abuse or neglect occurred then the report is substantiated or founded.</td>
</tr>
<tr>
<td>SUDI</td>
<td>Sudden Unexpected Death in Infancy</td>
</tr>
<tr>
<td>ToM</td>
<td>Theory of Mind</td>
</tr>
<tr>
<td>Triad</td>
<td>The 'Triad' of intracranial injuries (retinal haemorrhages, subdural haemorrhages and encephalopathy) often found in victims of NAHI.</td>
</tr>
<tr>
<td>UKCTG</td>
<td>UK Clinical Trials Gateway.</td>
</tr>
<tr>
<td>Unified Hypothesis</td>
<td>The hypothesis put forward by Geddes (2004), challenging the mainstream interpretation of the triad in NAHI. It suggests that brain hypoxia, infection or raised intracranial pressure can cause not only encephalopathy, but also subdural haemorrhage and by implication retinal haemorrhage.</td>
</tr>
<tr>
<td>WCPSRG</td>
<td>Welsh Child Protection Systematic Review Group.</td>
</tr>
</tbody>
</table>
Table of Contents

Selective Glossary of Terms and Abbreviations .................................................. viii

Table of Cases.......................................................................................................... xvi

Table of Statutes and other legislative instruments............................................. xxi

Table of Figures...................................................................................................... xxii

Table of Tables....................................................................................................... xxii

Chapter 1 - Introduction ........................................................................................... 1
  The key focus of this thesis .....................................................................................2
  The special nature of paediatric forensic practice .............................................. 4
  What do I mean by paediatric forensic expert evidence? ................................... 5
  The broader ecology of child protection and forensic clinical assessment ....... 13
  Problems with expert evidence generally ..........................................................21
  Wrongful convictions and flawed paediatric evidence ...................................... 27
  The key question that I address in this thesis ..................................................... 35
  Thesis structure .................................................................................................. 36

Chapter 2: Anatomy of a miscarriage of justice: the wrongful conviction of
Kimberley Hainey .................................................................................................... 39
  Confidentiality statement .................................................................................... 39
  Introduction ......................................................................................................... 40
  Section 1 – Declan’s death and his mother’s conviction for murder .................. 41
  Arrest and indictment .......................................................................................... 46
  Section 2 – the trial .............................................................................................. 47
  The forensic experts, their evidence, and the competing trial narratives .......... 47
  The expert evidence ............................................................................................. 48
  Prosecution and defence narratives – contrasting stories to fit the available facts ... 50
  Section 3 – The disputed expert evidence .......................................................... 54
  Declan’s bones and claims of evidence of neglect ............................................ 55
  Challenges to the Dundee experts’ opinions .................................................... 57
  Significant key events of the trial process ......................................................... 60
  The verdict ........................................................................................................... 62
  The appeal ........................................................................................................... 62
  Section 4 – Case commentary and signposts to next chapters ......................... 66
Chapter 5 - Medicine-based evidence in alleged abuse: a secure epistemic basis for expert opinion? ................................................................................................................................................................. 159
  Introduction .................................................................................................................................................................................................................................................. 159
  Section 1- The birth of evidence-based medicine .......................................................................................................................... 161
  Section 2 - A brief evolutionary history of the evidence base for child protection ............. 169
  Interim conclusion ................................................................................................................................................................................................................... 175
  Section 3 - A secure foundation for forensic expert opinion? ...................................................... 176
  The Cardiff ranking system – a valid epistemic warrant for abusive injury? ....................... 181
  Confession evidence as a validation of clinical features of abusive injury? ......................... 186
  Conclusion ................................................................................................................................................................................................................................. 190

Chapter 6: Semantics and ambiguities in expressing expert opinions .................. 193
  Introduction ................................................................................................................................................................................................................................. 193
  Section 1: Quantifying expressions of certainty in expert opinions ........................................... 195
  Judicial responses to Bayesian reasoning and LR framework evidence ............................... 204
  Section 2 - “Doctor – when you wrote that, what did you mean?” .......................................... 209
  Semantics and expressions of certainty in paediatric opinion evidence ............................... 211
  The challenge of developing an appropriate taxonomy of forensic certainty .......................... 220
  Section 3 - A Bayesian approach to forensic child abuse assessment? ................................. 224
  Conclusion ................................................................................................................................................................................................................................. 231

Chapter 7 – “Shaken baby” cases – a diagnosis on trial ........................................ 233
  Introduction ................................................................................................................................................................................................................................. 233
  Section 1 - “Shaken babies” – sound science or an unproven hypothesis? ............................. 239
  Section 2 - Expert evidence in criminal trials of SBS – are paediatricians and their institutions the problem rather than the solution? .............................................. 246
  Section 3 – Admissibility of expert evidence - a brief overview ........................................ 253
  General principles – the common law admissibility test ..................................................... 254
  Evolving approaches to admissibility of expert evidence in England and Wales ............ 260
  Controversy, expert polemic, and obscurantism in SBS – the courts’ responses ............... 263
  Section 4 - From problem description to problem solution ............................................... 270
  The truth, the whole truth…and SBS ......................................................................................... 271
  Some on-going challenges for the courts .............................................................................. 273
  Conclusion ................................................................................................................................................................................................................................. 278
Chapter 8 – Conclusion: some inconvenient truths, and thoughts about future directions

Introduction .................................................................................................................. 281
What was the research problem and how did I address it? ......................................... 282
Key themes specific to the paediatric forensic practitioner .............................................. 284
Relevance and implications of this study ........................................................................ 286
Implications for practice ................................................................................................. 286
Strengths and limitations of the research ...................................................................... 290
Recommendations for further research .......................................................................... 290
Closing comments ......................................................................................................... 292

Bibliography ................................................................................................................. 293
<table>
<thead>
<tr>
<th>Table of Cases</th>
<th>UK Supreme Court</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jones v Kaney [2011] UKSC 13</td>
<td>.................................................................................................................. 19</td>
</tr>
<tr>
<td>Kennedy v Cordia (Services) LLP [2016] UKSC 6, [2016] 1 WLR 597 (SC)</td>
<td>.......... 70, 255</td>
</tr>
<tr>
<td>England &amp; Wales</td>
<td></td>
</tr>
<tr>
<td>Deloitte LLP v HMRC Commissioners [2016] UKFTT 0479 (TC)</td>
<td>................................................................. 255</td>
</tr>
<tr>
<td>Lillie &amp; Reed v Newcastle City Council [2002] EWHC 1600 (QB)</td>
<td>................................................. 249</td>
</tr>
<tr>
<td>Milton Keynes BC v Nulty [2013] EWCA Civ 15</td>
<td>.................................................................................. 204</td>
</tr>
<tr>
<td>Pearce v Ove Arup Partnership Ltd &amp; Ors [2001] EWHC Ch 455</td>
<td>................................................................. 24</td>
</tr>
<tr>
<td>R v (B) (The boyfriend of Baby Peter's mother) (C) (Baby Peter's mother) and Jason Owen [2009] EWCA Crim 2259</td>
<td>................................................................. 17</td>
</tr>
<tr>
<td>R v Adams (No 2) [1998] 1 Cr App R 377</td>
<td>.................................................................................. 205</td>
</tr>
<tr>
<td>R v Al Alas and Wray [2012] EWHC 865 (Fam)</td>
<td>.................................................................................. 28</td>
</tr>
<tr>
<td>R v Anthony (Donna) [2005] EWCA Crim 952</td>
<td>................................................................. 27, 68</td>
</tr>
<tr>
<td>R v Arshad [2012] EWCA Crim 18</td>
<td>.................................................................................. 268</td>
</tr>
<tr>
<td>R v Barry George [2007] EWCA Crim 2722; (2007) 151 SJLB</td>
<td>................................................................. 204</td>
</tr>
</tbody>
</table>
R v T [2010] EWCA 2439 .................................................................206
R v Weller (2010) EWCA Crim 1085 ...........................................221
Re AB (Child Abuse: Expert Witnesses) [1995] 1 FLR 181 ................266
Re B (Non-Accidental injury) [2002] 2 FLR 1133 ................................219
Re L (a child) [2013] EWHC 4646 (Fam) ........................................221
Re O-M (children) (expert evidence: non-accidental injuries) [2009] EWCA Civ 1405 .......7
Re X (Non-Accidental Injury: Expert Evidence) [2001] 2 FLR 90, [2001] EWHC Fam 6 ...70
Squier v GMC [2016] EWHC 2739 ................................................272
Squier v GMC [2016] EWHC 2739 (Admin), [ 2016] ACD 139 ........74, 76, 78, 272

Scotland

Davie v Edinburgh Magistrates 1953 SC 34; 1953 SLT 54 ..................24, 256
Hainey v HMA [2013] HCJAC 47 ..................................................27, 40, 267
Jones v HMA [2016] HCJAC 65 ..................................................255
Liehne v HMA [2011] HCJAC 51 ..................................................passim
Main v McAndrew Wormald Ltd 1988 SLT 141 ..........................74
McTear v Imperial Tobacco 2005 2 SC 1 ...........................................70, 257
Q v P [2016] Fam LR 54 ................................ ................................ ................................ ............... 255
Wilson v HMA 2009 JC 336 ................................ ................................ ................................ ........ 70, 76, 256
Younas v HMA [2014] HCJAC 114 ................................ ................................ ................................. 128

Australia

Folbigg v The Queen [2005] HCATrans 657 ................................ ................................ ...................... 27, 69
Inquest into the death of Azaria Chantel Loren Chamberlain [2012] NTMC 020 ......... 27
R v Bonython (1984) 38 SASR 45 .................................................................................. 255
R v Matthey [2007] VSC 398 .................................................................................. 27
R v Phillips [1999] NSWSC 1175 .................................................................................. 27

Canada

R v Marquardt 2011 ONCA 281 .................................................................................. 27
R v Sherret-Robinson 2009 ONCA 886 ........................................................................ 27

United States

Addington v. Texas, 441 U.S. 418 (1979) ................................ ................................ .................. 214
Daubert v Merrell Dow Pharmaceuticals 509 U.S. 579 (1993) ................................ . 25, 184
Holbrook v Lykes Bros Steamship Co Inc. 80 F3d 777 (3d Cir 1996) ...................... 215

State v Edmunds 2008 WI App 33, 746 NW2d 590 ................................................................. 27

United States v Mornan 413 F.3d 372 (3d Cir. 2005) .......................................................... 214

United States v Two Elk. 536 F3d 890 (8th Cir 2008) .......................................................... 214
Table of Statutes and other legislative instruments

Scotland
Act of Sederunt (Sheriff Court Rules) (Miscellaneous Amendments) (No 2) 2013
Children and Young Persons (Scotland) Act 1937
Criminal Procedure (Scotland) Act 1995

England and Wales
Civil Procedure Rules (Part 35)
Criminal Procedural Rules 2015
Criminal Practice Directions 2015 ([2015] EWCA Crim 1567)
Family Procedure Rules 2010 (updated January 2017)

New Zealand
Code of conduct for expert witnesses, New Zealand High Court Rules 2016
Table of Figures

Figure 1: A paediatrician’s multiple interactions in a typical child protection case................10
Figure 2: Declan Hainey around his first birthday..........................................................43
Figure 3: Bedroom image ...............................................................................................44
Figure 4: Living room......................................................................................................45
Figure 5: Likely drug paraphernalia ..............................................................................45
Figure 6: How (expert) doctors think............................................................................101
Figure 7: Dror's taxonomy of sources of bias that may contaminate forensic judgments....107

Table of Tables

Table 1: Common heuristics that lead to decision-making errors.................................92
Table 2: Ranking system to assign cases to abused or not abused within included studies.181
Table 3: Standards for numerical and equivalent verbal expressions of likelihood ratios. ..202
Chapter 1 - Introduction

For over 25 years as a consultant paediatrician, one of my formal clinical responsibilities involved forensic paediatric assessments in cases of suspected child abuse or neglect as part of a specialist child protection service. Such cases usually involved multi-disciplinary working with other statutory agencies – principally police and social work services – and ranged from the relatively minor to the most serious, including inflicted life-changing injuries and child murders. In this role, I provided countless forensic medical reports, attended child protection case conferences, and spent many hours as a professional witness in the criminal and family courts in Scotland.

In addition to this role as an NHS paediatrician, I also acted on many occasions as an independent expert witness for both the prosecution and defence in criminal and civil actions involving children in courts across the UK, and in Scotland provided expert evidence to Fatal Accident Inquiries involving children. Thus, I have been able to observe and experience the interaction of paediatric clinical forensic science and the law at close quarters.

While I have seen much to admire on both sides of that interface, I have also encountered situations that left me frustrated or concerned, from lawyers who seemed poorly prepared and who, from my perspective, did not ask me the “right question” to allow what I considered important evidence to be adduced, to colleagues who were selective and dogmatic in their clinical interpretation of injuries, and unwilling to concede the possibility of alternative explanations. I have had to counsel and support colleagues traumatised by the experience of the adversarial system, and seen others angry that an accused had, in their eyes, “got away with it”.

It is that personal experience and perspective that I bring to this thesis.
The key focus of this thesis

The key question that I consider in this thesis is whether there may be particular frailties intrinsic to paediatric forensic practice that may undermine the reliability of paediatric expert opinion evidence in cases of alleged child murder or physical abuse.

In doing so I make no claim that there is, as yet, much direct empirical proof of the specific frailties I profess to expose. However, from the perspective of an experienced paediatric forensic practitioner, I argue that a number of issues particular to the paediatric setting may indeed potentially undermine such opinion evidence and thus may partly explain the risk of flawed expert evidence and associated wrongful convictions or family court judgment in such cases.

While negative professional perceptions of the current regulatory and legal framework that provides governance of the paediatric profession are a possible obstacle for such a self-reflective analysis, these are concerns that the UK paediatric child protection community barely acknowledge. And despite paediatric experts being associated with a number of high-profile miscarriages of justice, there is a notable lack of specific scholarship on this issue.

Nor does it seem to be a particular domain of expert knowledge that the courts and legal community generally recognise as being particularly problematic from an opinion evidence perspective. Of course, the potential conflict between scientific understanding of truth and legal understanding of truth has been the subject of prior research.\(^1\) This type of study is necessary to develop the law of evidence on the one hand and the regulation of forensic professions on the other. The thesis will argue that paediatric forensic evidence is of a different nature from laboratory-based forensic science and presents particular difficulties when considered against the conception of truth in law.

\(^1\) Susan Haack 'Of Truth, in Science and in Law' (2008a) 73 (3) Brook Law Rev 985.
From a methodological point of view, the approach I follow does not fall neatly within a particular category associated with the traditional taxonomies of legal research. I build my arguments on indirect evidence, drawing on research from other relevant expert forensic contexts or other areas of scholarship that I then apply to the paediatric forensic setting. Thus, the work borrows from empirical studies focussed on disparate areas of forensic science, from clinical paediatrics, and from the theoretical constructs of neuroscience and behavioural psychology. From this blend of the harder and softer sciences, and through the interpretative lens of my personal experience, I construct my claim that there are indeed intrinsic frailties in the foundations of paediatric expert evidence; I then set these against the normative expectations of the legal settings that such evidence is meant to serve.

This, in essence, is the original contribution this thesis offers in an area of scholarship that is largely unexplored, and which, in addressing both the clinical and legal dimensions of paediatric expert forensic practice, provides a novel and important cross-cultural perspective equally relevant to paediatric experts and to legal practitioners in the criminal and family courts.

This introductory chapter provides a general background to the issues I later address. First, I highlight a number of practical and professional issues particularly relevant to paediatric forensic practice, and explain the broader societal and multidisciplinary context in which child protection work is set, before describing the clinical process involved in “diagnosing” a child’s injury as abusive.

I then provide a short overview of established problems with expert evidence generally, before considering the specific issue of wrongful convictions linked to flawed paediatric evidence and the arguably disproportionate impact these child cases have had on the provision of expert evidence more generally.
Finally, having set the scene, I end the chapter by describing the structure of the remainder of this thesis, identifying the key themes of the other chapters in which I develop my argument that particular frailties do indeed threaten the reliability of paediatric opinion evidence.

So, given that the question this thesis pursues is founded on issues peculiar to forensic practice within a profession paediatric setting, let me first set the scene with an overview of this particular paediatric medical domain.

**The special nature of paediatric forensic practice**

One of the core arguments I develop in this thesis is that paediatric forensic practice is fundamentally different from more “conventional” forensic science disciplines. For example, most forensic scientists work within a relatively exclusive “forensic” environment – the scientists have no other role, and while many areas of forensic science represent “niche” disciplines, the rules that must be followed and the standards that must be met may be considered both generic and relatively well established.

In contrast, all paediatricians who work in child protection services will have a background in the therapeutic care of children. They vary in the proportion of their professional time that they commit to a child protection role – ranging from a relatively modest to almost full time professional commitment. The majority continue to have other clinical responsibilities.

Unlike conventional forensic scientists, in moving into the forensic domain, paediatricians must change their professional perspective in fundamental ways – in particular away from a caring and therapeutic role, with its associated professional values and norms. Instead they must switch into a forensic mode, where routine assumptions – for example, that a child’s parents are telling the truth – must be explicitly discarded, and where normal clinical practice
routines do not apply, and a variety of mandatory laboratory or other investigations must be initiated, often without clinical justification.\(^2\)

I will argue later that it is in this major transition from the clinical therapeutic function to the forensic role that difficulties for paediatricians may arise. However, I will also claim that potential problems with paediatric expert evidence may also arise from other sources.

**What do I mean by paediatric forensic expert evidence?**

Before going further, let me explain briefly what I mean by paediatric forensic expert evidence, a term I use frequently throughout this thesis. Here, I use the term “forensic” simply to mean that the clinical evaluation and judgments made, together with any written and / or oral evidence provided by a paediatrician (or any other relevant clinician), is for legal purposes.

And of course, as I discuss later (Chapter 7), when the paediatrician, or any other expert, moves into the forensic domain they acquire specific responsibilities and must work within a new set of rules.

In the context of suspected inflicted injury, the output of a paediatric forensic assessment involves the provision of two types of evidence – factual evidence and opinion evidence. The factual evidence encompasses the history provided by the parents (and sometimes the child), physical examination findings, medical treatment, and the results of laboratory, imaging or other appropriate investigations.\(^3\)

Founded on the paediatrician’s training, knowledge and experience, inferences drawn from these facts provide the basis for expert opinion evidence, generally focussed on the likely

---

\(^2\) For example, the mandatory expectation to arrange a radiological skeletal survey of a young child in cases of suspected abuse to screen them for covert injuries (*The radiological investigation of suspected physical abuse in children* (The Royal College of Radiologists 2017)), or the mandatory referral to appropriate statutory agencies if abuse is suspected.

\(^3\) In the context of a direct paediatric assessment of a child, this may additionally involve the collection of other forensic samples for laboratory or other testing. Examples include DNA samples from alleged human bites or microbiological samples to test for sexually transmitted infections in alleged sexual abuse.
causation of suspicious injuries. From an epistemological perspective, the core issue here is to
distinguish natural (illness or accidental injury) from unnatural causes for the clinical
findings. As I discuss later, this often involves plausibility judgments and opinions on
whether the history provided by a parent is consistent with the physical examination findings
and other relevant clinical facts of a case.4

Let me also clarify another area of potential confusion. In the broader context of paediatric
forensic expert opinion evidence, I recognise a distinction between the roles of the paediatric
forensic expert, usually working as part of an NHS clinical child protection service, who
usually provides initial face-to-face clinical assessment, and that of an expert brought in later
by the prosecution or defence to provide an independent (secondary) opinion. The clinical
expert will often not only deliver or co-ordinate direct clinical care for the child in question
but will also provide expert advice to police and social work investigations, and both factual
and opinion evidence to the courts.

The involvement of an independent paediatric expert is somewhat different. Thus, for
example, such an expert rarely meets and discusses a case with the parents of a child;5 nor
are they involved in joint discussions about a case with relevant specialists from other
clinical services, commonly radiology. Rather, their opinions depend on inferences derived
from secondary (hearsay) information extracted from written clinical records and reports/
statements provided by various statutory agencies and other witnesses, albeit usually highly
detailed. Thus, in theory at least, they should be less affected by any potential difficulties

4 For a useful overview of the various interactions of paediatricians with the law, see Sandra L. J.
Johnson, 'Legal and Forensic Medicine in Paediatrics' in Roy G. Beran (ed), Legal and Forensic
Medicine (Springer-Verlag 2013) 79.
5 T. J. David 'Avoidable pitfalls when writing medical reports for court proceedings in cases of
that might arise from the often blurred and potentially conflicting overlap of both therapeutic and forensic roles.  

However, while it may be argued that this level of detachment may protect an independent expert’s objectivity, I suggest that at best such protection is relatively modest. All independent experts must, by current standards, be currently or recently engaged in active clinical child protection work and be involved in peer review of their own and others’ cases. In the relatively small professional world of safeguarding and child protection, I suggest that these experts are still subject to many of the threats to objectivity that this thesis later considers, and so the issues I am discussing are relevant to both. Thus, unless otherwise stated, my use of the term “paediatric expert” in this thesis encompasses both roles.

Further, since such independent experts usually have no direct contact with an accused parent or other factual witness, they must depend on the recorded histories of an incident and the interpretations of another professional’s plausibility judgment of a carer’s explanation, and photographs and descriptions of injuries rather than direct clinical examination. These are thus not “clean sheet” assessments, but dependent on others’ perspectives.

Let me turn to describe briefly the practical processes involved in a paediatric forensic clinical assessment.

---

6 David M. Foreman and Juliet Williams 'Medical law and protection of children' (2008) 337 BMJ; S. A. Greenberg and D. W. Shuman 'Irreconcilable conflict between therapeutic and forensic roles' (1997) 28 (1) Prof Psychol Res Pr 50. While there is some concern that such a dual role is problematic (see, for example Mary Pillai 'Forensic examination of suspected child victims of sexual abuse in the UK: a personal view' (2005) 12 (2) J Clin Forensic Med 57; P. G. Richards and others 'Shaken baby syndrome' (2006) 91 (3) Arch Dis Child 205), in the context of child protection work, such a dual role is recognised and accepted by the courts: Re O-M (children) (expert evidence: non-accidental injuries) [2009] EWCA Civ 1405.

A variety of recommendations have been made to ensure that paediatric expert witnesses are up-to-date in their knowledge and clinical activity so that they are in a position to provide valid opinion evidence. See, for example, in relation to the English courts, Part 25 of the Family Procedure Rules 2010 (updated January 2017); available from: https://www.justice.gov.uk/courts/procedure-rules/family/practice-directions/practice-direction-25b-the-duties-of-an-expert,-the-experts-report-and-arrangements-for-an-expert-to-attend-court (accessed 7 May 2018).
The “diagnosis” of child abuse – moving from suspicion to substantiation

The key responsibility of the specialist paediatrician when faced with an allegation that a child has suffered abuse is to collate and evaluate all relevant information, clinically evaluate the child, and make a diagnosis, and thus contribute to the information that, as a whole, will substantiate or refute that allegation. Specific medical issues usually considered in child protection cases, as noted by David, are:

- Precise delineation of injuries such as bruises or fractures
- An attempt to identify the likely age of an injury
- An attempt to indicate the likely type and degree of force resulting in an injury
- Consideration of the type of mechanism which could have caused the injury
- Differentiation between natural disease and pathology resulting from abuse
- The differentiation between accidental and non-accidental injury
- The identification of neglect
- The delineation of the likely symptoms exhibited by the child in a period prior to presentation to the health or welfare services.

---

8 Here I am focusing on the role of the expert child protection paediatrician, in England and Wales often called the “named” paediatrician responsible for child protection within a hospital or other service setting. Across the UK, specialist safeguarding/child protection professionals provide expertise and have specific roles and responsibilities in safeguarding children. In England, Northern Ireland, and Wales, named and designated professionals perform this function, the “designated” doctors having a significant strategic role. In Scotland, where the term “safeguarding” is not used, lead clinicians in child protection fulfil specialist roles, often supported by Nurse Consultants and Child Protection Advisers. See Royal College of Paediatrics and Child Health, ‘Intercollegiate document - Safeguarding Children and Young people: roles and competences for health care staff’ (2010).

9 David, n. 5 above.

10 It is important to note that this is often conflated with such dichotomous terms “intentional”/“unintentional”, “accidental”. As David points out (n. 5 above) and I discuss later, attributing implied motivation and intent is outwith the expertise of the paediatrician, and such terms are pejorative and potentially misleading.

11 That may be used by the prosecuting authorities to delineate a time-window when an injury occurred, or to suggest a neglectful delay in seeking medical attention.
The clinical evaluation process that is involved is essentially similar to the assessment of any ill or injured child. Since for expert clinicians, such a diagnostic process often relies on intuitive pattern recognition (see later, Chapter 3), this immediately raises questions about whether this is an ideal process to apply in a forensic context. However, the context of a forensic child protection evaluation also brings with it elements that are significantly different from that applied in more conventional clinical care.

Observing that the supporting evidence base in clinical child protection work is limited (of which see later, Chapter 5), Hilary Cass, past President of the RCPCH recently noted that diagnosis in suspected child abuse:

…is dependent on the trigger of an initial index of suspicion, supported by a mix of corroborative detail, experience, triangulation, monitoring, review and above all the highest level of multiagency communication and cooperation.  

By triangulation she meant the need to integrate the history provided by the parents and other professionals, the expert’s clinical examination findings, and the subsequent deliberations of a multi-agency case discussion.

A significant number of actors contribute to the overall process of forensic evaluation, so such triangulation must take account of the information and perspectives arising from many relevant sources.

---

Figure 1: A paediatrician’s multiple interactions in a typical child protection case.

As is evident from Figure 1, the multi-agency and multi-professional environment in which paediatricians provide forensic opinion evidence is highly complex, with potential areas of territorial difficulty involving the overlapping of different and somewhat conflicting objectives of various statutory agencies, each shaped by constantly evolving legislation, policy and guidance.

For example, despite this close working arrangement, the police and social work in particular have different statutory responsibilities and perspectives that may have a direct impact on how a case progresses. Thus, from a social work standpoint, an admission of inflicted injury may be seen as a positive event that may justify keeping a family together with supervision

and support. In contrast, such an admission to the police almost always results in a decision to prosecute.  

As the quote from Dr Cass above emphasises, one of the core elements of multi-agency child protection investigation is information sharing. Thus, the paediatrician, in formulating an opinion, is expected to take account of information from a variety of external sources – GP, Health Visitor (HV), social worker etc. While background health and medical information may be relevant, any non-clinical information from such sources should not influence the paediatrician’s forensic interpretation of an injury. However, such mandatory information sharing brings up an important issue that I address in detail in Chapter 3 – the potential risk that clinically irrelevant contextual information may unconsciously influence the expert’s interpretation of the origins of a child’s injury.

It is notable that Dr Cass explicitly suggests that the “diagnosis” of abuse is highly dependent on non-clinical features particular to each case, and that the “medical” elements, sometimes involving input from a number of clinical disciplines, are only one aspect of the process. In this she indicates that substantiation of abuse is a multi-agency process. This immediately raises questions about the basis of the conclusions the expert paediatrician reaches in such cases. Here I respectfully suggest that she is confusing the paediatrician’s role as part of a safeguarding (child protection) multiagency risk assessment process, focussed on a child’s future welfare, with the narrower focus of the clinical forensic interpretation of alleged inflicted injury.

This apparent blurring of roles provides an example of the potential frailties underlying paediatric expert evidence I consider later in detail.

---

15 Cobley, n. 13 above.
Figure 1 also illustrates a separate important feature of the forensic evaluation of a suspicious paediatric injury – that a variety of specialists may provide opinions on specific clinical issues depending on the nature of the case. Such multi-disciplinary clinical inputs raise issues in relation to defining the boundaries of personal expertise and may pose difficulties for both clinicians and the courts in determining who is the most appropriate expert to speak to a particular disputed fact, and which opinion carries the most weight. This issue of the boundaries of expertise is a recurrent concern in relation to wrongful convictions involving paediatric expert evidence, and is a particular focus of a detailed illustrative case review I provide in Chapter 2.

And of course, the paediatrician’s opinion may eventually be put before a number of fact-finding tribunals with different roles and standards of proof. Thus, almost all cases of suspected child abuse result in a case conference, where the focus of decision-making is welfare and the child’s “best interests”. Here the burden of proof is undefined but a conclusion that a child is to be placed on the Child Protection Register must be justified by a finding that the child is at risk of serious harm.

Running in parallel, there may be a criminal investigation, and subsequent trial with a beyond reasonable doubt/ “sure” standard of proof, while if the facts of the case conference or other matters are disputed and the case is referred to the English Family Court (or a Sheriff court proof hearing in Scotland), a balance of probability standard applies.

---

17 For example, in a typical suspected “shaken baby” case (see later, Chapter 7), treatment and assessments from an intensive care specialist, neurologist, ophthalmologist, haematologist, metabolic specialist, radiologist, and often neurosurgeon would be routine.

18 In Scotland, any relevant statutory agency may request that a CPCC be convened if there is concern that a child is at risk of significant harm, although they are convened and chaired by social work services.

19 See: National Guidance for Child Protection in Scotland (The Scottish Government 2014). I also acknowledge, for completeness, the additional role in Scotland of the Scottish Children’s Reporter Administration (SCRA) and the associated Hearings System. However, it would be very unusual for paediatricians to attend and provide evidence at SCRA lay hearings, but much more common to attend “Grounds for Referral” hearings before a sheriff if the facts were disputed.

20 In the English courts, “sure” is now the preferred term: see, for example, R v Majid (Abdul) [2009] EWCA Crim 2563, paras. [11] – [15].
Notwithstanding the variable standards of proof that apply, the outcomes of any of these disparate tribunals are promoted as a key indicator confirming that the evidence base underlying such forensic clinical interpretations is clinically valid. In Chapter 5 I argue that such a claim is fundamentally flawed.

Separately, in presenting the conclusions of their forensic evaluation of a case to such disparate tribunals with different standards of proof, it is self-evident that how the expert expresses the certainty with which they have concluded or rejected the possibility that a suspicious injury is abusive may well be crucial to the ultimate decisions the relevant fact-finders make. Yet the reality is that there is no agreed semantic hierarchy that paediatric experts may adopt in expressing such certainties in their conclusions. This is an issue I consider in detail in Chapter 6.

So, as I hope I have already begun to reveal, beneath the surface of paediatric expert forensic evaluation may lurk a number of potential threats to the reliability of the expert evidence provided in such cases.

Having provided a short overview of the process of paediatric forensic evaluation of a child with a suspicious injury, it is also important to appreciate the wider context in which such evaluations are set.

**The broader ecology of child protection and forensic clinical assessment**

In relation to my general claim that there are particular potential systemic frailties that may undermine paediatric forensic opinion evidence in alleged child abuse, it is also important to appreciate the wider setting in which such forensic work is provided.

*The demands of policy and expectation*

Child protection, in its broadest sense, is a challenging and highly emotive issue in our progressively risk-averse society. Concern about the possibility that children have been or
are at risk of physical or sexual abuse or neglect is at an all-time high.\textsuperscript{21} Safeguarding
vulnerable young people poses major challenges for our society and the families that live
within it, for policy makers, and for the various professional agencies with a statutory duty to
respond appropriately.\textsuperscript{22}

In order to understand the context of paediatric expert evidence in the prosecution of cases of
physical abuse, it is appropriate that I outline, albeit very briefly, the burden of child
maltreatment and the professional responses to it.

\textit{The epidemiology of child maltreatment}

Each year in the UK and other English-speaking common law jurisdictions, between 1.5%
and 5\% of all children are reported to child-protection services; 1\% of the child population
has maltreatment substantiated.\textsuperscript{23} This figure seriously underestimates the true burden of
child abuse. It is estimated that each year between 4–16\% of children are physically abused,
while 10\% are neglected or psychologically abused.\textsuperscript{24} Further, up to a quarter of young
adults report some form of sexual abuse during childhood.\textsuperscript{25} Thus there is no doubt that the
problem of child maltreatment is very significant.

\textsuperscript{21} For example, referring specifically to trends in England, see Home Office Safeguarding Unit,
\textit{Reporting and acting on child abuse and neglect: Government consultation} (HM Government 2016)
1, at [9].
\textsuperscript{22} See, for example, current guidance for England: \textit{Child sexual exploitation: definition and guide for
practitioners} (Department for Education DFE-00056-2017 1.
\textsuperscript{23} Ruth Gilbert and others 'Recognising and responding to child maltreatment' (2009b) 373 (9658) The
Lancet 167. In general, four types of child abuse are widely recognised: physical, sexual,
psychological (emotional), and neglect. They are not mutually exclusive and commonly children
experience more than one form.
\textsuperscript{24} R. Gilbert and others 'Burden and consequences of child maltreatment in high-income countries'
(2009a) 373 (9657) Lancet 68.
\textsuperscript{25} Lorraine Radford, Susanna Corral, Christine Bradley, Helen Fisher, Claire Bassett, Nick Howat
Pressure to get it right - the search for certainty

We also live in an age of outrage, and a huge 24-hour media industry shapes our society’s views and influences policy, none more so that when dealing with cases of child abuse and child murder. When professionals involved in such cases appear to have made misjudgments, there is an instant, overwhelming and invariably critical media response, irrespective of any valid justification, the reaction often supplemented by highly intimidating internet trolling.

When vulnerable children are murdered, rather than focus blame on those who were ultimately convicted, almost invariably such events are retrospectively characterised as a combination of individual professional and system failures. Accusations that previous lessons have not been learnt are usually part of the critique, for these cases are perceived to follow a long line of similar “failures” and subsequent furore, dating back at least 30 years.

One effect of societal anger and politicians’ unease has been the evolution of linked child protection legislation with progressively prescriptive procedural frameworks for relevant

---

26 Mark Manson. ‘Living in the Age of Outrage’ Observer (20 June 2016).
28 Sandra Laville. ‘Paediatricians shy away from dealing with abuse as backlash from parents grows’ The Guardian (7 June 2004) Children.
30 Nigel Parton ‘From Maria Colwell to Victoria Climbié: reflections on public inquiries into child abuse a generation apart’ (2004) 13 (2) Child Abuse Rev 80. Prominent events that provoked major inquiries and major system change were the deaths of Maria Colwell (1973), Kennedy McFarlane (2000), Victoria Climbié (2000) and Peter Connolly (2007), and the Cleveland, Orkney and Newcastle Upon Tyne inquiries following allegations of multiple cases of sexual abuse in the family home and in day care, among others. For summaries of the reviews and responses to some of the major incidents see, for example: “Messages from reviews and inquiries”, Scottish Government: http://www.scotland.gov.uk/Publications/2003/05/17127/21840 (accessed 1 August 2018).
statutory agencies across the various nations of the UK – a so-called “protocolisation”
approach.\(^{31}\)

Even in cases that do not involve child death, when things go wrong mandatory serious case
reviews are conducted,\(^{32}\) often relatively judgmental in nature, and the actions and decisions
of those from the various agencies involved – social work, health, education, and police – are
robustly scrutinised.\(^{33}\) Despite evidence that some of these events were unpredictable, or that
professional errors were defensible,\(^{34}\) recriminations abound.

**Shifting thresholds and swinging pendulums**

Concern about unmet need, on-going risk and a drive for a more effective multi-agency
response is clearly evident at national policy level, as the Scottish Government’s review of
child protection - “It's everyone's job to make sure I'm alright” - makes clear.\(^{35}\)

Thus there is an explicit expectation that both the public and relevant professionals should
recognise and respond to children who might be abused or neglected.\(^{36}\) And perhaps

---

Risk Regulation Regimes* (Oxford University Press Oxford 2004) 165. Such an increasingly
prescriptive approach may defend the relevant agency, while focusing any blame when things go
2008) 130.

32 Sharon Vincent and Alison Petch, *Audit and Analysis of Significant Case Reviews* (The Scottish
Government 2012).

33 Dept. for Education, *Working Together to Safeguard Children: A guide to inter-agency working to
safeguard and promote the welfare of children* (HM Government 2013) 68; Available at:
https://www.education.gov.uk/aboutdfe/statutory/g00213160/working-together-to-safeguard-children

34 Eileen Munro 'Avoidable and unavoidable mistakes in child protection work' (1996) 26 (6) Br J Soc
Work 793.

35 The Scottish Government, "It's everyone's job to make sure I'm alright"- Report of the Child
Protection Audit and Review (The Stationery Office 2002).

36 Matt Forde. 'Time to step up and speak out for vulnerable children' *Glasgow Herald* (29 January
2014) Agenda 15; this exhortation by Matt Forde, NSPCC Scotland for the public to report without
delay any concerns about possible abuse or neglect of children typifies the current mood, and has
resulted in a significant increase in referrals to NSPCC and other relevant agencies: Anon. 'NSPCC
reports 58% rise in calls referred to social services' *The Scotsman*, 13 Oct 2016.
inevitably, one effect of society’s anxiety and political disquiet has been an increasing demand for aggressive intervention by the various responsible agencies.

Not surprisingly, the central policy changes and generally prescriptive processes developed in response by the relevant agencies have had a major impact on child protection activity, with significant rises in referrals and assessments because of child protection concerns. In the five fiscal years following the ‘Baby P’ case, care applications in England rose by 70%, and at the time of writing, the English family justice and children’s social care systems are facing a crisis in the face of a relentless rise in care proceedings.

It seems clear that policy changes and a lowering of the threshold for intervention underlie that increase, the rise in the number of children entering the care system reflecting a new, less tolerant, perspective.

37 The NSPCC provide secular trend statistics on a wide variety of national child protection activity trends. For example, the numbers of children on child protection register in England are available at: https://www.nspcc.org.uk/services-and-resources/research-and-resources/ which links through to statistics for all four nations (accessed 14 March 2018).


39 R v (B) (The boyfriend of Baby Peter’s mother) (C) (Baby Peter’s mother) and Jason Owen [2009] EWCA Crim 2259. See also:


The increased pressure on doctors to recognise and report possible maltreatment

The pressure to be suspicious and “think dirty” is not just a societal one. There is mounting pressure on doctors in all clinical settings to recognise child abuse and not “miss” a case, with prescriptive national guidelines underpinning that pressure. Current General Medical Council (GMC) guidance makes it clear that all doctors must raise concerns with appropriate agencies if they feel that a child is at risk of, or experiencing, abuse or neglect, even if their concern is merely based “on a hunch”. There appears to be little room for professional discretion and no place for error despite the often highly complex situations involved.

The impact of these such pressures on paediatricians

The impact on paediatric professional practice of high profile media storms in the aftermath of claims of missed opportunities to prevent abusive child deaths or serious injury cannot be overstated. Cases such as that of “Baby P” (Peter Connolly) exemplify that trend, and

---

45 Rebecca Smith. 'Doctors should report child abuse suspicions "on a hunch"' The Telegraph (10 July 2012).
46 While there are no specific mandatory regulations in the UK requiring professionals to report suspicions about possible child abuse to the authorities, professional guidelines across a number of professional regulatory authorities such as the GMC emphasise that this is expected. Mandatory reporting of concern about possible child abuse by a variety of professional groups (medical, school and child care, law enforcement, and a variety of other public service employees) was established in all US states by 1967; Children’s Bureau, Mandatory Reporters of Child Abuse and Neglect (U.S. Department of Health and Human Services 2012) 1, and now in most Australian states: Ben Mathews and Deborah Scott. 'Mandatory reporting of child abuse and neglect' (2013); <http://www.afis.gov.au/cfca/pubw/factsheets/a141787/index.html> accessed 18 May 2018.
47 Sven Günther. 'The recent history of child protection in Scotland and the impact of intra-familiar child abuse inquiries such as Orkney on today's child protection work' (Masters, Glasgow Caledonian University 2000).
48 Clare Dyer 'Great Ormond Street and Baby P: was there a cover up?' (2011) 343 BMJ 286.
illustrate starkly the expectations involved and the pressures professionals from all agencies face in managing risk and avoiding harm to vulnerable children.

Public vilification of paediatricians from an aggressive media for, on the one hand, “missing abuse”, or on the other for labelling a child’s injuries as abusive, only to have that opinion rejected in court, together with an associated aggressive professional regulatory response, has been a source of great professional insecurity and stress. Not surprisingly there has been a severe impact on the number of paediatricians prepared to give evidence in the English Family Courts and, in the face of a fivefold rise in complaints about paediatricians’ actions in child protection cases, recruitment to paediatric child protection posts has been very difficult.

Given this highly pressurised policy environment in which clinical child protection practice sits, for those paediatricians who take on such responsibilities and have to make complex clinical judgments in relation to possible abuse, it would be hardly surprising if, in the face of such background pressures, they were to adopt a defensive approach, with a lower threshold not only to consider the possibility of abuse but also to conclude that an injury was abusive. Conversely, a perceived risk of professional censure might induce a more cautious interpretation.

49 See, for example, the experiences of Professor David Southall: Clare Dyer ‘David Southall: anatomy of a wrecked career’ (2012) 344 BMJ 18.
50 Cass (2014), n. 12 above.
51 Bearing Good Witness: Proposals for reforming the delivery of medical expert evidence in family law cases - A report by the Chief Medical Officer (Department of Health 2006). Arguably this reluctance has been reinforced by the Supreme Court ruling in Jones v Kaney [2011] UKSC 13 that expert witnesses are no longer immune from being sued for the evidence they give in court.
52 Zosia Kmietowicz ‘Complaints against doctors in child protection work have increased fivefold’ (2004) 328 (7440) BMJ 601.
Currently there is little empirical evidence to indicate any secular trend, one way or the other, in paediatricians’ diagnostic thresholds for substantiating child abuse, although significant variations in individual diagnostic thresholds make such studies difficult. However, in Scotland in the spring of 2009 there was a notable rise in hospital admissions of infants with injuries attributed to physical abuse, a change that coincided with the death of Peter Connolly and the subsequent media frenzy around it. This may indicate that paediatricians’ clinical diagnostic thresholds are flexible and respond to external influences.

And of course, these external pressures are just one source of potential threat to diagnostic objectivity. Other factors might alter such diagnostic thresholds. As I noted earlier, all paediatric forensic experts have a professional background imbued with the values of caring and child advocacy that are intrinsic to their general professional role, which includes general obligations to identify children’s health and emotional needs and risks of harm, and to prioritise their needs. As I discuss later (Chapter 4), such empathic engagement with child patients may be in tension with the objectivity demanded by the forensic role. Might paediatricians, working in an explicitly child-centred environment with such deep-rooted professional values, unconsciously respond in alleged abuse cases by modifying their diagnostic thresholds or the way they express the expert opinions that follow? Before focussing further on such potential frailties with paediatric expert evidence, it is appropriate that I first consider the wider “expert evidence” context.

Problems with expert evidence generally

Whether in a criminal or civil context, legal disputes involving expert evidence are superficially simple: a fact-finder must settle a dispute in one or other party’s favour by deciding between competing narratives; unlike other witnesses, when the issues in dispute involve matters outside common knowledge, expert witnesses may exceptionally provide the decision maker with an opinion on the technical merits of each narrative. However, since the courts began to seek the help of experts to advise them on technical or other specialist matters over two centuries ago, there have been difficulties.  

For example, faced with conflicting expert opinions, how are decision makers meant to judge the merits of each view, given they lack the requisite expertise in the first place? And, of course, there have been many other problems: experts usurping the decision-makers role, especially if they have acquired “celebrity” status; providing partial evidence in a “hired gun” role, influenced by “adversarial bias”; or simply offering scientifically flawed or, rarely, deliberately deceptive evidence.


57 Perhaps typified by the first Home Office pathologist Sir Bernard Spilsbury, who became the face of forensic medicine to the British public and whose opinion was accepted without question for most of his career. It is likely that a number of innocent individuals were hanged as a result. See Ian Burney and Neil Pemberton 'Bruised Witness: Bernard Spilsbury and the Performance of Early Twentieth-Century English Forensic Pathology' (2011) 55 (1) Med Hist 41.


59 See, for example, Joseph Sanders 'Science, Law and the Expert Witness' (2009) 72 LCP 63.

60 See, for example, Paul Giannelli 'Scientific Fraud' (2010) 46 (6) Crim Law Bull 1313.
And there are legal system problems as well as those linked to individual experts, notably a frequent mismatch in resources between prosecution and defence in relation to access to and funding of experts.\footnote{Keith A. Findley 'Innocents at Risk: Adversary Imbalance, Forensic Science, and the Search for Truth' (2008) 38 (No 893) Seton Hall L Rev 893, at 906-7.}

In broad terms, problems with expert evidence may be considered to fall into two areas: experts’ failure to comply with their expected responsibilities, and with the courts or other fact-finding tribunals failure to initially exclude, or later expose, unreliable expert testimony.

Henderson and Seymour have succinctly summarised the key, albeit overlapping, issues:\footnote{Emily Henderson and Fred Seymour, \textit{Expert Witnesses Under Examination in the New Zealand Criminal and Family Courts} (School of Psychology, University of Auckland 2013), 5-6.}

- Experts may present flawed opinion evidence to the courts, whether from deliberate dishonesty, incompetence and/or bias, or other reasons
- Fact-finders, whether juries or judges, are not competent to assess expert evidence and thus recognise this flawed testimony
- Irrespective of the quality of the evidence, juries in particular may respond to the authority or charisma of the expert, who may thus usurp their fact-finding function
- The traditional adversarial trial model and process is poorly equipped to handle expert evidence
- Experts themselves are concerned about how they, and their evidence, are used (and abused) by lawyers and the courts.

Given these difficulties, justice has not always been served, and innocent individuals and their families have suffered as a result. While much blame for wrongful convictions has lain with individual expert failures, previous calamities such as the “Birmingham Six”,\footnote{R v McIlkenny [1992] 2 All ER 417; (1991) 93 Cr App R 287.} the “Maguire Seven”,\footnote{R v Maguire (1992) 94 Cr App R 133 CA.} the case of Judith Ward,\footnote{65} and the associated government inquiries they
provoked also reveal that such miscarriages of justice may arise from systemic problems with expert evidence across the criminal justice process.  

Of course, much has changed in the last 25 years and, partly related to the exculpatory evidence provided by the emergence of DNA profiling, we have a greater understanding of the causes of wrongful convictions. Although the research base has often focused around analysis of individual or small groups of cases, we now know, for example, that eyewitness identification, informant testimony, and “false” confessions are consistent problematic areas.

On a broader front and more relevant to the issues this thesis considers, it has become increasingly clear that the evidential reliability of much traditional forensic science expertise is less certain than once thought – particularly ballistics, fibre, bite mark and hair comparison evidence, and even fingerprint evidence.

As a significant literature continues to expand in relation to problems with expert evidence generally across a wide variety of domains, new questions have emerged. For example,

---

67 I acknowledge that DNA exonerations principally apply to the factually innocent, and do not reflect broader issues in relation to miscarriages of justice; see Daniel S. Medwed (ed), Wrongful Convictions and the DNA Revolution: Twenty-Five Years of Freeing the Innocent (Cambridge University Press, 2017) 407.
70 See, generally, Gary Edmond 'The science of miscarriages of justice' (2014) 37 (1) UNSWLJ 376; Gary Edmond 'Advice for the courts? Sufficiently reliable assistance with forensic science and
setting aside concerns about the application of questionable scientific techniques, a major related issue that has emerged from more recent wrongful conviction scholarship is the recognition that a variety of cognitive biases may influence the objectivity of all the actors in the justice process, including forensic experts.  

Given these longstanding concerns, the common law in a number of jurisdictions provides explicit codes of conduct in relation to the role and responsibilities of expert witnesses, prominently focused on the expectation of the expert’s objectivity and impartiality, and their role in assisting the court rather than the party that engaged them. In England, for example, such guidance is supported by a number of Practice Directions that cover a wide variety of relevant matters such as the format and content of experts’ reports, often including mandatory boilerplate statements that the expert must include confirming their understanding of their responsibilities. Such codes of conduct are often supplemented by guidance from the expert’s relevant professional body.

---


Inevitably, it is appellate case law that tends to expose problems with expert evidence, reflecting not only individual failures, but also the shifting landscape of scientific and other knowledge, and the challenges the courts face in responding to the uncertain reliability of novel hypotheses or the waning of established dogma.

Thus we have seen, for example, the development in the US courts of the judicial gatekeeping role in scrutinising expert evidence linked to the so called “toxic tort” civil litigation that resulted in the Daubert trilogy and associated standards for admitting expert testimony, 75 and the response to uncertain reliability in the rejection of ear print evidence by the English court in R v Dallagher. 76

A key theme that has become explicit in all common law jurisdictions is that the courts demand that experts must only provide opinions that are within the bounds of their expertise. A failure to meet such expectations was recently claimed to have contributed to the wrongful conviction of solicitor Sally Clark for the murder of her infant sons. 77

In Clark (hereinafter to refer to her second appeal), the success of the appeal was primarily based on the exposure of “new” evidence, previously wrongly withheld by a pathologist, that possible infection might explain one infant’s death. However the Court also held that flawed statistical evidence provided by a renowned paediatrician Professor Sir Roy Meadow against the probability that Sally Clark’s two infant sons had died natural deaths from Sudden Infant Death Syndrome (SIDS) was likely to have influenced the jury to convict her of their murder, this ground having been rejected at her first appeal. 78

---

77 R v Clark [2003] EWCA Crim 1020, [2003] 2 FCR 447. On 9 November 1999 at Chester Crown Court (unreported), Sally Clark had been convicted of killing her children Christopher and Harry when aged 11 weeks and 8 weeks respectively.
78 R v Clark (Sally) 2000 WL 1421196.
Of course, such problems are not just about expert failings, for questions arise here about why the prosecution felt it appropriate to adduce such statistical evidence in the first place from a general paediatrician. Again, cases such as Clark expose the difficulties for lawyers and judges in deciding who is the most appropriate expert to speak to a particular issue, a key theme of this thesis that I consider again in Chapter 2.

But as I discuss later, in the context of paediatric child protection, the professional interactions and dependency intrinsic to such multidisciplinary work make simple professional labels somewhat irrelevant, and the courts recognise that in the context of certain particularly complex issues such as “shaken baby syndrome”, individual expert knowledge may extent much wider than conventional professional boundaries (Chapter 7).

It is not my intention here to provide any further overview of the problems with expert evidence in general. The scholarship is extensive and expanding, with a significant body of work of a generic nature in respect of the role and responsibility of the expert witness in civil and criminal trials, including recently those of paediatricians. Many of the problems posed by the use of such evidence are well recognised, notwithstanding it is clear that there are many issues to resolve.

---

79 The Ikarian Reefer, n. 72 above.
80 RCPCH/ Family Justice Council, Paediatricians as expert witnesses in the family courts in England and Wales: Standards, competencies and expectations <https://www.rcpch.ac.uk/sites/default/files/2018-08/Paediatricians%20as%20Expert%20Witnesses%20in%20the%20Family%20Courts.pdf> accessed 15 August 2018
Wrongful convictions and flawed paediatric evidence

While such cases are relatively uncommon and, as Clark above illustrates, the clinical forensic issues they raise are highly specific to a child/infant population, as I discuss below these recent paediatric cases have arguably had a significant and disproportionate impact on expert evidence jurisprudence more generally.

Clark was only one of a series of prominent wrongful convictions involving paediatric experts called as witnesses in cases of alleged serious inflicted infant and child harm in Anglo-American common law jurisdictions in recent years. Thus appellate courts have dealt with notable appeals against conviction for murder or manslaughter/culpable homicide of infant(s) and young children in Scotland (Liehne, Hainey), England and Wales (Clark, Cannings, Anthony, Kai-Whitewind, Harris, Rock, Cherry and Faulder, Henderson, Butler and Oyediran), Canada (Sherret-Robinson and Marquardt), and Australia (Phillips, Folbigg and Matthey). Similar cases have been seen in the US.

---

82 For an overview of cases involving flawed testimony by paediatric medical experts in the England and Wales, see K Campbell and C Walker, ‘Medical mistakes and miscarriages of justice: Perspectives on the experiences in England and Wales’ in Kent Roach (ed), Pediatric Forensic Pathology and the Justice System: Independent Research Studies vol. 2, Goudge Report (Gov. of Ontario 2007). There have been a series of high profile unsafe convictions involving unreliable expert forensic evidence related to a number of areas of expertise (see, for example: Mike Redmayne, Expert Evidence and Criminal Justice (1st edn, Oxford University Press Oxford 2001) 98-99; Bibi Sangha, Kent Roach and Robert Moles, Forensic investigations and miscarriages of justice: the rhetoric meets the reality (Irwin Law Toronto 2010) 447.


85 R v Marquardt 2011 ONCA 281; R v Sherret-Robinson 2009 ONCA 886.

86 Folbigg v The Queen [2005] HCATrans 657; R v Matthey [2007] VSC 398; R v Phillips [1999] NSWSC 1175. Failures of forensic evidence more generally were involved in the infamous “dingo baby” case. Lindy Chamberlain was jailed for her 9-week-old daughter Azaria’s murder in 1982 but released in 1986. The case involved a criminal trial, three inquests, a Royal Commission, a pardon (Chamberlain v The Queen (No.2) 1984] HCA 7, (1984) 153 CLR 521). This case was only recently finally resolved in 2012, over three decades after the event, when a third inquest accepted that a dingo had killed the baby: Inquest into the death of Azaria Chantel Loren [2012] NTMC 020.

87 For example, that of Audrey Edmunds (State v Edmunds 2008 WI App 33, 746 NW2d 590).
These cases prominently involved alleged murder by suffocation (where SIDS was a competing explanation), or so-called ‘shaken baby syndrome’ (SBS) to explain death or severe brain damage, although flawed paediatric expert evidence focused on a variety of other issues have also been associated with wrongful convictions. In all such cases, the outcomes depended almost exclusively on medical opinion evidence on the cause of a baby’s death or serious injury.

While the cases I have quoted above were all in the criminal courts, similar cases of equal complexity involving paediatric expert opinion evidence come even more frequently before the family courts, where the focus is child welfare rather than criminal prosecution, and, as I noted above, the standard of proof is lower. Here the implications of the judgments for all concerned are equally stark – where the potential gross injustice of a child being wrongly removed from its parents must be balanced against the risks of leaving a vulnerable child at home to be further abused. When the paediatric expert evidence on which such care cases turn is flawed, there are sometimes irreversible consequences of the utmost gravity for innocent parents.

---

89 Around 10,000 public law childcare cases are dealt with each year by the Family Courts in England and Wales. At least 25% involve expert evidence, typically from paediatricians or child psychiatrists at the ‘threshold’ stage, or adult psychiatrists involved in ‘disposal’ decisions: (Bearing Good Witness: Proposals for reforming the delivery of medical expert evidence in family law cases - A report by the Chief Medical Officer (Department of Health 2006), paras 2.10-2.11).
90 See, for example, R v Al Alas and Wray [2012] EWHC 865 (Fam), involving the death of a child with suspected “shaken baby” head injury, who turned out to have rickets, and where the initial care provided to the infant may have contributed to the diagnostic confusion.
91 Nadia Khomami. ‘Parents cleared of abuse told they may not be reunited with their child’ The Guardian (9 Oct 2015).
And of course, unreliable paediatric expert evidence may not only result in wrongful convictions or unjustified decisions in care proceeding, but also raise questions about the legitimacy of some acquittals. In some of the notable paediatric cases above, the reasons for the flawed convictions seem, at least superficially, relatively straightforward. Thus, as I noted earlier, in withholding evidence that might have allowed defence counsel to argue there was a possibility of an innocent explanation for an infant’s death, the pathologist Dr Williams in Clark demonstrated an apparent egregious failure to understand the legal process and the professional responsibilities of expert witnesses in relation to criminal trial evidence. Or, as seen in the false certitude provided by Professor Sir Roy Meadow in his interpretation of the likelihood of an unnatural explanation for clusters of recurrent SIDS cases in Clark and Cannings, an expert may adhere to scientifically invalid dogma, and/ or provide flawed evidence outwith the boundaries of their expertise. These examples of failings of expert witnesses noted above might be regarded as “conventional” and well-recognised shortcomings that are an explicit breach of experts’ established duties (see above).

However, the media storm, public vilification of “experts” and political tension that ensued in the aftermath of Clark and Cannings seemed to expose pre-existing concerns about expert

---

92 Perhaps best typified by the exculpatory “expert” evidence provided by Dundee biochemist Dr Colin Paterson in relation to so-called “temporary brittle bone disease” in many criminal and family courts in the UK and US before this postulated condition was discredited and he was censured by the courts and the General medical Council: (Owen Dyer ‘GMC strikes off proponent of temporary brittle bone disease’ (2004b) 328 BMJ 604; C. R. Paterson and E. A. Monk ‘Temporary brittle bone disease: relationship between clinical findings and judicial outcome’ (2011) 3 (3) Pediatr Rep e24).  
93 Clare Dyer ‘Pathologist in Sally Clark case suspended from court work’ (2005) 530 BMJ 1347.  
94 Here Dr Williams arguably demonstrated the fundamental difference between the clinical and forensic settings, his judgment that a positive bacterial culture was clinically irrelevant and a likely “contaminant”, and often unreported in routine clinical laboratory work, contrasting with the strict expectation in a forensic setting that all information that forms the basis of opinions must be explicit.  
95 See Clark, n. 77, and Cannings, n. 84 above.  
evidence more generally, and despite their modest numbers the effect of these cases has been far reaching.

**The impact of the infant homicide appeals**

Following the successful Cannings appeal, the then Attorney General, Lord Goldsmith, instructed a review of all convictions since 1994 for murder, manslaughter or infanticide of a child under two years old by a parent or carer, to ensure that there were no further cases of the type described in Cannings. While 28 cases out of the 297 reviewed raised concerns, only three involved recurrent SIDS like those of the Cannings’ children, and all were referred back to the mothers’ lawyers, resulting in Donna Anthony’s successful appeal.  

A further 97 cases raised a separate issue: these were SBS cases and had involved convictions in the face of established but now disputed medical opinion on the origins of brain injury and other putative features of SBS. In the event, four cases (hereinafter Harris) went to the English Court of Appeal (Criminal Division) (CACD) who allowed the appeal in whole or in part in three.

I discuss these appellate cases in detail within the specific issues around paediatric evidence and SBS later, in Chapter 7.

Following these successful criminal appeals, the then Children’s Minister Margaret Hodge ordered a review of civil cases in which children had been taken into local authority care primarily based on disputed expert medical evidence. In the event, there were only two appeals brought in relation to care orders made in the Family Division and both were dismissed.

---

97 Anthony, n. 84 above. Donna Anthony had her conviction for killing her two children quashed after spending seven years in prison, after the CCRC send her case back to the Court of Appeal in the light of the Cannings judgment.

98 Harris, n. 84 above.

99 Mrs Margaret Hodge, Minister for children: Expert Witnesses (Family Court Cases), Hansard 23 Feb 2004: Column 37.
The wider consequences of the flawed paediatric expert evidence cases

These wrongful convictions involving paediatric expert evidence had a much wider impact than the narrow area of paediatric forensic testimony. Since such convictions were almost exclusively based on opinion evidence, they reignited more widespread concerns about many aspects of expert witness evidence. Thus, they prompted a call for reform of the admissibility of expert evidence in criminal cases in England and Wales by the House of Commons’ Science and Technology Committee, culminating in a major review of expert evidence in criminal trials by the English Law Commission, where Clark, Cannings and the other paediatric cases are prominently cited as a justification of proposals for reform.

These paediatric cases particularly exposed the specific challenge for the courts when faced with conflicting expert evidence associated with evolving scientific theories or hypotheses, or with the possible falsification of established scientific dogma.

This first issue had been dramatically demonstrated in the 1980s following the sudden “diagnosis” of probable sexual abuse in 121 children in a small geographical area over a few months, and the subsequent Cleveland Child Abuse Inquiry. While the main focus of the inquiry was how the various statutory authorities – principally police, health and social work - had responded to this “epidemic” rather than the validity of interpretation of the clinical findings, a key professional issue was the apparent over-zealous reliance by paediatrician Dr Marietta Higgs and colleagues on the then novel (and today generally unsupported) clinical

---

The Children Act, already in development, was heavily influenced by the Cleveland scandal and its aftermath; in particular there was move from “state paternalism” to an explicit recognition of the importance of family life and a child’s right to maintain family bonds. Of course, how the courts reconcile the opinions of the zealots and the naysayers in relation to evolving (or regressing) scientific theories will always be challenging. As the rigid application of “Meadow’s law” or the eventual rejection of “temporary brittle bone disease” as an exculpatory explanation for infants with unexplained fractures show, such transitions take time and inevitably wrong judgments will (in retrospect) be made during such movements in scientific understanding.

While in some jurisdictions rules of evidence are in place to prevent or at least minimise the risk of such errors, when experts are allowed to present flawed testimony based on unreliable or evolving science, such events may also be considered to represent failures by the legal actors involved in managing such evidence. As I discuss later (chapter 7), currently we appear to be in another transition phase - focused on the reliability or otherwise of expert evidence in “shaken baby” cases – the courts faced with deciding between the highly polarised conflicting views of various experts on the cause underlying such clinical findings.

103 While reliance on the RAD test as diagnostic confirmation of sexual abuse was clearly invalid, the referrals for clinical examination were founded on other professional concerns. Later, a confidential report sent to the Department of Health by the then medical officer of health in the North East, Sir Liam Donaldson suggested that in up to 70% of the cases, the diagnoses of sexual abuse may have been correct.


106 O Dyer, 'GMC strikes off proponent of temporary brittle bone disease', vol 328 (2004b) 604; Richard Nicholl and others ‘Can we have a permanent end to ‘temporary brittle bone disease”?’ (2012) 97 (8) Arch Dis Child 762.
Paediatric expert evidence – important but relatively unexplored territory

As I noted earlier, there is a dearth of literature specifically centred on the paediatric expert witness in the criminal courts. The sparse published literature on paediatric expert evidence has primarily been dominated by comment on individual cases, or groups of cases, and concentrated on the failures of individual expert witnesses, or on the contested scientific reliability of forensic evidence in relation to a particular contested issue, such as SBS (see later, Chapter 7).

Similarly, responses to flawed expert evidence in the family courts, although such cases traditionally have had a much lower profile for reasons of confidentiality, have also tended to focus on individual expert failings.

While the Goudge Inquiry in Ontario, Canada, although highly focused on the impact of the personal failures of paediatric pathologist Charles Smith, took a wider view and explored the possibility of more general systemic weaknesses in paediatric forensic pathology that

---

111 See, for example, Jo Delahunty and Kate Purkiss 'What price justice? Experts or treating clinicians? LB Islington v Al Alas and Wray ' (2012) (July) Fam Law 832.
112 Goudge, n. 101 above.
may have contributed to the miscarriages of justice for which he was responsible, it did not address issues most relevant to paediatricians providing expert opinion evidence. A broader examination of issues particular to paediatric expert evidence has not taken place.

Summary

This introduction has provided a brief overview of the clinical context, processes and the complex professional and wider community setting in which paediatric experts provide forensic opinions in relation to suspicious injury or unexpected death in infants and young children. It shows that there are some significant differences in the provision of paediatric expert evidence compared to those of more conventional forensic science.

Despite an expanding scholarship on many aspects of expert evidence generally, and the small but significant number of high profile wrongful convictions involving paediatric experts, there is a minimal literature acknowledging the possibility that there may be issues of concern that are relatively specific to paediatric expert evidence.

This thesis argues that there are such specific factors intrinsic to the paediatric forensic setting that may impact on the decision-making and judgments of expert paediatricians, and which may influence the reliability of the opinion evidence they provide.

These are the issues that I address in the subsequent chapters of this thesis.

---

The key question that I address in this thesis

As I noted above, the key question I consider here is whether there are particular frailties underlying paediatric expert evidence that might undermine the reliability of such opinion evidence, and thus threaten the safety of criminal convictions or civil fact-finding dependent on such evidence.

Drawing on indirect scholarship from a number of fields, I argue that there are indeed significant and largely unrecognised systemic frailties specific to the provision of paediatric expert opinion evidence in alleged child abuse cases that may partly explain previous failings of the justice system.

Four key issues emerged from my review of the existing scholarship and deficiencies relevant to my key research question. I consider these in detail in the subsequent chapters of this thesis:

1. That the process of multiagency investigation, the nature of the clinical assessment process, and related interactions provide a largely unavoidable exposure to contextual and other cognitive biases and emotional contamination that may significantly undermine expert objectivity

2. That there are fundamental questions about the reliability of some of the clinical evidence base that should support the objectivity of experts’ conclusions and opinions.

3. Recognising that scientific truth is not the same as legal truth, and that doctors’ use of scientific evidence to make diagnoses differs from the way that the courts use evidence to make judgments, that ambiguities and lack of an agreed semantic hierarchy within which paediatricians may express their level of confidence in the conclusions and opinions that flow from their forensic assessments leave room for significant (mis)interpretation.
4. Of course, when wrongful criminal convictions or flawed family court judgments occur, such failures do not simply nor necessarily reflect poor paediatric expert testimony, but a wider failure of the legal process. Thus, I also consider whether the established safeguards provided by conventional rules of evidence in relation to admissibility and reliability of expert testimony are particularly suited to significant aspects of paediatric forensic evidence.

More broadly, a theme that runs through this thesis, and the various specific issues that I consider in detail, is that paediatric expert evidence is intrinsically different from other forms of expert evidence, and this special nature presents particular challenges for the experts, and for the courts.

**Thesis structure**

This thesis consists of eight chapters.

Here, in Chapter 1, I have described my own involvement with the issues this thesis considers and provided an overview of the particular contexts within which paediatric expert forensic assessments take place. I illustrated the paucity of specific scholarship on this topic and the background to the key question this thesis addresses - whether there are specific systemic problems in the provision of paediatric expert evidence to criminal and civil hearings involving alleged serious harm to babies and children that might explain the wrongful convictions (or flawed judgments) associated with such cases.

In Chapter 2, I provide a critical analysis of one of these cases - the wrongful conviction of Kimberley Hainey for the murder of her son Declan, where I acted as an expert witness. I use this Scottish case, not previously formally explored in the literature, as an exemplar to identify key themes that I address in subsequent chapters.
Thereafter, in the following two chapters, I focus on various aspects of human cognition and emotion. From different perspectives, both chapters explore potential biases relevant to paediatric child abuse work that may threaten expert objectivity.

In Chapter 3, I discuss the largely intuitive nature of most human reasoning, and relate this to the narrower area of clinical judgment and forensic opinion forming in the context of paediatric forensic assessment. I describe how intuition and heuristics are now recognised as fundamental to the medical diagnostic process, and contrast that with the different expectations of forensic expert assessment. I suggest that the implicit influence of paediatric professional culture and values, together with case-specific contextual features, may potentially undermine the objectivity of paediatric expert opinion evidence.

In Chapter 4, I extend my discussion of human cognition by exploring how our emotions affect how we think and form judgments. I argue that the intrinsically empathic nature of paediatric clinical practice and the highly emotive nature of cases involving the alleged murder or inflicted injury in babies and young children may threaten the paediatric expert’s objectivity in the forensic case assessment and the opinion evidence they provide.

In Chapter 5, I examine the available evidence base on which expert opinion evidence on alleged child abuse must be founded. Here I question the scientific and epistemological basis of some key areas of paediatric expert testimony, given that the “gold standard” claimed to validate clinical features of abusive injury is principally founded on putative perpetrator confessions, and on a circular argument founded on the outcomes of criminal trials, civil family court judgments and child protection hearings, rather than empirical science.

In Chapter 6, I consider the “are you ‘sure’ doctor?” question. In the absence of an agreed verbal scale by the child protection community, I discuss the ambiguities and particular semantic difficulties faced by the forensic paediatrician in expressing appropriate degrees of certainty in the conclusions they reach and how fact-finders may (mis)interpret such opinion evidence.
In Chapter 7, recognising that wrongful convictions also arguably represent a failure by the justice system, I use the example of alleged “shaken baby syndrome” (SBS) cases to demonstrate that in such complex cases all the various potential concerns about paediatric expert evidence that I identified in the previous chapters coalesce, and that they are manifest not only at an individual but at professional institutional level. Exposing the particular challenges the courts face in managing expert evidence in such cases, I argue that given evolving uncertainties about the origins on the triad of cranial injuries found in alleged infant shaking (see later), in Scotland such cases should no longer be prosecuted based on the presence of the triad alone.

Finally, in Chapter 8, I summarise the key conclusions of this thesis, discuss the implications for paediatric child protection practice, and suggest some solutions and areas for future research.

In the final analysis, my claim to an original contribution is that this thesis represents the first systematic examination and integrated account of potential frailties that may underlie core elements of paediatric expert forensic assessment and associated opinion evidence in relation to alleged serious child abuse. The work provides a novel insight into an important but largely unexplored area of forensic practice and the findings are highly relevant to both paediatric experts and legal practitioners.
Chapter 2: Anatomy of a miscarriage of justice: the wrongful conviction of Kimberley Hainey

Confidentiality statement

As an expert witness for the Crown in this case I had access to confidential documents such as Ms Hainey’s maternity record and primary care reports in order to provide opinion evidence on the various questions I was asked to address. In this chapter I have been careful to ensure that any personal information related to Ms Hainey that I discuss in this chapter is restricted to that available in the public domain, either revealed in the original criminal trial itself, the published appeal, or in the later Serious Case Review and Fatal Accident Inquiry related to the death of Declan Hainey.

Further, although I was supplied with an extensive confidential forensic photograph collection by Police Scotland in order to prepare my expert report, the photographs I use here to illustrate this case study were included in those released by the Crown Office to the media following judicial review and the opinion of Lord Woolman, on 12 January 2012.

In the preparation of this chapter, with the generous permission of Lord Woolman and the anonymous agreement of a Scottish Appeal Court judge, I was provided with Lord Woolman’s confidential report to the Appeal Court in relation to Ms Hainey’s appeal. This report was made available to me on the basis of my proposal that if any information therein were used in this chapter, such information would be redacted from any subsequent published thesis, or other publication, whether in digital or hardcopy format.

Thus, with the agreement of the University of Edinburgh School of Law Ethics Committee, and with the support of my examiners, while an unredacted copy was used for viva examination purposes, all relevant information related to Lord Woolman’s report, albeit sparse, is redacted in this chapter.

Introduction

“[T]he law holds, that it is better that ten guilty persons escape, than that one innocent suffers.” Lord William Blackstone.

As I noted in the preceding introduction, the core question that this thesis seeks to address is whether there are fundamental problems with the provision of paediatric expert evidence that might explain why prosecutions involving death or serious injury in young children are relatively prominent among wrongful conviction cases. It seems appropriate therefore that I begin with a detailed examination of such a case – the wrongful conviction of Kimberley Hainey in 2011 for the murder of her son Declan. I have chosen this case for a number of reasons.

First, because of my personal involvement in this case as a prosecution expert witness, I can offer an insight that may not be available to an external observer. Second, since this case has had little formal analysis elsewhere, and none from a paediatric forensic perspective, such an analysis will also contribute to the novel scholarship this thesis offers. Third, the case not only allows me to set the general scene but to identify some of the significant issues particularly relevant to paediatric expert evidence that I address in subsequent chapters.

My analysis of the case begins in Section 1 with a narrative describing the accepted facts of the case leading to the original conviction of Kimberley Hainey for the murder of her son Declan, following the discovery in March 2010 of Declan’s partly decomposed and mummified remains lying in his cot, some eight months after he was last seen alive.

---

118 I acknowledge that the terms “miscarriage of justice” and “wrongful [or unsafe] conviction” are not synonymous, that the former is a result of the latter, and that neither term necessarily implies factual innocence, although in many cases that will be the case. For a useful judicial interpretation by the Supreme Court of the meaning underlying these phrases, see R (Adams) v Secretary of State for Justice [2011] UKSC 18, [2012] 1 AC 48.
119 Hainey v HMA [2013] HCJAC 47.
In **Section 2**, I describe the initial forensic information and involvement of the various expert witnesses, before moving on in **Section 3** to provide an overview of the key evidential themes and prosecution and defence counsel narratives, with a focus on the most crucial disputed issues. In **Section 4**, I consider the case in the broader context of this thesis, and identify some key issues that I then explore in subsequent chapters.

**Section 1 – Declan’s death and his mother’s conviction for murder**

*Background and timeline to the discovery of Declan’s body*

Kimberley Hainey was born in 1974. Her early life seems relatively uncomplicated. She had barely known her father, and had little contact with him, but had a reasonably good relationship with her mother, albeit there were periods of intermittent disconnection. She also had close relationships with an uncle and an aunt. She had secure employment with an insurance company, a mortgage and owned small flat. By all accounts, she had a stable life.

In 1999, apparently following the death of her uncle, her life changed significantly. She became depressed, developed a drug habit and began drinking heavily. She lost her job, and with it her flat. As her addiction developed, her lifestyle became chaotic. She had no permanent home and “disappeared” for long periods, with spells of minimal family contact for up to a year. By 2005 she was regularly using heroin and cannabis. She had three periods of inpatient psychiatric care over this period, and during the final one formed a relationship with another psychiatric patient, David Gibson. That brief and ultimately acrimonious

---

121 The information in this brief section is derived the case papers supplied to me in order to prepare my expert report for the Crown, from Sheriff Ruth Anderson’s Determination following the Fatal Accident Inquiry (FAI) into the death of Declan Hainey, in particular, Sheriff Anderson’s Findings of Facts Section, at [1] – [8]. Available at: [http://www.scotland-judiciary.org.uk/10/1308/Fatal-Accident-Inquiry-into-the-death-of-Declan-Hainey](http://www.scotland-judiciary.org.uk/10/1308/Fatal-Accident-Inquiry-into-the-death-of-Declan-Hainey) (accessed 3 April 2018). These facts were also detailed in Hainey, n. 119 above, at [9].
relationship lasted some months, during which Declan was conceived in the summer of 2007.\(^{122}\)

During the early part of the pregnancy, Ms Hainey continued to smoke and abuse drugs; she was malnourished and unkempt. Although she did eventually engage with the local supervised methadone programme, her co-operation and engagement with various health and other support agencies was generally poor and inconsistent. Yet although significant concerns were identified in relation to Ms Hainey’s ability to care for her baby once born, a pre-birth case discussion concluded that there were ‘no significant child protection concerns at the moment’.\(^{123}\)

Declan Hainey was born uneventfully on 17 April 2008. Although he was rather small, he appeared perfectly healthy. Ms Hainey and her mother became reconciled immediately following Declan’s birth and Declan and Ms Hainey went to live in the new grandmother’s home. During the following months, Declan thrived and showed normal developmental progress.

In September 2008, Declan, now aged five months, and his mother left the grandmother’s home and moved to a new flat at 45 Bruce Road, Paisley. Shortly thereafter, Ms Hainey gradually reduced contact with her family and disconnected with the health, social and drug support services. She lied about her whereabouts at times, did not attend various planned meetings, and there was evidence of deterioration in her self-care likely linked to drinking and drug use.

---

\(^{122}\) David Gibson expressed no interest in the outcome of the pregnancy, and never met his son Declan at any stage.

\(^{123}\) FAI, n. 121 above, at [26]. Serious criticisms of the actions and co-ordination of care and information-sharing were major issues identified by an independent Serious Case Review commissioned by Renfrewshire Child Protection Committee commissioned Significant Case Review in 2010, and in the subsequent FAI determination by Sheriff Anderson in 2014. These are not the focus of this review and will not be further discussed.
However, there was occasional extended family contact so that on 17 April 2009, Kim Hainey and Declan went out for a meal with the rest of the family to celebrate Declan’s first birthday. That was the last time any member of his extended family was to see Declan alive. The photograph of Declan below was the last one to be taken. His likely age in that photograph would become a significant issue later, given subsequent events.

Figure 2: Declan Hainey around his first birthday.

Around the time of Declan’s first birthday, Ms Hainey formed a new relationship with a Robert Milton. Notably from then, Declan was seen in public less and less. Kim Hainey offered various explanations for his absences, most explanations later shown to be untrue. Further she would regularly stay overnight with Robert Milton at his home, sometimes for two consecutive nights. It seems clear that at these times Declan was being left entirely alone in the family flat in Paisley.
In the latter part of 2009 and early 2010, Ms Hainey repeatedly evaded efforts by the family health visitor to visit and assess Declan. Social workers were also thwarted in multiple attempts to see him. By March 2010 there was mounting professional concern about the situation.

On 30 March 2010, when he would have been 23 months old, the health visitor alerted his grandmother that social workers were about to report the family as “missing”. When relatives visited and gained entry to the family flat, they found Declan’s decomposed and partly mummified body. It was clear he had been dead for many months.

*The death scene*

![Figure 3: Bedroom image](image)

Declan was found lying on a bare plastic mattress in his cot, partly covered by a yellow towel and dressed in an unbuttoned babygro, with one bare leg exposed. He was not wearing a nappy. Close to his head lay a soiled nappy and a baby bottle containing curdled milk.

A cellophane wrapper from a cigarette paper lay between his legs.
The flat, previously redecorated for the family’s move there, and initially clean and kept in good condition, was now a squalid rubbish-scattered home, the baby’s playpen and other areas strewn with empty cider and vodka bottles and beer cans (Figure 4). Used milk bottles and discarded mouldering food cartons littered every surface. Likely heroin smoking paraphernalia were found (Figure 5).

Figure 4: Living room

Figure 5: Likely drug paraphernalia
It was evident that Kim Hainey was not living in the flat. When her estranged father, alerted to the situation, contacted her by telephone two days after Declan’s body was found, Kim Hainey still maintained her story that everything was fine, telling him that Declan was at his nursery before abruptly ending the conversation.  

**Arrest and indictment**

Kimberley Hainey was arrested and charged with Declan’s murder. It is relevant to note in particular the terms of the final charge in her indictment:

> “Between 1st September 2008 and 30th March 2010, both dates inclusive, the precise dates being to the prosecutor unknown, at 45 Bruce Road, Paisley and elsewhere, having sole custody, charge and care of Declan Hainey, your son, born 17th April 2008 then residing with you at said 45 Bruce Road, you KIMBERLEY MARY HAINEY did, on various occasions wilfully ill-treat and neglect said Declan Hainey, fail to provide him with adequate nourishment and fluids, leave him alone and unattended within said 45 Bruce Road for excessive periods of time and fail to seek to provide medical aid and care for him, and did thereby, and by other means to the prosecutor unknown, cause unnecessary suffering and injury to the health of said Declan Hainey whereby he died and you did murder him.”

The indictment in charge 1 also contained an alternative to the charge detailed above.

Essentially this alternative charge was of neglect. I have provided the specific detail of these indictments because they go to the heart of the forensic evidence that would become crucial to the initial murder conviction and a major issue in a later appeal. A number of other charges were laid on the indictment, but these are not relevant to my present analysis.

---

124 FAI, n. 121 above at [113].
125 The charge had originally included various words and phrases: the word "assault" before the words "wilfully ill-treat and neglect", the words "leave him in wet and soiled nappies for excessive periods of time" after the words "within said 45 Bruce Road for excessive periods of time", and "cause him to ingest diamorphine and amphetamine" before the words "and fail to seek to provide medical aid". The words in question were deleted by the Advocate Depute at the close of the Crown case.
126 Under the Children and Young Persons (Scotland) Act 1937, section 12(1).
127 As noted by the Appeal Court, there were four other charges originally on the indictment. The second was a charge of attempting to defeat the ends of justice. Kim Hainey was found guilty of that
Section 2 – the trial

The trial took place in late 2011 at the High Court of Justiciary in Glasgow.

Having set out the core facts of the case leading to Kimberley Hainey’s arrest, in this section I describe the initial forensic information, the involvement of the various expert witnesses, and the prosecution and defence narratives that were developed by the respective counsel at the trial.

The forensic experts, their evidence, and the competing trial narratives

Here I provide an overview of the key forensic evidence adduced and the case narratives developed by the prosecution and defence counsel from their respective perspectives. I do so as an “insider”, having had detailed access to all the prosecution evidence in order to provide a pre-trial expert report on some specific issues, pre-trial precognition meetings with both prosecution and defence counsel, and having appeared in Court over two days as an expert witness for the Crown.

Further, having anticipated some of the technical problems the conflicting expert evidence might raise, I also took the opportunity to personally observe from the public benches the closing arguments from both prosecution and defence counsel.128

Of course, in recounting this case it is important to acknowledge that such an analysis must inevitably reflect my personal perspective on the case, seen through the lens of my own experience of it, and restricted by the information I was and became privy to. Doubtless other relevant actors will have different perceptions coloured by their own roles and personal

charge. Two charges of offences involving contravention of the Misuse of Drugs Act 1971 were withdrawn from the indictment by the Crown. A further charge of theft was also withdrawn by the Crown. She was found guilty of a sixth charge involving her having failed to comply with a bail condition.

128 This was first time in a long career as an expert witness that I had re-attended a hearing or trial in this way. However, I recognised the controversial issues of expert evidence that were at play here, and that a better understanding may allow the case to become part of this thesis. I note that there was a little concern by counsel when I informed them of my plan, and I was asked to sit, as far as possible, outside the eye line of the jury members.
experiences. However I have made every effort to ensure that the facts of the case that I provide are accurate and verifiable.

Before dealing with the expert evidence, I note for completeness that a number of Crown witnesses provided evidence on a number of factual matters. In particular, various family members, health professionals, social workers and neighbours provided evidence in relation to Ms Hainey’s lifestyle, relationship with Declan, Declan’s various noted absences and Ms Hainey’s explanations of these, and other relevant issues. I will not discuss these further here, although much of this information was summarised in the later FAI.  

The expert evidence

A variety of experts provided opinion evidence, principally focused on the timing of Declan’s death, and whether or not any natural or unnatural explanation for the death could be inferred from the information available.

Timing of Declan’s death - forensic entomology

Dr John Manlove, forensic entomologist, opined that Declan’s body had been lying in his cot from or before the autumn of 2009, thus some seven months or so before he was found.  
Such timing correlated with the evidence from various lay witnesses, and it was particularly noted that Declan was not seen in public after the summer of 2009.

The cause of death - forensic pathology

Forensic pathologist Dr Julie McAdam and specialist paediatric pathologist Dr Clair Evans performed a joint forensic autopsy two days after Declan’s body was found.  

---

129 See FAI, n. 121 above.  
131 In the UK autopsies following sudden infant deaths follow the “Kennedy protocol”, and involve a series of a variety of laboratory and other investigations as well as physical examination of the remains; Baroness Helena Kennedy, ‘Sudden Unexpected Death in Infancy: a multi-agency protocol for care and investigation’ (The Royal College of Pathologists and The Royal College of Paediatrics and Child Health 2004) 73.
state of decomposition and partial mummification of his body indicated to them that death had occurred at least several months earlier. However, the condition of the body severely limited useful interpretation of the autopsy examination. Thus, although post mortem radiological examination did not reveal any bone injury, neither natural disease nor an unnatural cause of death such as inflicted injury could be determined with certainty. Nor could any useful opinion be provided in relation to the child’s nutritional state at the time of death. In the circumstances, the pathologists’ joint opinion was that Declan’s cause of death must therefore be regarded as “unascertained”.

At the trial, another two pathologists provided expert opinion evidence: Professor Neil Sebire, Great Ormond Street Hospital, London, and Dr David Manghan, a bone pathologist at the Royal Orthopaedic Hospital, Birmingham.

Professor Sebire provided advice on the terminology and classification of sudden deaths in infants and young children. He also commented on some of the antenatal and environmental risk factors that were associated with Sudden Infant Death Syndrome (SIDS), some of which were present in this case. In commentating particularly on the incidence of

132 Dr Julie McAdam/ Dr Claire Evans, Post Mortem Report PM2010/0546, 18 May 2010. Generally, the term “unascertained” is used when the pathologist identifies features in the clinical history or at autopsy that raise the suspicion of a possible unnatural death but which are insufficient to explain the death; Stephen J. Gould, Martin A. Weber and Neil J. Sebire 'Variation and uncertainties in the classification of sudden unexpected infant deaths among paediatric pathologists in the UK: findings of a National Delphi Study' (2010) 63 (9) J Clin Path 796; Martin A. Weber and Neil J. Sebire, 'Post-mortem Investigation of Sudden Unexpected Death in Infancy: Role of Autopsy in Classification of Death' in Elisabeth E. Turk (ed), Forensic Pathology Reviews Vol. 6 (Humana Press/ Springer 2011) 27.

133 SIDS refers to sudden unexplainable infant deaths and is only appropriate after a detailed autopsy and other investigations reveal no identifiable explanation for an infant death (M. Willinger, L. S. James and C. Catz 'Defining the sudden infant death syndrome (SIDS): deliberations of an expert panel convened by the National Institute of Child Health and Human Development' (1991) 11 (5) Pediatr Pathol 677). SIDS is grouped with other sudden explainable infant deaths under the broader heading Sudden Unexpected Death in Infancy (SUDI). None of the usual labels used in classifying such deaths – SIDS, SUDI, or Sudden Unexpected Death in Childhood (SUDIC) were appropriate in Declan’s case, given the limited information available from the autopsy due to decomposition.

134 Notably maternal smoking. Other relevant factors in this case are also noted risk factors identified in criminally suspicious child deaths, and include maternal health issues and drug and alcohol abuse (Michelle Ann McManus and others 'The Co-occurrence of Risk Factors for Intra-familial Child Homicides and Suspicious Child Deaths in England and Wales.' (2015) 1 (1) J Invest Child Deaths 60.
an unexplained but possibly “natural” death in a child of Declan’s age, he suggested that he would expect “about one boy a year in Scotland at the age of fifteen months to die due to unexplained causes”. Thus, such deaths are very rare but not unknown.

Before moving on to discuss the evidence provided by other experts, where the prime controversies lay, it is important to understand that evidence in the context of the prosecution and defence narratives. Thus, I turn first to summarise those.

**Prosecution and defence narratives – contrasting stories to fit the available facts**

*The Crown case*

The core of the Crown case was that Declan’s mother had murdered Declan by wilfully and recklessly failing to provide him with appropriate essential care, so that he had died as a result of her neglect. Prosecuting counsel Advocate Depute Andrew Stewart developed a case founded on a claim of “overwhelming circumstantial evidence – [a] jigsaw [that] made sense as a whole” – so that it was Ms Hainey’s “wicked recklessness” in her neglect of Declan, rather than an intention to kill, that provided justification for a valid murder conviction.

Acknowledging a “good start” after Declan’s birth, the availability of a supportive extended family, and her initial co-operation with third sector drug/alcohol support agencies, Ms Hainey’s relapse in alcohol and drug abuse, her related withdrawal of contact with her family, and lack of co-operation with routine health visitor and other agencies surveillance were indirect markers of her lack of appropriate care of Declan. The evidence of several witnesses had confirmed that Declan was only seen in public sporadically after his first birthday and showed that he was left for progressively long periods alone and unsupervised.

---

135 Hainey, n. 119 above, at [23].
136 This and subsequent quotations in this sub-section are from my contemporaneous notes of Mr Stewart’s closing speech to the jury.
while she was elsewhere drinking alcohol and using heroin and other drugs. It was clear that such repeated abandonments posed progressively serious risks to Declan’s life – and “[it] was clear from Dr Beattie’s expert evidence that in such circumstances there could be a downward spiral leading to Declan’s death if left [for] long enough”.

The AD argued that the defence suggestion of SIDS as the cause of death (see later) should be discounted. SIDS was usually much commoner in infants younger than Declan; after the first birthday, it was extremely rare. Even if there might be a single such death each year in Scotland involving a child of Declan’s age – in the context of all the other adverse circumstances, he asked: “was Declan that one boy?”

And if Declan’s death was “natural”, why did his mother not contact anyone? Instead she concealed the body and embarked on a course of deception. This had two effects – it allowed her to fraudulently continue to receive child benefit that she spent on drugs, and her actions also ensured that the cause of Declan’s death could not be determined due to the degree of decomposition of his remains.

Instead the AD argued that there was some expert evidence, albeit indirect, that supported his contention that Declan had been neglected (see later).

Given all the facts of the case – the pieces of the jigsaw – he submitted that a clear picture emerged - that due to her wicked recklessness Declan died, and thus Kim Hainey was guilty of murder. While the jury could consider the lesser crime of culpable homicide, the Crown argued that this lesser crime did not reflect the severe and intentional recklessness of Kim Hainey’s conduct.
The Defence narrative

At the trial, defence counsel Edward Targowski QC presented an alternative narrative of the relevant events leading to the discovery of Declan’s body.

While in her earlier adult life she had been an independent and successful young woman with a responsible job, a mortgage, and stable lifestyle, he suggested, nevertheless, that Ms Hainey was psychologically fragile, as evidenced by her response to the death of her uncle, with whom she had a close and dependent relationship. It was her significant adverse psychological reaction to this event that led to her abusing alcohol and drugs, and to depression, resulting in hospitalisation. As a consequence, she lost her job and flat.

However, she was fundamentally a well-motivated and responsible individual. In spite of her problems, when she became pregnant with Declan she gradually pulled herself together, worked co-operatively with the relevant support agencies, and was stable on the methadone programme.

It was generally acknowledged that once Declan was born she had been a loving and attentive mother. Her new flat was “spotless” and early progress was good. Indeed, unlike many heroin users, she managed to come off both heroin and methadone, so that her supervisory drug worker felt able to discontinue follow-up. There had been no professional concern about her care of Declan.

Further, there was no positive evidence of any traumatic or other unnatural cause of Declan’s sad death. Indeed, an alternative natural explanation for Declan’s death and other associated findings was available –the possibility that Declan had died of SIDS could not be discounted. Mr Targowski’s position was that Ms Hainey’s abreaction to Declan’s unexpected death had provoked a downward spiral into alcohol and drug abuse.

---

Ms Hainey’s evidence at trial

Kim Hainey chose to give evidence herself. Her position was that she had never left Declan alone for long periods. She stated that Declan had died unexpectedly overnight sometime during the summer of 2009. She found him dead in his carrycot one morning. She had attempted mouth-to-mouth resuscitation, but did not consider seeking any help from neighbours or the emergency services. She was in denial and could not accept that he was dead. However, despite rejecting the reality of that fact, her life fell apart and it was then that conditions in her flat deteriorated. It was only sometime after his death that she began to abuse alcohol and drugs again.
Section 3 – The disputed expert evidence

In the face of two conflicting narratives to explain Declan’s death, it was necessary for the prosecution to offer positive evidence of unlawful killing. Given that the pathologists’ expert position was that the cause of death could not be ascertained, here the Crown relied on expert evidence that Declan had suffered from episodes of significant malnutrition to support the core argument that he had died from wilful neglect. In the section below, I provide a rather detailed overview of the relevant expert evidence in relation to that aspect of the Crown case. I justify that because this was to become the most crucial issue both at the trial, and at the subsequent appeal.

Here the Crown’s key witness was Susan Black, professor of human anatomy and forensic anthropology at Dundee University, and a highly experienced expert witness. Dr Craig Cunningham, then a lecturer in Professor’s Black’s department, had provided joint expert reports with Professor Black, and was also a key Crown witness.

Shortly after Declan’s body was found, the investigating police team contacted Professor Black and her Dundee team, primarily to seek an opinion on Declan’s age at the time of his death, based on their expertise in forensic osteology. Later, that request apparently evolved, and they were asked to identify any skeletal trauma/pathology that may have contributed to his death, and to identify any potential indicators of neglect from the skeleton. Their skeletal assessment was founded on examination of post-mortem radiographs and computed tomography (CT) images, together with the physical examination and further radiological imaging of a number of Declan’s bones removed to Dundee for preparation and examination.

138 See, for example, Louise Scheuer and others, Developmental Juvenile Osteology (Elsevier Science San Diego 2000) 587.
139 Statement and Report for Strathclyde Police (SP/CC.SB/0410), dated 4 May 2010), Dr Craig Cunningham, Professor Sue Black.
Over time Professor Black and Dr Cunningham provided four joint pre-trial expert reports, before giving expert evidence in the criminal trial. Their evidence focused on the three issues noted above – age of death, the presence of any skeletal trauma or pathology, and indicators of neglect.

**Declan’s age at death**

Their uncontested opinion of the age of death – at between one and two years old – was consistent with other expert opinions and fitted with lay witness observations as to when Declan was last seen alive. These combined opinions suggested that Declan’s death was most likely to have occurred during July/August of 2009.

Further in their initial report they found no obvious skeletal pathology or trauma that might have contributed *directly* [my emphasis] to Declan’s death.

**Declan’s bones and claims of evidence of neglect**

It was the third area on which they reported which would prove the most significant and controversial, both at the original trial and at the subsequent appeal, for it was the Dundee anthropologists’ interpretation of some subtle bone features they claimed to find on radiological imaging and direct examination that provided the peg upon which the prosecution hung the possibility of neglect as a marker of an unnatural cause to explain Declan’s death.

“*Harris lines*”

Initially Professor Black and Dr Cunningham reported the presence of two skeletal growth arrest lines – so-called Harris lines – at the distal end of a forearm bone (and later from Declan’s femur) on the original post-mortem radiographs. They suggested that these

---

141 Growth arrest lines can develop at the ends of long bones (the metaphysis) where bone growth occurs. Such indicators of slowing of longitudinal growth may develop after physiological stress or trauma. They were first recorded by Harris in 1926 (H A Harris 'The growth of the long bones in childhood: With special reference to certain bony striations of the metaphysis and to the role of the
Harris lines might be markers for two episodes of pre-death skeletal stress from neglect and malnutrition.

These witnesses had also arranged further radiological imaging of Declan’s bones after they had been removed to Dundee. Part of this imaging process had involved digital image manipulation to enhance their clarity. Such manipulation would itself become a contentious issue during later cross-examination.

Following this further imaging, they now reported up to six growth arrest lines in Declan’s right tibia (shin). In a further written joint expert report, Professor Black and Dr Cunningham’s interpretations of these radiological findings was that these features represented up to six episodes of growth arrest at regular intervals during Declan’s short life.

**Cortical erosion and “poor bone formation”**

Additionally, after laboratory preparation and physical examination of the transferred bones, Professor Black and Dr Cunningham also commented on the “quality” of Declan’s bones. Based on comparison with an 80-year-old anatomical teaching example taken from a two-year-old Portuguese child held in a local anthropathological collection, they asserted that Declan exhibited widespread evidence of “poor quality” bone formation, that his bones were “not robust”, this view reinforcing their opinion that Declan had suffered regular periods of malnutrition and neglect.

---

142 The Scheuer collection of juvenile skeletal remains stored in Dundee University.

Challenges to the Dundee experts’ opinions

I was called as a Crown witness in order to speak on a number of matters in relation to normal child growth, development and nutrition. More specifically, I was also asked to comment on Declan’s general health from the limited routine medical records available prior to his death, and to address generally the dangers to a young infant if abandoned for prolonged periods of up to 48 hours. During cross-examination, I was also asked about my experience of the reaction of parents to the sudden death of a child.

However, most relevant to the ultimate outcome of the case, both in my pre-trial report and in oral evidence in court, in the light of Professor Black and Dr Cunningham’s joint reports I was asked to comment from a paediatrician’s perspective on the clinical significance of the finding of Harris lines in bone x-rays of young children. Here my opinion was clear – they were of very uncertain significance, and although historically malnutrition was one of the putative explanations proposed to explain them, and may be one explanation, there were likely many other benign and likely unknown causes. From a clinical point of view, they did not carry any significance, and in the context of a child protection case, the presence of Harris lines could not be usefully interpreted.144

Professor Black and Dr Cunningham’s interpretation of the Harris lines, and the suggestion of malnutrition-related poor bone formation, was also explicitly rejected by a defence expert witness, bone pathologist Dr Mangham. As far as he was concerned, Declan’s bones were unremarkable – the so-called “cortical erosion” reflected normal anatomy – and he dismissed any signs of malnutrition in the scans or x-rays.

At the trial, the Dundee experts were subjected to rigorous cross-examination, where some significant concessions were made.

First, both experts were forced to acknowledge that they had no formal medical training or experience of the medical care of children. Thus, it was suggested that they were not in a position to offer expert opinion on clinical matters in this case such as the significance of Harris lines or putative cortical erosions. Notwithstanding that challenge to their expertise, they were also asked to explain why they reached such apparently significant conclusions about these lines, when they were not even commented on by the paediatric radiologist reporting the post-mortem skeletal survey imaging to the pathologists, and “a highly experienced paediatrician” considered the clinical significance of Harris lines as very uncertain.\(^{145}\)

Dr Cunningham’s status as an expert, and the reliability of his evidence, was subjected to particular challenge.\(^ {146}\) He was forced to concede that in relation to the tests he and Professor Black had performed, their interpretation, and the contents of the various joint reports they had produced, his position was that of a subordinate trainee, learning from Professor Black. Thus, his status as an independent expert witness was seriously undermined. He acknowledged not only that he had no clinical medical training, but also that he had received no formal training in medical CT interpretation, the imaging technique he had used on Declan’s bones.

He conceded that he and Professor Black’s approach to the interpretation of the Harris lines was flawed, given that it had been partly based on digital manipulation of the images, while the variation in the number of Harris lines they had found in different long bones undermined their theory that they had been caused by Declan experiencing repeated episodes of malnutrition.

In addition, both experts were accused of failing in their duty of objectivity and impartiality in their apparent selective bias in their citation of paleontological scientific literature to

\(^{145}\) Hainey, n. 119 above, at [26].
\(^{146}\) ibid, [38] – [40].
support their conclusions. \textsuperscript{147} Professor Black, while conceding knowledge of these publications, defended her position, asserting that she considered them to represent “poor science”, thus not worthy of citation.

Dr Christina Papageorgopoulou, Institute of Anatomy, University of Zurich was called by the defence to provide expert rebuttal evidence of the uncertain nature of Harris lines from a palaeontological perspective. She quoted a number of peer-reviewed papers to support this view.

Given that uncertainty about the significance of Harris lines has been a well-established issue of academic debate within the relevant paleontological and related communities well before Declan Hainey’s death, at first glance Professor Black’s apparent selective referencing or failure to acknowledge disputed scientific interpretations may appear to be a significant error. However, as I discuss later in this thesis, for the expert witness, judgments about relevant contradictory science and which alternative views to include in opinion evidence is not straightforward.

However, while apparently accepting my opinion as to the non-specific nature of the Harris lines, Professor Black continued to assert in re-examination her view that they indicated that Declan had been subject to significant episodes of neglect and nutritional deprivation. As the Advocate Depute made clear to the jury in his closing speech, this opinion evidence of Professor Black and Dr Cunningham in relation to the sinister implications of the presence of Harris lines was a crucial part of the Crown’s case.

\textsuperscript{147} Hainey, n. 119 above, at [37], quoting C. Papageorgopoulou and others 'Harris lines revisited: prevalence, comorbidities, and possible etiologies' (2011) 23 (3) Am J Hum Biol 381, and M. P. Alfonso, J. L. Thompson and V. G. Standen 'Reevaluating Harris lines--a comparison between Harris lines and enamel hypoplasia' (2005) 29 (2) Coll Antropol 393.
Significant key events of the trial process

"No case to answer" motion

At the end of the Crown case, prosecuting counsel deleted the word assault from the original indictment. In response, given that the cause of Declan’s death was classified as “unascertained”, defence counsel submitted a “no case to answer” motion on behalf of Ms Hainey. He argued that the jury could not, on the evidence, conclude that Declan had died an unnatural death; nor that his death was the result of his mother’s conduct; nor that she had the necessary state of mind to be guilty of murder.

Following counter arguments from the Crown, the trial judge rejected this defence submission on the grounds of the extensive circumstantial evidence available to the jury.

A challenge to the Dundee experts’ relevant expertise and objectivity

Critical to the ultimate outcome of this prosecution, and particularly relevant to the focus of this case analysis, was another motion by Mr Targowski, at the completion of his defence case. He submitted that Lord Woolman should direct the jury to disregard all the expert evidence relating to Harris lines and cortical erosion given by Professor Black and Dr Cunningham. There were three elements to this submission:

- They lacked the requisite expertise – they had no medical qualifications or experience
- Their evidence fell below the standard required of an expert witness – they had failed to refer to peer-reviewed scientific papers that offered alternative benign interpretations of the significance of Harris lines
- That Professor Black’s evidence was not corroborated by Dr Cunningham. In particular, that he, as an author of their joint reports, had accepted that significant elements of their analytical processes were unconventional, and he conceded that their conclusions were flawed.\textsuperscript{149}

\textsuperscript{148} In terms of section 97 of the Criminal Procedure (Scotland) Act 1995.

\textsuperscript{149}
Lord Woolman rejected this submission.

Lord Woolman directed the jury members “to scrutinise the
chapter of the evidence relating to Harris lines and *cortical erosion* [my emphasis] with great care. [They] must decide what weight, if any, it should receive.\(^{153}\)

**The verdict**

Kimberley Hainey was found guilty by majority of Declan’s murder at the High Court of Justiciary in Glasgow on 16 December 2011. On 12 January 2012 she was sentenced to a mandatory life sentence with a minimum term of 15 years. In respect of a second charge - attempting to defeat the ends of justice - she was sentenced to 7 years imprisonment; this sentence was to run concurrently with the life sentence.\(^{154}\)

Ms Hainey subsequently appealed her conviction.

**The appeal**

Ms Hainey’s appeal was founded on two inter-related grounds:\(^{119}\)

1. That the trial judge should have acceded to the defence motion that the jury should be directed to disregard the evidence of Professor Black and Dr Cunningham, and
2. The trial judge had failed to give appropriate and adequate directions regarding the scientific and medical evidence, as set out in *Liehne*.$^{155}$

In the event, the High Court of Justiciary (HCJ) allowed the appeal and the verdict was quashed. The court held that, particularly in cases such as this, largely based on expert evidence, it was:

---

\(^{153}\)ibid, at page 17.

\(^{154}\)Kimberley Hainey was also found guilty by a unanimous verdict of two lesser charges irrelevant to my focus here.

\(^{155}\)See *Liehne v HMA* [2011] HCJAC 51. This case involved the successful appeal against her 2006 conviction by Jennifer Liehne for the culpable homicide of her 7-month-old daughter Jaqueline almost 30 years previously. The Crown case turned primarily on expert medical evidence that a variety of hospital visits by the infant leading up to her death were the result of repeated deliberate obstruction of her airway by her mother. There were multiple medical experts for both prosecution and defence. The Court of Criminal Appeal quashed the conviction, and held, *inter alia*, that the trial judge had erred in failing to provide appropriate guidance to the jury in dealing with the conflicting and highly complex expert evidence in the case.
of the utmost importance that the experts chosen by the prosecutor to provide
evidence supportive of a charge of murder should have the relevant qualifications,
competence, expertise and experience to speak to the matters they are invited to give
evidence about.\(^{156}\)

In relation to the evidence provided by Professor Black and Dr Cunningham, the court
acknowledged that, as forensic anthropologists, providing an opinion in relation to Declan’s
age at death was clearly within their expertise. However, since they were not medically
qualified, they were thus not competent to provide clinical opinions on the significance of
possible cortical erosion and or the Harris lines in relation to malnutrition and neglect.\(^{157}\)
These were issues that fell within the expertise of medical practitioners, and it was notable
that none of the medical expert witnesses supported their interpretation.

Lord Woolman was severely censured in respect to how he had dealt with the question of
these two witnesses’ status as experts and the breadth of their expertise. The court
emphasised that, in the absence of any established pre-trial judicial gatekeeper process, the
trial judge had a duty, in the interests of a fair trial, to ensure that any witness proffered to
provide expert opinion evidence on a particular matter had the requisite competence and
expertise:

Putting matters colloquially it cannot be right for a trial judge to allow an obvious
"quack" doctor to speak to a subject in a supposed expert way in relation to which he
has no qualifications, and to allow his evidence to be placed before a jury with the
simple direction that it is a matter for them to assess his competence.\(^{158}\)

This comment resulted in extensive media attention particularly focused around Professor
Black and the “quack” epithet. In the resulting furore, questions were raised in the Scottish
Parliament in relation to these adverse judicial comments apparently directed at a highly
respected academic figure. Very unusually, the HCJ quickly issued a note of apology and

---

\(^{156}\) Hainey, n. 119 above, at [49].

\(^{157}\) However close reading of the judgment might suggest that Professor Black may be entitled to
comment on Harris lines in other contexts.

\(^{158}\) Hainey, n. 119 above, at [49], per Lord Clarke.
clarification, emphasising that the “quack” comments were intended to be of a general nature rather than focused on particular individuals involved in *Haine*y.159

Returning to the appeal judgment in *Haine*y, the court emphasised that should it emerge in the course of a trial that a witness had provided opinion evidence beyond the boundaries of their expertise, the judge should direct the jury to disregard that evidence.160 In the instant case, this was particularly relevant to Dr Cunningham, whose status as an expert witness on any of the relevant issues was questionable.

The court also held that other significant elements of Lord Woolman’s charge to the jury had been inadequate. Given the complex expert evidence, the court held that he should have adopted the approach set out in *Liehne*,155 and provided some rehearsal of the medical and other scientific evidence at trial. Notably, Lord Woolman had failed to direct the jury to disregard the cortical erosion evidence of Professor Black and Dr Cunningham, since both witnesses had accepted in evidence that neither was competent or qualified to speak to that medical topic.

In contrast, his charge seemed if anything, to have highlighted that flawed evidence. Despite acknowledging in his charge that the Advocate Depute had not referred to the cortical erosion evidence in his closing speech, the trial judge had instructed the jury to “scrutinise the chapter of the evidence relating to Harris lines and cortical erosion (my emphasis) with great care. You must decide what weight, if any, it should receive”.161 The court held that this amounted to a material misdirection.

---

160 As a side issue, whether, in reality, jury members can fully ignore the impact of such evidence is questionable; see, for example, Nancy Steblay and others 'The Impact on Juror Verdicts of Judicial Instruction to Disregard Inadmissible Evidence: A Meta-Analysis' (2006) 30 (4) Law Hum Behav 469.
161 Lord Woolman’s charge to the jury, pages 16-17.
Further, given that the advocate depute had clearly highlighted in his closing speech that Professor Black and Dr Cunningham’s evidence in relation to the Harris lines was a crucial part of the Crown’s case, in failing to provide in his charge any detailed rehearsal of the medical and other evidence that could be considered to undermine their evidence – in particular the contrary opinion of other experts, Dr Cunningham’s ultimate withdrawal from his original position, and evidence led on the possibility that Declan had died a natural death – was also a highly significant judicial error.

The court’s overall conclusion was that the misdirection and failures in direction by the trial judge were highly material and had resulted in a miscarriage of justice.
Section 4 – Case commentary and signposts to next chapters

Having set out the case details in *Hainey* and described the key problems at the trial at first instance that led the HCJ to uphold Kimberley Hainey’s appeal and quash her murder conviction, in the final section of this chapter I turn to explore in more detail potential problems with flawed forensic testimony in cases of alleged infant murder or abuse, using *Hainey* to illustrate some of the key themes that I will explore in subsequent chapters.

The general claim I make is that cases involving death or alleged serious injury in young children are intrinsically different and thus pose particular problems. As I discuss below, in such cases there may be particular difficulties in distinguishing an unnatural death from natural illness, so the cases often turn on complex evidence from a variety of paediatric experts, the prosecutions are often dependent on circumstantial evidence, and the highly emotive nature of such cases may impact on the objective reasoning of fact-finders (and experts).

*Child cases are different*

As is perhaps self-evident, the prosecution of cases of alleged homicide in young children are atypical compared to more “conventional” homicide cases. Many of the features that typify such cases are evident in *Hainey*.

Unlike most violent deaths, most child victims of homicide are killed by a parent or other caregiver, and in cases involving infants, where prosecuted, the accused in almost invariably a parent. Again, in contrast to the usual marked preponderance of male perpetrators (and

---

162 Here, my use of the term ‘homicide’ covers the offences of murder, manslaughter/culpable homicide and infanticide.

163 In Scotland during the ten-year period between 2005-06 and 2014-15, a total of 43 children under the age of 16 years were victims of homicide. Where there was an accused person, 70% involved a parent and of the 14 victims aged under one-year-old, the main accused was a parent in all but one case. See Homicide in Scotland, 2014-15 report, available at: [http://www.gov.scot/Publications/2015/09/8172/downloads](http://www.gov.scot/Publications/2015/09/8172/downloads) (accessed 4 April 2018).
victims) associated with violent crime, in suspicious deaths involving infants, mothers are almost as frequently prosecuted as fathers.\footnote{164}

The reasons why mothers may kill their young children vary, but there are five patterns (or motivations) recognised in the published literature.\footnote{165} While some reflect maternal mental illness or deliberate acts founded on maternal self-interest or spousal revenge, most are the unplanned outcome of cumulative progressive child abuse and/ or neglect, the latter reminiscent of the alleged circumstances in \textit{Hainey}. There are also significant challenges specific to the forensic investigation of possible infant homicides. In most adult homicides the immediate cause of death is usually relatively obvious.\footnote{164} In contrast, in suspected homicides involving infants, it is much less common for the pathologist to find clear-cut physical injuries that indicate an obvious violent assault as the cause of death.\footnote{166} Rather, the findings are often non-specific, or open to varied interpretation.\footnote{164}

Given that the prosecuting authorities must rely almost exclusively on expert forensic opinions related to potentially contested clinical findings, in differentiating between a likely homicide and a natural death in such cases, prosecuting decisions are often highly dependent on supportive extraneous circumstantial evidence. Indeed, without reasonably cogent circumstantial evidence, such prosecutions may not be pursued.\footnote{167} Given these particular challenges for the prosecution, it is hardly surprising that there have been a number of

\footnote{164} Fiona Brookman and Jane Nolan 'The Dark Figure of Infanticide in England and Wales: Complexities of Diagnosis' (2006) 21 (7) J Interpers Violence 869.\footnote{165} Susan Hatters Friedman and Philip J. Resnick 'Child murder by mothers: patterns and prevention' (2007) 6 (3) World Psychiatry 137; P. J. Resnick 'Child murder by parents: a psychiatric review of filicide' (1969) 126 (3) Am J Psychiatry 325.\footnote{166} Brookman and Nolan, n. 164 above. These ambiguous features may involve suspected suffocation, or shaking injuries, or other “non-specific” causes of death.\footnote{167} For example, the Crown Prosecution Service’s guidance in relation to prosecution of “shaken baby” cases, which emphasises the need for corroborative circumstantial evidence, an issue I discuss in detail in Chapter 7.
wrongful convictions linked to such cases. Again, atypically, women have been prominently involved in them, and the circumstantial evidence used to bolster such prosecutions is often based on gendered and role stereotypes.

Circumstantial evidence in alleged infant homicide – of the “mad” and the “bad”

Not surprisingly, and reflecting the highly emotive context, cases such as Hainey are almost invariably associated with a significant media profile, sensationalised and potentially highly prejudicial reportage, and an attendant moral panic and clamour for justice. There is arguably an inherent danger here that emotion and societal anger may influence any legal decisions, including those of prosecutors, even if the evidence is equivocal. Commonly, as in Hainey, in building a circumstantial case around the accepted and disputed facts, the prosecution narrative presents the accused as the antithesis of the normative (idealised) model of motherhood, as determined by societal stereotypes. As Cossins points out, in such cases the stereotypical cultural model of the “good” mother is set against prosecution narratives of “bad” (or “mad”) mothering. Here, atypical lifestyle features, or even normal maternal behaviour, may acquire a sinister significance.

---

168 Prominent UK cases involving infant deaths include wrongful convictions such as Clark, Cannings, Anthony and Liehne noted above. See also, for example: R v Holdsworth [2008] EWCA Crim 971.
170 The alternatives of “mad” or “bad” reflect a longstanding recognition that the criminal justice system treats women and men differently in relation to the prosecution and sentencing of crime, notably filicide (see: for example, Ania Wilczynski 'Mad or bad? 'Child-Killers, Gender and the Courts' (1997) 37 (3) Brit J Criminol 419. However, for women the narrative may be best characterised as - if not “mad”, they must be “bad”.
172 Perhaps typified in the post-Savile era by a number of recent acquittals of high profile individuals prosecuted on the basis of unsubstantiated claims of sexual abuse.
In *Hainey*, in the absence of *direct* pathological evidence of unnatural death, the prosecution relied primarily on circumstantial evidence that Kimberley Hainey had been criminally neglectful and so responsible for Declan’s death. Here the prosecution case narrative and legal argument developed against Ms Hainey was formulated to transform circumstantial evidence into indicators (or “proof”) that she was guilty. Thus a damaging narrative was developed involving drug and alcohol abuse, deceit and dishonesty, and, of course, selfish child abandonment. Although we can never know, it seems reasonable to speculate that the jurors' assessment of Ms. Hainey as a mother was a significant factor in the decision to find her guilty of murder.

Of course, having highlighted some of the somewhat atypical aspects of investigation and prosecution of alleged infant homicide cases, it is important that I acknowledge that convictions in such cases have to be predicated on a conclusion by the relevant jury that the deaths were unnatural, conclusions strongly dependent on the inferences drawn from expert evidence.

Thus, I now turn to briefly highlight some key elements of the (flawed) expert evidence in *Hainey* that point to specific difficulties in relation to paediatric expert evidence that I address much more fully in later chapters.

---

175 In *Hainey* the prosecution argued that the “Harris lines” pointed *indirectly* to a possible cause of death.

176 There are strong echoes here with the negative presentation of maternal behaviour in the Australian case of Kathleen Folbigg, convicted in 2003 for the multiple filicide of her four children over a decade: (Folbigg, n. 86 above). See also Sharmila Betts and Jane Goodman-Delahunty 'The case of Kathleen Folbigg: how did justice and medicine fare?' (2007) 39 (1) Aust J Forensic Sci 11.

177 However, occasionally such juror information becomes available. In the Chamberlain case, n. 86 above, it seems clear that at least one juror was strongly influenced to convict Lindy Chamberlain based on her perception of her as a mother. See: E Cunliffe, 'Weeping on cue: the socio-legal construction of motherhood in the Chamberlain case' (LMM, University of British Columbia 2003).
The expert evidence – signposts for further analysis

The expert issues and related flaws in Clark, Hainey and similar infant cases appear to be focus around three inter-related themes:

1. individual expert(s) failings or misconduct; (2) expert bias; and (3) the admission of unreliable expert evidence.

So, with those three overlapping themes in mind, I wish to consider why the Hainey prosecution went wrong, and to highlight key issues relevant to paediatric expert evidence that I consider in subsequent chapters.

Failures by individual experts

As I noted in Chapter 1, the obligations of an expert in providing opinion evidence are long established in the civil procedures of the English and Scottish courts. These obligations are equally relevant in criminal proceedings, as recently emphasised by the English Court of Criminal Appeal in a group of consolidated appeals in R v Harris, involving a number of so-called “shaken baby syndrome” convictions. These duties of experts are approved and equally relevant in Scotland.

There are two areas in that list of expert duties that seem particularly relevant to Hainey.

---


**Failure to stay within the boundaries of relevant expertise**

An expert should make it clear when a particular question or issue falls outside his expertise. As the above quotation indicates, one of the explicit expectations of an expert is that they must only provide opinion evidence derived from a field of knowledge that they are “qualified” to opine on, and it is the responsibility of the party that instructs them, and the presiding judge, to ensure that a particular witness is an expert in that field.

Such questions of expert qualification may seem (superficially) straightforward; in clinical negligence cases, for example, generally only evidence from medical practitioners with qualifications and experience in the relevant specialty area would satisfy that requirement, the expertise derived from formal training and applied practice, often encompassed within formal active membership of an appropriate professional body.

But of course, in many circumstances it is not quite so simple and recognising the boundaries of personal expertise and when it is appropriate to defer to the opinions of others may not be straightforward.

In *Hainey*, Professor Black and Dr Cunningham were considered to have stepped outside the boundaries of their expertise when they provided testimony in relation to the putative Harris lines (and cortical erosions) as markers of neglect that linked indirectly to the cause of Declan’s death. And albeit that Professor Black apparently accepted my opinion that the cause and significance of such lines were highly uncertain, neither deferred to my view. Rather both experts reverted to their original conjoined opinions that they considered the Harris lines as markers of neglect.

Thus, not only were they considered to have strayed outwith the bounds of their expertise, but the trial judge was held at fault for not recognising this failing and intervening, even when formally requested by the defence to exclude their evidence on those matters.

---

182 Harris, n. 84 above, at [271].
Thus while the Hainey appellate court considered it quite appropriate for the Dundee team to provide an opinion on Declan’s age at the time of his death - an issue that fell within their expertise in forensic anthropology, given their lack of medical qualifications and expertise, the court rejected the Dundee anthropologists expert status “as qualified to speak to the status of [Declan’s] health particularly in relation to how death might have come about.”

Rather the court explicitly held that the question of whether the two key disputed bone issues - possible cortical erosion and Harris lines - could be linked to the cause of Declan’s death - were clearly set within the expert ambit of medical practitioners rather than anthropology.

Further, the court held that these incompetent opinions were significant, concerned that despite robust challenge at cross-examination and contradictory opinions from other competent medical experts, these apparently unqualified views might well have influenced the jury to reach a conclusion of Ms Hainey’s guilt.

**Whose science is it anyway?**

This raises an interesting and broader question, for it is clear that the topic of Harris lines clearly straddles two very different knowledge domains with apparently diametrically opposite views about their interpretation, albeit comparing findings in the bones of the living or recently dead, compared with ancient bones.

The judgment in *Hainey* seems to raise this issue of ownership of an area of expertise quite explicitly, so the Harris lines issue in *Hainey* could be viewed as raising separate important questions about expert evidence and the boundaries of expertise:

---

183 Hainey, n. 119 above, at [49].

184 This dismissal of the Dundee experts’ opinions in Hainey is considered by Raitt to reveal the court’s fundamental misunderstanding of the appropriate application of forensic anthropology (Fiona E. Raitt, *Evidence: principles, policy and practice* (2nd edn, W. Green Edinburgh 2013) 306, at [4-56]).

185 Hainey, n. 119 above, at [47].
• Does a particular group of experts “own” a particular field of expert knowledge?
• If there is some overlap, which field of science (and the associated literature) is it appropriate to apply to a particular disputed fact?
• Or does it depend on the context of how that knowledge is being applied?

It seems that the latter question was the crucial one in Hainey, since the court held that opining, even indirectly, on a possible cause of death was a medical matter.

Despite the apparent ease with which the court in Hainey determined from which knowledge domain experts should be drawn to provide assistance to the court’s decision-making on that particular issue, albeit not further explained, such boundaries are often not so easily defined by broad professional labels. And even within the medical domain, the criminal courts seek to define very precisely where the boundaries of relevant forensic expertise lie and restrict expert opinion evidence to those with the appropriate qualifications and narrow areas of practice.  

However, in some clinical practice areas, such as child protection, such boundaries may be especially difficult to define. As I noted in Chapter 1, in such a multi-disciplinary professional environment, it is common to encounter multiple expert opinions from different but overlapping clinical perspectives. It is here, in the fuzzy overlap of clinical expertise, where the paediatrician often acts as a generalist expert, bringing together the various views of others to provide an integrated conclusion about the cause of an infant’s injuries.

---

186 R v Clarke & Morabir [2013] EWCA Crim 162. Here the CACD held that the court at first instance had quite appropriately restricted an eminent expert pathologist specialising in fractures and metabolic bone disease from providing an opinion as to cause of an alleged murder victim’s death, agreeing that the expert “did not have the experience or expertise to consider all of the causes of death” (at [77]), unlike a Home Office registered forensic pathologist.

187 As illustrated, for example, in Liehne, where multiple clinical and pathology experts were involved.

188 It is clear, however, that it is quite appropriate for an expert to refer to published literature from another area of expertise and to form an opinion partly founded on that literature, as long it has been made clear where their own expertise lies, and their use of such literature has a direct bearing upon their area of expertise. However, it is only appropriate for an expert to provide an opinion founded on
However, if experts do not recognise the boundaries of their expertise, or choose to ignore them, they may be drawn to provide unsound but highly influential evidence.

This has been recently illustrated in the high-profile GMC professional disciplinary case that followed significant judicial concern about the expert evidence of paediatric neuropathologist Dr Waney Squier in relation to her evidence in several “shaken baby” cases. In a protracted disciplinary hearing, the Medical Practitioners Tribunal Service had found, *inter alia*, that Dr Squier had inappropriately provided expert evidence from a variety of domains, including paediatric neurosurgery, biomechanics, paediatric medicine, neuroradiology, and forensic medicine when she had no formal training or clinical responsibilities in these fields, and thus could not be regarded as an expert in them, albeit she had a significant knowledge of published literature in those fields.

While useful guidance on the criteria with which a *medical* expert may judge the limits of their personal professional expertise may be derived by a careful reading of the GMC erasure determination of Dr Squier, for child protection paediatricians providing expert opinion evidence in relation to child death or injury, defining the appropriate limits of expertise may be particularly problematic both for the experts, and the courts. Indeed the court in Dr Squier’s successful appeal appeared to acknowledge that in relation to some particular contested paediatric forensic issues, the conventional boundaries of expertise should be considered extremely flexibly.

---

190 Her appeal was against erasure was successful; her licence to practise was restored but she is barred from child abuse expert witness activity.
191 Based on Dr Squier’s MPTS hearing determination (note 189 above), expert clinical status would be expected to reflect a background of formal training in the area of clinical practice, on-going professional development in that field, registration by an appropriate professional body as qualified in that clinical field, and that involvement in the clinical area would be part of your regular practice.
This problem of reconciling disparate expert views about a disputed issue that may arise from different knowledge domains is a particular challenge for the courts in relation to causation in “shaken baby” cases, and it is an issue I return to in Chapter 7.

Overlapping paediatric expertise and conjoined opinions

A separate but related issue that Hainey illustrates is the potential problems that conjoined expert opinions may pose. Paired working is customary in many aspects of child abuse assessment, and in some specific circumstances clinical forensic examinations are often performed jointly. Thus in many areas of Scotland, a senior child protection paediatrician and a police Forensic Medical Examiner, albeit with different clinical backgrounds, provide joint examinations and reports.

Similarly, as can be seen in Hainey, in deaths involving children where homicide is a possibility, it is mandatory for the autopsy to involve two pathologists - a forensic (Home Office accredited) pathologist and a paediatric pathologist – who prepare joint reports but individually may provide opinion evidence at any subsequent trial. Although such so-called “double doctor” autopsies are principally orientated to ensure that any significant findings may be established by corroborated evidence, the pathologists provide complementary competencies and offer integration of expertise, the paediatric pathologist focused primarily on disease exclusion (natural causes) and the forensic pathologist focussed on unnatural explanations.

Such conjoined reporting is not a problem if both experts are of equal status and their respective roles are clear. However, as seen with the expert anthropologists in Hainey, the

---

192 While the original expectation was that such joint examinations were only necessary when an individual doctor did not have all the necessary skills, and the examiners would provide complementary expertise, increasingly such joint examinations are the norm. See: Guidelines on Paediatric Forensic Examinations in Relation to Possible Child Sexual Abuse, Royal College of Paediatrics and Child Health and the Faculty of Forensic & Legal Medicine, October 2012; available at: www.fflm.ac.uk/wp-content/uploads/documentstore/1352802061.pdf (accessed 1 June 2018).
validity of an expert consensus when there is a significant differential status of the experts involved in such joint assessments may be at least questionable.

**Failure to justify objectively the basis of an expert opinion**

[A]n expert witness should in particular explain why any material relevant to his conclusions is ignored or regarded as unimportant. It is generally expected that for an expert opinion to be truly of assistance to a court, the expert must not only give an opinion, but should explain the basis for that opinion. The expert’s reasoning should be supported by a balanced literature review that acknowledges any relevant contrary published work that may undermine the basis for their conclusion, and explain why this contrary literature has been rejected.

Setting aside the issue of boundaries of expertise, the *Hainey* judgment highlighted the apparent selective nature of the osteo-archeological academic literature cited by Professor Black and Dr Cunningham in their joint reports to support their conclusions. Justified by Professor Black as simply rejecting “bad science”, such apparent “cherry picking” of supportive literature has been a feature in other high profile wrongful conviction cases.

This raises a broader issue that experts face – how to appropriately interpret and apply the evidence base in relation to a particular disputed issue. Science thrives on debate and academic disagreement, and most knowledge is in a continuous state of flux. As I argue later in Chapters 5 and 7, some of the evidence base that is purported to underpin expert evidence

---

193 Wilson v HMA, n. 181 above, at [60].
194 For example, in England the procedural rules relevant to expert evidence in the Criminal, Civil and Family Courts all emphasise that an expert must make clear that where there is a range of opinion on any question to be answered by the expert, summarise the range of opinion, and explain the reasoning behind the conclusion they have reached. See, for example, the particularly detailed instructions in relation to the content of an expert report approach specified by for the Family Court Practice Direction (25b): [http://www.justice.gov.uk/courts/procedure-rules/family/practice_directions/practice-direction-25b-the-duties-of-an-expert,-the-experts-report-and-arrangements-for-an-expert-to-attend-court#IDAFOIT](http://www.justice.gov.uk/courts/procedure-rules/family/practice_directions/practice-direction-25b-the-duties-of-an-expert,-the-experts-report-and-arrangements-for-an-expert-to-attend-court#IDAFOIT) (accessed 5 April 2018).
195 *Hainey* appeal, n. 119 above, at [37].
196 It was a notable criticism of Dr Squier in relation to her evidence in several “shaken baby” cases; see Squier v GMC, n. 189 above.
in child protection is of questionable reliability, and some has certainly been exposed as “bad science”.

Yet while it may be argued that in such circumstances, cross-examination and the alternative views of other experts usefully exposes such scientific disputes and provide a counterbalance to expert subjectivity, this depends on the non-scientific lawyer becoming aware that such disputes exist. Given that by definition, lawyers are not scientific experts, it does seem that an expert has a clear duty to the courts to acknowledge such contrary opinions, whether they agree with them or not. Clearly at a minimum there would be an expectation that experts should acknowledge when relevant areas of their evidence are controversial.

**Expert bias**

Let me turn to the question of expert bias, the second key theme attributed to wrongful convictions in cases of infant harm. While a number of forms of expert bias are recognised, a question I want to raise here is whether child death (or serious injury) cases pose particular risks of bias (or at least influence decision-making more generally) because of the particular nature of such cases.

It seems trite to acknowledge the highly emotive nature of cases involving children, especially suspected homicide, but given that, it seems reasonable to suggest for those professionals directly involved in such cases, including expert witnesses, it may be difficult to maintain a professional distance and remain objective. There is certainly some evidence

---


that professionals dealing with infant homicides find them significantly more emotionally challenging than adult cases.\footnote{199} And in the context of a trial itself, given the established vulnerability of jurors’ objectivity to sometimes gruesome and distressing evidence,\footnote{200} it also seems reasonable to suggest that cases such as 

\textit{Hainey} are very likely to induce negative bias towards defendants, given the adverse reactions that such evidence may generate.\footnote{201}

Are expert witnesses (paediatric doctors or others) involved in such cases also vulnerable to a loss of objectivity in their forensic interpretation of clinical and other relevant features? Such bias can, of course, be explicit or implicit. Both are known to occur in the context of forensic testimony. Explicit biases may include an unjustifiable promotion of a particular personal theory to explain forensic findings.\footnote{202} It is also now evident that implicit bias is increasingly acknowledged as a pervasive issue relevant to many domains of forensic science and testimony.\footnote{203}

Might such biases at least partly explain one of the most perplexing questions that arises from an expert witness’ perspective in \textit{Clark} and \textit{Hainey} — how did two such highly experienced expert witnesses as Professors Meadow and Black — indeed, one honoured as a

\begin{itemize}
\item [\footnote{200}]{See, for example, Kayo Matsuo and Yuji Itoh 'Effects of Emotional Testimony and Gruesome Photographs on Mock Jurors' Decisions and Negative Emotions' (2016) 23 (1) Psychiatr Psychol Law 85. The nature of the charges in Hainey, and the evidence that had to be led - the jury shown photographs of Declan’s mummified and partly decomposed body, for example - meant that this was inevitably a highly emotional trial.\footnote{201} D. A. Bright and J. Goodman-Delahunty 'Gruesome evidence and emotion: anger, blame, and jury decision-making' (2006) 30 (2) Law Hum Behav 183.\footnote{202} Typified, perhaps, by the “temporary brittle bone disease” diagnosis promoted by the subsequently discredited Biochemist Dr Colin Paterson (R. Nicholl and others, 'Can we have a permanent end to ‘temporary brittle bone disease’?' (2012) 97(8) Arch Dis Child 762; Colin R. Paterson ‘Temporary brittle bone disease: the current position’ (2011) 96 (9) Arch Dis Child 901), or Dr Squier’s unshakeable claim that a “shaken baby” suffered from perinatal HIV encephalopathy in the face of irrefutable evidence to the contrary are good examples; see Squier, n. 189 above).\footnote{203} See, for example, I. E. Dror, D. Charlton and A. E. Peron 'Contextual information renders experts vulnerable to making erroneous identifications.' (2006) 156 Forensic Sci Int 74.}
\end{itemize}

78
Knight and the other, more recently, a Dame, largely in recognition of their respective forensic expertise – apparently get things so wrong, and become the subject of judicial criticism?

These were “celebrity” experts, highly regarded in legal circles and often pre-emptively snapped up early in a case by prosecuting teams, before the defence had the opportunity, or the agreed funding, to instruct them. 204

Perhaps, as some have suggested, 205 celebrity status may become part of the problem with some expert witnesses - too eminent, authorities unto themselves, where prosecuting counsel would defer to their opinion without question, rather than probing the basis for their opinions, or the boundaries of their expertise. 206

Certainly, Professor Meadow had a long-established association with the prosecuting authorities in cases of suspected infanticide in recurrent SIDS cases, both as an educator to medical profession, the judiciary and the police in relation to the issue, and as an expert witness in that regard on many occasions. 207 Indeed, given his promotion of the aphorism that became known as “Meadow’s law”, one could argue that his bias in relation to that concept was explicit.

Both these eminent experts were unaware of or chose to ignore published literature that undermined their opinions. Does this reflect another form of bias – that once each had reached a conclusion did these experts close their minds to literature that challenged it?

Given the emerging evidence that bias from a number of sources is a pervasive threat to objectivity in many forensic domains, what about the situation of the paediatrician, toiling at the coalface of clinical child abuse work, imbued with the values of child advocacy and therapeutic care of children, while expected to provide expert witness testimony to the

204 See Professor Black’s evidence to the House of Commons Science and Technology Committee, ‘Forensic Science On Trial’ report, n. 100 above, at [141].
205 Sutherland, n. 109 above.
206 What lawyer, for example, untrained in science and medicine, is going to question the draftsman of the eponymous Meadow’s Law: ‘One sudden infant death is a tragedy, two is suspicious and three is murder, unless proven otherwise’: See Meadow, n. 105, above.
207 Noted in his GMC hearings.
family and criminal courts? This question of expert bias in a paediatric setting is a key focus of Chapter 3.

**The admission of and failure to manage unreliable expert evidence**

The third theme I noted earlier in relation to issues underlying flawed convictions in cases of infant harm is the admission into court of unreliable expert evidence.

In relation to the particular focus of this thesis, such unreliability generally arises either because of intrinsic weaknesses in the science itself or its application in an inappropriate context, or the opinion may be provided by a witness who lacks the appropriate expertise to speak to that issue. And of course, these two issues may coincide.

As I stated in my written and oral evidence in *Hainey*, the issue of Harris lines and their theoretical implications in relation to neglect or malnutrition does not appear at all in any published child protection literature. Indeed, they are simply ignored in paediatric clinical practice. Does this mean that they are truly insignificant in the context of forensic child abuse practice, or does the absence of a paediatric literature simply indicate that the issue has not been considered? I return to consider the strength or otherwise and the difficulties inherent in the currently available evidence base on which paediatric expert forensic opinions must be founded in Chapter 5.

While the responsibility for ensuring that a witness is appropriately skilled to provide the evidence to be adduced and for ensuring that the evidence itself is scientifically valid must fall to the personal responsibility of the witnesses themselves in meeting their duty to the court, the lawyer who seeks to adduce their evidence also shares this responsibility.

However, it is ultimately for the judge to apply their discretion to ensure that only sufficiently reliable expert evidence is put before the court.

I suggest that the admission of the flawed expert evidence in *Hainey* represented failures by all three of these actors. Here the key prosecution witnesses who purported to provide incriminating expert testimony applied scientific knowledge of at least arguable forensic
validity from one field of science to a disputed fact that sat within another scientific sphere and interpreted that misapplied science to reach conclusions they were not qualified to make. The prosecutor who led this witness’ evidence apparently failed to recognise the limitations of the witness’ evidence and resisted a challenge to its reliability. Similarly, the trial judge, *inter alia*, also allowed this unreliable evidence to be adduced. At least in *Hainey*, unlike many other paediatric high-profile cases, the criticism was not just levelled at the witnesses. Indeed, and unusually, the coruscating censure of Lord Woolman was as overt as the indirect but severe criticism of the expert witnesses. *Hainey* is also significant for the suggestion that emerged in the dicta suggesting the introduction of significant procedural changes in appropriate Scottish cases. The problems in relation to expert evidence exposed by *Hainey* led the court to suggest that a pre-trial hearing of complex and disputed medical evidence, a procedure already adopted by the English courts, might be considered in relevant Scottish cases, the purpose to narrow down the contested medical issues focused around expert (medical) evidence.

The suggestion of such pre-trial hearings had been raised before, notably, in the context of complex child death cases. And in English civil litigation cases, such a meeting of experts to reach a consensus and to narrow areas of disagreement — is well established. Indeed such a scheme involving a pre-hearing meeting of experts is now also expected within the

---

208 Generally direct public criticism of the lawyers and judiciary involved in such cases is at most, muted, if not absent. See, for example, the report of the House of Commons Science and Technology Committee, noting, in relation to the miscarriage of justice in *Clark*, that “little attention was given, at least in public, to the lawyers and judges involved, who may have been able to prevent the miscarriage of justice from being carried out, but failed to do so.” This reflected evidence to the committee by Professor Sir Alec Jeffreys (of DNA fingerprinting fame) that the flaws in Professor Sir Roy Meadow’s statistical evidence [should have been] “tracked right at the beginning”, describing it as “a failure not only of the experts but also of the courts”; House of Commons Science and Technology Committee, ‘Forensic Science On Trial’ report, n. 100 above, at [169].
209 R v Henderson; R v Butler; R v Oyediran [2010] EWCA Crim 1269.
210 Baroness Kennedy, n. 131 above, page 8.
Scottish Proof Hearing process in relation to disputed medical evidence in child protection cases.

However, as Maher has pointed out, the dicta in *Hainey* lack clarity. If the HCJ proposes that in selected criminal cases the trial judge should provide a “gatekeeping” role in relation to the reliability of complex forensic or medical expert evidence, are our judges up to this task? For as noted by Edmonds, and evident in *Hainey*:

> [M]any...judges are not in a position to respond ...to the assessment of validity and reliability or appreciate the subtle methodological and statistical issues associated with forensic science and forensic medicine...[rather] judges are reluctant to exclude incriminating expert evidence (if it may have some probative value).

Important questions arise here. What threshold level of reliability must be met before it is considered safe to admit such evidence before a jury? And should there be different thresholds for expert evidence based on “established” as opposed to “novel” science?

This issue of the relationship between discretionary admissibility and reliability thresholds is an important one that I consider in detail in Chapter 7.

---

212 Under the Act of Sederunt (Sheriff Court Rules) (Miscellaneous Amendments) (No 2) 2013 (SSI 2013/139), which came into force on 3 June 2013. However, unlike the “hot-tub” process, in this and the English civil pre-trial meeting of experts [Woolf reforms], the judge or lawyers are absent, the experts themselves conducting a roundtable discussion, and producing an agreed list of the agreed and disputed facts relevant to the areas of expert evidence. In contrast, in the now well-established Australian process known colloquially as “hot-tubbing”, and more formally termed *concurrent expert evidence* (CEE), following such a pre-trial expert meeting, the experts provide concurrent opinions to the fact-finder on disputed matters. In particular, enthusiasts of the process suggest it is particularly useful “where there are complex scientific issues”; see lecture by Justice Pater Garland to the Oxford Law Faculty, 1 December 2015, recounting the New South Wales experience of CEE; available at: [http://www.supremecourt.justice.nsw.gov.au/pages/search.aspx?k=concurrent%20expert%20evidence #k=concurrent%20expert%20evidence](http://www.supremecourt.justice.nsw.gov.au/pages/search.aspx?k=concurrent%20expert%20evidence #k=concurrent%20expert%20evidence), accessed 3 May 2018.

213 Maher, n. 120 above.


215 Such indicia of reliability might be adopted from those suggested by the Law Commission of England and Wales; see: The Law Commission (2011), n. 101 above, at [148].
Conclusion

In this chapter, from the unusual perspective of an insider, I have provided a detailed overview and analysis of a recent high profile wrongful conviction in Scotland involving an alleged child murder.

As we have seen, *Hainey* offers an insight into the emotive environment and technical complexities such child cases typically involve. Although each case is fact-specific, underlying each unique narrative are common themes that can be usefully considered in relation to the specific challenges that the use of paediatric expert evidence poses in such cases.

In this chapter I identified within a real case issues that I suggest are particularly relevant to the difficulties intrinsic to paediatric expert evidence. In the following chapters I examine these in detail. The overlapping issues I highlighted include potential biases and emotional responses that may undermine expert objectivity, the quality of the science (evidence base) on which expert opinions must be founded, and the challenges court’s face in recognising and excluding unreliable expert evidence. It is these core issues that I consider in future chapters.

In the next two chapters my focus is the forensic diagnostic process and issues specific to paediatric practice that might undermine the objectivity essential in the provision of expert evidence in cases of alleged infant harm.

In Chapter 3, I first focus on cognitive aspects of human decision-making and how these relate to medical diagnosis, before exploring potential sources of cognitive bias in paediatric forensic practice. Then in Chapter 4 I turn to consider whether emotional issues such as empathy and the caring values intrinsic to paediatric practice may separately threaten the objectivity that should underlie paediatric forensic interpretation.
Chapter 3: Unavoidable cognitive biases: the soft underbelly of paediatric forensic evidence?

Cognitive biases are tendencies to process information in ways that are influenced by various motivational factors, social factors, and a tendency to employ various mental shortcuts (heuristics), some of which lead us to erroneous conclusions.216

Introduction

As I discussed in Chapter 1, the core theme of this thesis is to identify a number of frailties peculiar to paediatric expert opinion evidence that might explain the relative prominence of wrongful criminal convictions involving alleged child murder or abuse that have occurred in most common-law jurisdictions, and where forensic paediatric evidence has been the focus of particular concern.217 At its heart, one can argue that such criticisms reflect a mismatch between a normative expectation (at least by the courts) of how experts in child cases should have performed when providing opinion evidence to assist a trier of fact,218 and how the experts did in fact perform.

In my analysis of Hainey in the preceding chapter I raised issues that were not mentioned in the HCJ’s review of the case: I speculated that in such cases expert witness’ opinions may have been influenced by a variety of contextual but irrelevant factors – emotional identification with the child victim, perhaps a degree of adversarial bias, or other factors that may have led to the selective use of supportive literature to justify their opinions, the latter representing an example of confirmation bias or the related “myside” bias.219


217 Law Commission (2009), n. 101 above, at paras [2.16 – 2.24], where various child homicide cases are the most frequent examples used to illustrate problems with the use of expert evidence in the English courts.


219 Myside bias occurs when individuals evaluate or generate evidence, and test hypotheses in a manner biased toward their own prior opinions and attitudes. See: Keith E. Stanovich, Richard F.
This possibility would not be unusual. It is now recognised that in most, if not all, areas of forensic practice, including forensic anthropology, \(^{220}\) a variety of cognitive biases may commonly interfere with professional judgments, influencing the way forensic practitioner’s conclusions are developed and expressed. \(^{221}\)

Despite the increasing recognition of the threats to expert objectivity such cognitive biases pose, this is an issue that, with notable exceptions, \(^{222}\) has barely been acknowledged in the specific domain of paediatric forensic practice.

Thus, in this chapter I focus on the largely intuitive nature of expert clinical diagnosis and the reasoning process involved in paediatric forensic assessment and explore the potential for a number of irrelevant contextual factors to induce (unconscious) judgment errors.

Later, in Chapter 4, I address a separate but related question – whether the emotional impact of child cases may independently influence expert reasoning and undermine the expected objectivity of forensic assessment.

---

West and Maggie E. Toplak 'Myside Bias, Rational Thinking, and Intelligence' (2013) 22 (4) Curr Dir Psychol Sci 259.


These issues have received minimal attention in relation to paediatric expert evidence, yet, as I will show, there is a strong theoretical case that such factors are likely to be operative and need to be taken account of in the context of such forensic case assessments.

In exposing these potential risks, these two chapters provide a significant contribution to the novel scholarship that this thesis provides.

**Chapter structure**

This chapter is presented in four asymmetric sections.

First, in order to put into context the particular issues that may underlie unsound paediatric clinical judgments and thus the flawed expert opinions that may lead to unsafe convictions, it is necessary to appreciate how human judgment and decision-making works more generally. Thus, in **Section 1**, I begin with an overview of the evolution of current understanding of human judgment and decision-making.

Against this background, in **Section 2** I consider the cognitive and other processes involved in expert judgments generally, before focusing on expert medical judgments and decision-making, particularly the clinical diagnostic process.

Having laid this groundwork, in **Section 3** I briefly discuss the general issue of cognitive biases and how they relate to forensic expertise before moving on in **Section 4** to examine such potential biases in the context of clinical child protection work. Here I explore the judgments that are involved when paediatric forensic clinical decisions are made, and the various factors that might undermine the objectivity of the conclusions paediatric expert’s reach.

Setting aside explicit emotional factors that I deal with in Chapter 4, I argue that the largely intuitive nature of much applied expertise, together with the unavoidable exposure to a number of biasing contextual factors particularly associated with child abuse investigations might undermine the judgments on which paediatric forensic opinions are based.
First, let me set the scene with a necessarily brief description of the evolution and current understanding of the cognitive processes involved in human judgment and decision-making more generally.

**Section 1  Human judgment and decision-making under uncertainty**

All of us make many decisions every day. Most are unremarkable and barely register – what to wear, whether to go by train or car. Other decisions we make are highly significant and may have major implications for other people as well as ourselves – whether to apply for a new job overseas, or to develop a new relationship. Some may be fundamental – whether to have surgery, or chemotherapy, or neither. Some of these decisions are simple binary choices between alternatives; others may involve a series of sequential judgments.

Many important decisions we make involve uncertainties, and often we must base decisions on incomplete information, or on the advice of others. This is where judgment comes in. For each decision that we make (or conclusions we reach) follows a judgment of some kind, where we consider the pros and cons of the various options. Whether we recognise it or not, our decisions are highly sensitive to external influences. They also tend to reflect our personal (internal) preferences and values, so whether a choice we make is “correct” – i.e. rational from others’ perspectives – must be judged against these subjective factors.

**A brief evolutionary history of human judgment and decision-making**

It is not possible (or justified) within the constraints of this thesis to provide a detailed literature review of the vast and wide-ranging scholarship on human judgment and decision-making (JDM) under conditions of risk and uncertainty that has evolved in the last half

---

223 Of course, not all judgments result in decisions, but each decision implies some judgment; Michael Lewis, *The Undoing Project: A Friendship that Changed the World* (Allen Lane London 2017) 362, at page 251.

However by way of background and to identify key themes, in this section I provide a very brief and selective overview of key relevant decision science concepts in order to provide a theoretical framework against which to consider the highly relevant cognitive processes involved in the applied expertise of clinical diagnosis and paediatric forensic judgments.

JDM scholarship can be separated into three interrelated themes that focus on the normative (how individuals ideally should judge and decide), the descriptive (how and why people make decisions in reality), or the prescriptive aspects of human judgments (how judgment and decision-making processes might be improved). My focus here is particularly on the normative and descriptive models.

While much of the early focus of such scholarship was on economic choices, it has become clear that the issues revealed are relevant to judgments in all human behavioural domains, with the same cognitive processes applicable irrespective of the setting.

---


226 I acknowledge that there are a number of other significant theoretical models of the dynamics of human decision making and choice – for example, the information processing model (John W. Payne, James R. Bettman and Eric J. Johnson, The adaptive decision maker (Cambridge University Press Cambridge 1993) 261.

227 Reid Hastie and Robyn M. Dawes, Rational choice in an uncertain world: The psychology of judgment and decision making (2nd edn, Sage London 2010) 392.
Changing concepts of human rationality

Unbounded rationality: optimisation

Although understanding had been evolving for centuries, until the mid 1900s, the dominant model of human decision-making was based on a neo-classical theory of unbounded rationality, typified by optimisation models such as Bayesian inference and expected utility maximisation. These theories, founded on probability estimates and formal algorithms, assumed that the decision-maker, nominally with unlimited time and cognitive abilities, would assess the various probabilities and the likely benefits (gains) of each option and so make a logical choice that would maximise the possible benefit (utility) that could be achieved. In the face of absent information or uncertainties, a group of algorithms characterised as optimisation under constraints were developed.

These theoretical models of decision-making became well established. Extensively applied to operational planning under the pressurised conditions and enforced interdisciplinary working of the Second World War, research on applied probabilities and decision-making rapidly progressed, reflected in the emergence of artificial intelligence and game theory in the post war years. Indeed, such was the confidence in the theory that in the 1950s it

---

231 See, for example, W. Edwards 'The theory of decision making' (1954) 51 (4) Psychol Bull 380.
232 This approach is now applied in many modern domains such as air traffic control or production planning. The mathematics is highly technical (and beyond my personal cognitive abilities!), but see, for example, Amine Lamine and others 'Solving constrained optimization problems by solution-based decomposition search' (2016) 32 (3) J Comb Optim 672.
provided the rationale for the “mutually assured destruction” deterrence doctrine of the nuclear umbrella under which we still shelter.\textsuperscript{233}

However, notwithstanding their widespread application and (continued) use, it became clear that in real life the outcomes of much human decision-making did not match those predicted by these normative probabilistic models.

Part of the reason for this is a simple but fundamental distinction: the cognitive processes involved in decisions under conditions of risk – when the probabilities of various outcomes can be calculated, is very different from decisions made under conditions of uncertainty, the condition under which most real-life decisions are made.\textsuperscript{234} As I discuss below, in the face of uncertainty, we start to take short-cuts and a variety of factors influence our decision making.

*Irrationality: systematic deviations and biases*

Descriptive empirical work by behavioural psychologists Amos Tversky and Daniel Kahneman from the late 1960s onwards led to a paradigm shift in the understanding of JDM. They showed that that most human reasoning did not obey formal rules of inference founded on a logical calculus. Instead, they suggested that when we make decisions, rather than use a deliberative and cognitively demanding process, we often use intuitive heuristics - mental short-cuts or simple “rules-of-thumb” - that allow us to solve problems quickly, effectively shortcutting the formal cognitive processes associated with the expected utility model. Thus, the *heuristics and biases* (HB) model of JDM was born.

As their work progressed, it became clear that such intuitive decision-making came at a price; they showed that certain heuristics often led to a number of predictable errors of


\textsuperscript{234} Kirsten G. Volz and Gerd Gigerenzer 'Cognitive Processes in Decisions Under Risk are not the Same as in Decisions Under Uncertainty' (2012) 6 Front Neurosci 105.
rationality – i.e. biases. In their initial studies, Kahneman and Tversky identified the first of a later long list of biases that influence the way people intuitively assess probabilities or frequencies in various experimental scenarios (Table 1). Thus they found, for example, that our estimation of the likelihood of a future event happening is markedly affected by the ease with which we can retrieve similar events from memory - a process they labelled “the availability heuristic”, and that we tend to adhere to the first conclusions that we reach, even in the face of later contradictory information, a biasing effect they termed “anchoring”.

As theories of cognitive psychology have evolved, many other heuristics and their biased outcomes have been described in a wide variety of settings.

<table>
<thead>
<tr>
<th>Cognitive biases</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Availability bias</td>
<td>A mental shortcut that relies on immediate examples or past experiences</td>
</tr>
<tr>
<td>Anchoring</td>
<td>Tendency to place too much weight on the first information obtained or first conclusion reached</td>
</tr>
<tr>
<td>Premature closure</td>
<td>Failure to consider reasonable alternatives after making an initial judgment (diagnosis)</td>
</tr>
<tr>
<td>Confirmation bias</td>
<td>The tendency to search for, favour, and recall information that confirms pre-existing hypotheses; giving disproportionately less consideration to alternatives</td>
</tr>
</tbody>
</table>

Table 1: Common heuristics that lead to decision-making errors

235 Amos Tversky and Daniel Kahneman 'Judgment under uncertainty: Heuristics and biases' (1974) 185 Science 1124. A decision or judgment was termed “biased” if the factual information available did not normatively justify the choice made – an “error” of rationality.
236 ibid. See also: Amos Tversky and Daniel Kahneman, 'Judgements of and by representativeness' in Daniel Kahneman, P. Slovic and Amos Tversky (eds), Judgment Under Uncertainty: Heuristics and Biases (Cambridge University Press 1982) 84.
Through a series of subsequent studies, they exposed a number of factors that may distort rational judgments. They showed, for example, that individuals evaluating information and making decisions respond to irrelevant information that should have no rational bearing on the judgment, or by the way information is presented for consideration. This latter mechanism - the framing effect - is now recognised to be relevant in a wide variety of settings, from informed consent and physician decision-making in healthcare, to the semantics and possible question order in national referenda.

Despite widespread interest in their research, the ideas behind their work were not universally accepted, for their emphasis on biases and irrationality seemed to paint a very negative picture of human cognition. Prominent among such criticisms was their use of the traditional optimisation criteria founded on probability, statistics and logic as a normative standard of rationality (to define the “correct” choice or decision). This was not a new criticism.

Years before the emergence of the heuristics and biases paradigm, social scientist Herbert Simon had challenged the assumptions behind the normative models of unbounded rational

---

239 Daniel Kahneman and Amos Tversky 'Prospect Theory: An Analysis of Decision under Risk' (1979) 47 (2) Econometrica 263. This article is the most cited of all in the economic literature.

240 Dilip Soman, 'Framing, Loss Aversion, and Mental Accounting' in Derek J. Koehler and Nigel Harvey (eds), Blackwell Handbook of Judgement and Decision Making (Blackwell Publishing 2004) 379.


244 Alberto Nardelli. 'Turn of phrase: EU referendum result may come down to wording of question' The Guardian (3 June 2015).

245 In Kahneman and Tversky’s model, much of their descriptive experimental work compared subjects’ probability judgment outcomes with those predicted by Bayesian statistics, a comparison that produced significant criticism and debate. See, for example, Gerd Gigerenzer 'On narrow norms and vague heuristics: A reply to Kahneman and Tversky (1996)' (1996) 103 (3) Psychol Rev 592; see also Alex Stein 'Are people probabilistically challenged?' (2013) 111 Mich Law Rev 855.
He had argued that generally a decision-maker’s ability to make a fully rational choice must be limited, among other things, by necessarily incomplete knowledge and limited information, and by our restricted cognitive processing capacity that constrained such a challenging mental exercise. Here he introduced the concept of *bounded rationality* to acknowledge that in reality our decisions must be made within significant constraints. He posited that instead of continuing to strive, in theory indefinitely, for the perfect solution - "optimizing" or "maximizing" the outcome - individuals choose an acceptable - “good-enough” - solution to a problem, a process he termed “satisficing”.

Reflecting Simon’s views, Tversky and Kahneman’s critics argued that, unlike the restricted conditions of their behavioural laboratory, most judgments and decisions are made in much less controlled circumstances, and often with limited information. Since the conditions demanded by rational decision theory cannot be met in such circumstances, they reasoned that the “correctness” of a decision should not be judged against that normative standard of the optimisation model.

**Fast and frugal heuristics**

Gigerenzer and colleagues proposed an alternative way to consider human reasoning – the *fast-and-frugal heuristics framework* – a model that supports rather than discourages the use of heuristics. They base their model on the concept that we have a collection of specialised

---

246 Herbert A. Simon 'Theories of Decision-Making in Economics and Behavioral Science' (1959) 49 (3) Am Econ Rev 253. See also Herbert A. Simon 'Motivational and Emotional Controls of Cognition' (1967) 74 (1) Psychol Rev 29; Herbert A. Simon, *Reason in Human Affairs* (Stanford University Press Stanford, CA 1983) 128. Simon, of course, had a huge impact over a wide range of scholarship, ranging from computer science and artificial intelligence to psychology and business administration but it was for his pioneering work on decision making in economic organisations that he was awarded the 1978 Nobel Prize in Economic Sciences.


cognitive mechanisms - an “adaptive toolbox” – a collection of fast and frugal heuristics available to an individual that may be selected and applied depending on the judgment they have to make. “Fast” because we are able to make decisions rapidly using a limited amount of reasoning, and “frugal” because the process generally works reasonably well despite limited information and does not demand much cognitive effort. These heuristic mechanisms are also considered to be ecologically sensitive: evolved over time, a variety of heuristics may come into play depending on the circumstances. Thus a key feature of this model is flexibility. Some particular heuristics may be best suited to particular information environments – so that in order to achieve good decision outcomes, individuals are expected to select decision strategies (heuristics) that as far as possible match the specific situation (ecology) in which such a judgment is set. Put simply, the selected heuristic should be ecologically rational.

Let me provide a real-world example of the importance of the ecological context to heuristic selection. When NHS24 took over the responsibility for providing coverage across Scotland for the telephone triage of GP out-of-hours calls it was necessary to develop algorithms to support call handlers in directing patients to the right outcome. I was involved in the development and introduction of these algorithms.

The outcomes of such telephone triage assessments include instructions to the caller to call an emergency ambulance, see a GP within four hours, or self-care. The decision trees that underpinned this guidance were modelled, albeit indirectly, from the “fast and frugal” heuristics paradigm to guide the decisions the call handlers make.

250 Gerd Gigerenzer and Peter M. Todd, 'Fast and frugal heuristics: The adaptive toolbox' in Gerd Gigerenzer, Peter M. Todd and & the ABC Group (eds), Simple heuristics that make us smart (Oxford University Press 1999) 3.
The first responsibility of a call handler is to classify the level of illness the caller has, with the priority to discover if the call relates to potentially serious illness. Thus a short series of initial questions look for positive cues that may indicate that critical situation – conscious level, breathing difficulty, bleeding etc. This process uses the “one-good-reason” heuristic (or “take-the-best heuristic”) associated with Gigerenzer’s adaptive toolbox, and involves examining a short series of cues, then stopping when we encounter a strong or convincing cue.

These emergency screening questions work well with older children and adults. However, markers of serious illness in young children are different, so a separate algorithm with addition questions was developed for pre-school children. However, such was the concern about missing serious illness, it was decided that if calls were received about febrile children under three years, they should simply be referred for face-to-face consultation in case serious infections were missed.

Thus it can be seen that the use of such heuristics must take account of the specific situation for which they are used – they must be ecologically rational. Such tailoring to the contexts in which they are used also depends on the learning and experience of the individual, that in some circumstances may be considered expert knowledge.

While the key here is in selecting the appropriate tools (heuristics) for the decision task at hand, in many less controlled circumstances the selection process is mostly unconscious and the mechanisms underlying such selection remain unclear.

Such fast and frugal decision making approaches have been applied in a variety of healthcare settings, for example in emergency triage as noted above, and to develop decision trees to

---

252 Gerd Gigerenzer and D. G. Goldstein, ‘Betting on one good reason: The Take the Best heuristic’ in Gerd Gigerenzer, Peter M. Todd and & the ABC Group (eds), Simple heuristics that make us smart (Oxford University Press 1999) 75.
guide diagnostic decision making. However, as I discuss later, while such intuitive reasoning may work well when applied to a conventional clinical diagnostic problem solving, such an approach may be inappropriate and produce “errors” if applied in the very different ecological setting of forensic assessment.

**Duel Process theories of human cognition**

As the HB scholarship gained momentum, and variants such as the fast and frugal model fed into that descriptive scholarship, parallel work in a variety of domains explored the cognitive mechanisms underlying these various decision outcomes.

These studies suggest that generally human reasoning involves a blend of two separate cognitive processes, typified by Kahneman’s description of “thinking fast and slow”.

While the various versions of such dual-process theories of human cognition vary somewhat with their research domain, and there is not, as yet, an accepted universal theory of cognitive dual processing, the general view is that human reasoning involves a rapid intuitive or implicit process with an ancient evolutionary history, combined with a more modern (from an evolutionary perspective) slower, conscious and deliberative analytical process.

These two complementary processes are applicable within a broad range of reasoning and decision-making settings, including that involving applied expertise, so let me turn briefly to consider human cognition as it applies to expert judgment, including that applicable to clinical decision-making.

---

255 For a historical review, see, for example, Keith Frankish and Jonathan St B. T. Evans, 'The duality of mind: An historical perspective' in Jonathan St B. T. Evans and Keith Frankish (eds), *In two minds: Dual processes and beyond* (Oxford University Press 2009).
256 For a useful review, see J. S. Evans 'Dual-processing accounts of reasoning, judgment, and social cognition' (2008) 59 Annu Rev Psychol 255.
As we will see, while expert decision-making has some distinct features compared to “normal” settings, it has also much in common with the generic features I have discussed already.

Section 2 - Expert judgments

Cognitive scientists have studied how expertise is acquired and applied for over 50 years, research ranging across very diverse expert domains that includes chess grand masters, pilots, fire fighters, a variety of medical specialty areas, and even burglars. Despite this wide diversity of expertise, there is common ground on some of the key cognitive mechanisms that underlie applied expertise.

As Martire and Edmond recently noted, two apparently disparate theoretical perspectives on the development of expertise and the success (or otherwise) of expert decision-making in complex environments – the HB approach noted above, and that of the Naturalistic Decision Making (NDM) model, have gradually coalesced. Both schools of thought agree that intuitive responses dominate the (initial) process of applied expertise, but they have different perspectives on the potential implications, the HB tradition viewing intuitive responses as a potential source of decision errors, while the NDM has a more positive perception and seeks to develop intuition to maximise correct judgments.

The process of expert medical clinical diagnosis generally involves a recognition-primed model that reflects the common ground of both the HB and NDM decision-making

---

259 For a useful overview of the evolution and research perspective of NDM, see: Gary Klein ’A naturalistic decision making perspective on studying intuitive decision making’ (2015) 4 (3) J Appl Res Mem Cogn 164.
In the final part of this section I bring together these various theoretical perspectives that coalesce within current understanding of expert medical diagnostic judgments.

**Clinical expert reasoning and the diagnostic process**

[S]omething besides hard science is involved in many medical decisions, both diagnostic and therapeutic, and … these variations are not necessarily rational responses to differences between patients.

Much has been learned about the development and application of cognitive reasoning as doctors mature from medical novice to clinical expert. For expert doctors, diagnostic reasoning is arguably their most important professional responsibility, and although most experts perform very well, diagnostic error rates remain significant. In contrast to the controlled laboratories of Tversky and Kahneman, real-world clinical diagnostic decision-making usually occurs in complex and often unstructured information environments, and the diagnostic process involves significant uncertainty. In such challenging clinical settings, Bayesian (optimisation) strategies are rarely adopted, and if

---

used, are often misapplied. Rather it is now generally accepted that expert diagnostic judgments primarily involve the use of a variety of heuristics - the “fast and frugal” process that I noted previously. Here a “recognition primed” decision-making approach is used, a judgment process also notable in some non-medical expert domains. Medical experts’ diagnostic reasoning usually relies on a fast automatic (type 1) process based on pattern recognition to suggest a likely diagnosis, often despite incomplete information, the frugal element representing the ability of the clinician to satisfice acceptable probability assessments with the least information (cues) possible. For experienced clinicians, this involves recall of similar cases or the identification of particularly indicative features, focused around the concept of “illness scripts” that are stored and retrieved from long term memory.

It is generally only when faced with novel, more complex or otherwise difficult problems that clinical experts revert to a conscious cognitively demanding (type 2) analysis. Such a clinical diagnostic process is often termed “hypothetico-deductive”, and involves four key

---


269 G. Klein, R. Calderwood and A. Clinton-Cirocco 'Rapid decision making on the fire ground: The original study plus a postscript' (2010) 4 JCEDM 186.


272 Croskerry, for example, has provided an overview of the application of such dual process thinking in the clinical context of emergency medicine: Pat Croskerry ‘Clinical cognition and diagnostic error: applications of a dual process model of reasoning' (2009a) 14 (1) Adv Health Sci Educ Theory Pract 27.
elements: information gathering, hypothesis generation, hypothesis testing, and reflection. 

Here initial clinical information is used to create diagnostic hypotheses – a differential diagnosis. Subsequently information – from contextual features, the patient’s history, examination, and perhaps laboratory and other tests - is used to refine the initial hypotheses, and so reach a diagnostic conclusion.

This dual process model of clinical reasoning in making diagnoses and problem solving is illustrated in Figure 6, replicated from Croskery et al.

Figure 6: How (expert) doctors think.

One key clinical issue here is balancing the likelihood of a common benign problem with the danger of rejecting the possibility of a much less likely diagnosis that is however potentially lethal if missed. Thus diagnostic practice is not simply a likelihood judgment. This "can't

275 Pat Croskery, Geeta Singhal and Silvia Mamede 'Cognitive debiasing 1: origins of bias and theory of debiasing' (2013) 22 ((Suppl 2)) BMJ Qual Saf i58. The illustration represents a concept of the dual process model for clinical decision-making. (T represents a toggle function, allowing the decision maker to move back and forth and between Type 1 and Type 2 processes).
276 The possibility of serious bacterial infection in febrile children with particular rashes mandates the empirical use of aggressive antibiotic therapy, although many are likely to have underlying benign
rule out” effect is a major influence on initial diagnostic and treatment decisions and is a particular feature in paediatric practice where distinguishing benign from serious conditions may be particularly difficult in the early phase of an illness. Thus the diagnostic process is dynamic, where a patient’s progress, response to treatment, or subsequent test results allows revisions of initial or sequential diagnostic conclusions.

When expert clinicians make diagnostic errors, by definition they are rarely due to deficits in knowledge. Rather they are usually multifactorial, involving a combination of system failures – the effect of interruptions, fatigue and other environmental factors, and flawed judgments from individual who have failed to correctly synthesise the available information. Inconsistency in response to the same information is a well-established key theme.

Most of these flawed diagnostic judgments arise from one or more cognitive biases (type 1 failures), although analytical (type 2) reasoning does not guarantee a correct conclusion is reached. As noted in the list below, many of the heuristics underlying such diagnostic errors reflect the cognitive biases noted in all human decision-making situations, although some may be renamed to reflect more specifically the medical decision-making setting:

- Triage: this priming heuristic results when an initial diagnostic label applied by one individual is accepted without question by others


---


279 P. E. Meehl, Clinical versus statistical prediction: A theoretical analysis and a review of the evidence (University of Minnesota Press Minneapolis, MN 1954) 149.

• Availability: tendency to consider a diagnosis more likely if easily retrieved from memory

• Framing: the tendency to reach different conclusions depending on how the information is presented

• Anchoring: to quickly and firmly latch on to a single diagnosis, and ignore other possible options; this is closely related to another heuristic - Premature Closure.

• Confirmation bias: tendency to look for data to corroborate rather than refute the original conclusion

The effect of such biased reasoning will, in some cases, have grave clinical consequences for a patient; when such clinical misjudgments occur in the context of paediatric forensic assessment, they may have equally serious implications for both a child and its parents.

**A reflective pause**

In the above two sections, I have provided an overview of the evolution of, and current understanding of, the processes of human judgment and decision-making, and how that scholarship relates to clinical expert diagnostic judgments.

Although I have attempted to be brief and selective, I defend the volume and detail provided so far in this chapter because - in the context of this thesis and my attempt to expose a variety of potential threats to objective paediatric forensic judgment - I suggest that the topic that this chapter now goes on to discuss is arguably one of the most novel and important matters this thesis addresses, and so it is important that it is firmly set within an appropriate theoretical framework.

As I have discussed, as our understanding of human cognition has developed over the last 60 years or so, it has become clear that much human judgment is intuitive and may be affected by factors that, from a rational (objective) perspective, should not influence the decisions we make. Such responses seem to reflect the human condition, and our evolutionary past, and in
many circumstances such intuitive thinking can be extremely useful both in commonplace and emergency situations. The literature above also demonstrates that, like human judgments more generally, the expert clinical diagnostic process, which involves the integration of specialist knowledge with case-specific information and clinical examination, is also dependent on the use of unconscious heuristic mechanisms that are vulnerable to cognitive frailties and biases due to contextual and other factors that may result in errors of clinical judgment.

I now move on in Section 3 to discuss how the scholarship I have considered so far may be related to potential frailties in paediatric expert forensic testimony.

---

Section 3  Cognitive biases and paediatric forensic judgments

The big problem, as it is everywhere, is with unconscious bias. I dare say that we all suffer from a degree of unconscious bias, and it can occur in all sorts of manifestations. It is almost by definition an unknown unknown, and therefore extraordinarily difficult to get rid of, or even to allow for.

In this section I argue that within the particular setting of the child protection environment there are a number of specific influences that may come together to exacerbate the risk of cognitive biases when paediatricians make forensic judgments and providing related expert opinions in cases of alleged inflicted injury. Such biases may contribute to the risk of flawed expert opinions and miscarriages of justice. It is these barely acknowledged factors that I now seek to expose.

The Myth of the Unbiased Expert Witness

Experts are supposed to be impartial and experts believe they are impartial and objective. However, in reality, experts are far from being impartial and far from being objective.

As I noted at the beginning of this chapter, a series of unsafe convictions across many jurisdictions have revealed the epistemic frailty of much established forensic science. The impact of these exposed vulnerabilities on the worldwide forensic science community cannot be overstated, and in the aftermath, there have been calls for a major reform of the forensic sciences themselves and their relationship with a variety of components of the criminal justice system.

---


284 See NAS report, n. 221 above. In one study of the origin of wrongful convictions, flawed (as opposed to fraudulent) scientific evidence was a factor in more than half of the cases (Michael J. Saks 'Scientific Evidence and the Ethical Obligations of Attorneys' (2001) 49 Clev St L Rev 421.

285 See Œ O'Brien et al, 2015, n. 69 above.
One of the key problems that the current “crisis” in the forensic sciences has revealed is that forensic experts are highly vulnerable to biases arising from human factors intrinsic to interpretive judgments. An explanation for some of these errors may be found in cognitive neuroscience.

As Dror has highlighted, we do not experience the world as it really is. Rather, our brains actively process and organise what our senses record, filtering out and suppressing much information, making assumptions, and imposing order on the world around us. These active processes reflect our intrinsic cognitive mechanisms and are unrelated to the actual information we are processing. They are as much physiological as psychological but feed into the intuitive nature of much of the expert reasoning that I highlighted earlier.

While these cognitive mechanisms allow us to process our information load efficiently, such efficiency may come at a price. Because we “see” the world through the prism of our values, individual past experiences, specialist training, and professional expectations and relationships, as I discuss below such personal perspectives may make us vulnerable to the influence of a variety of other biasing effects.

In tandem with the exposure of the subjective frailties of much traditional forensic “expertise”, an evolving scholarship has emerged focused on the human factors involved

286 For a useful overview of the emerging concern about the reliability of some well established forensic science areas, see generally, Mnookin (2008), n. 58 above, and House of Commons Science and Technology Committee, 'Forensic Science On Trial' report, n. 100 above. Such reports reveal that forensic experts have, inter alia, misunderstood their role and the boundaries of their expertise, misunderstood the limitations of their techniques, misinterpreted results, and provided biased conclusions to the courts. This has ranged from the exposure of the unreliability of such iconic techniques as fingerprint identification, as exposed in the inquiry into fingerprint evidence in *HM Advocate v McKie* (1999, unreported): see: http://www.thefingerprintinquiryscotland.org.uk/inquiry/21.html (accessed 14 May 2018), to longstanding concern about the independence of scientists whose forensic laboratories were incorporated with police services, with inherent risks of context bias (M. J. Saks and others 'Context effects in forensic science: A review and application of the science of science to crime laboratory practice in the United States' (2003) 43 (2) Sci Jus 77.

in forensic science perception and interpretation – so called cognitive forensics. This has revealed that forensic interpretations are highly vulnerable to a number of cognitive frailties. In particular, the potential influence of irrelevant contextual case information has been recognised, although “[t]he traps created by such biases can be very subtle, and typically [the forensic expert] is not aware that his or her judgment is being affected.” \(^{289}\) Contextual contamination is just one of the factors that may bias forensic judgments; a variety of sources may influence expert objectivity at various stages in the decision-making process (Figure 7).

![Figure 7: Dror's taxonomy of sources of bias that may contaminate forensic judgments.](image)

As understanding of the human factors that may underlie biased forensic science interpretative judgment has evolved, a number of tools have been developed to mitigate such


\(^{289}\) NAS Report, n. 221 above, page 145.

effects. In particular, variants of *sequential unmasking* - a technique to blind forensic practitioners to extraneous information for as long as possible - have been applied in a number of forensic domains.

**Section 4 - Contextual factors and risk of bias in child abuse assessments**

In the final section of this chapter, building on current understanding of the highly intuitive nature of the expert clinical diagnostic process and the multiple sources of potential bias revealed by the evolving forensic cognitive scholarship that I described earlier, I now consider how that scholarship may apply in the context of child abuse forensic judgments.

Here I argue that there are inherent procedural features and other factors particular to clinical child abuse evaluation that, unlike some other forensic fields, unavoidably expose the paediatrician to potentially biasing contextual information and which might heighten the risk of flawed paediatric expert opinion evidence. Here I focus particularly on the potential biasing effect of commonly associated adverse psychosocial factors, and a variety of other relevant heuristics that may be triggered during the process of forensic assessment. I consider these within the context of the multi-agency investigation and the actual working processes that are followed. I show that such factors may not only influence referrals for expert evaluation but that they have the potential to shape the paediatrician’s opinion of the likelihood that a child’s injury is accidental or inflicted. This is an area of paediatric forensic work that has hitherto been largely unexplored.

---

My focus in this section is on issues that may arise primarily, but not exclusively, in direct case assessment by the paediatrician. Thus I set my argument only within the third and fourth elements towards the top of Dror’s taxonomy pyramid (Figure 7 above) – i.e. the potential impact of irrelevant case information and base rate expectations, both of which relate directly to the assessment of a paediatric forensic case.

For reasons of both clarity and logic I consider these two potential biasing factors and their likely impact on forensic interpretation within the setting and natural sequence of the child protection assessment of a child with alleged inflicted injury that I described in chapter 1. The discussion moves from referral, to parental interview, information gathering, clinical examination, then multidisciplinary investigation, and finally to expert opinion.

**Case specific factors and the forensic substantiation of child abuse**

The first issue I want to address is potential selection bias in cases referred for expert assessment and whether adverse social and other characteristics prominent in some children’s families might influence the forensic interpretation of equivocal injuries, where associations must be distinguished from causation.

**Separating associations and correlations from causation**

Referrals of a young child for clinical assessment and forensic opinion because of suspected physical abuse may arise from a number of sources. Sometimes the concern arises within a hospital setting, usually following an emergency department attendance, occasionally from an inpatient ward referral. However, the majority of expert paediatric assessments occur following referrals from a community setting – usually from a social worker based within the

---

local community service, less commonly from a health visitor (HV) or general medical practitioner (GP) in a primary medical care setting.  

Despite strong encouragement to report any concern about possible abuse, there is some empirical evidence, albeit sparse, that UK health professionals use some discretion in the cases they select to refer for child protection evaluation. Unsurprisingly, such hesitation is usually displayed when clinical features are nuanced rather than clear-cut. Similarly, in the USA and some Australian states, despite rules of mandatory reporting if the professional has “reasonable suspicion”, there is evidence that doctors and others do not always report injuries that they suspect are caused by abuse.  

A number of issues underlie such discretionary behaviour. Thus, for family doctors, anticipated damage to an established relationship with the child’s extended family, and the potential for reputational damage within the local community if their concern is

294 A. M. Kemp and others 'Bruising in children who are assessed for suspected physical abuse' (2014a) 99 (2) Arch Dis Child 108.  
296 Across the various US states legislation, the term 'reasonable suspicion' is a commonly used term to indicate the threshold for reporting possible child maltreatment. However the interpretation of that term varies widely among doctors, and it is evident that there is no true consensus on the threshold for reporting concern; see B. H. Levi and K. Crowell 'Child abuse experts disagree about the threshold for mandated reporting' (2011) 50 (4) Clin Pediatr (Phila) 321; Benjamin H. Levi, Georgia Brown and Chris Erb 'Reasonable suspicion: A pilot study of pediatric residents' (2006) 30 (4) Child Abuse Negl 345.  
unsubstantiated, are significant influences on such decisions. Even paediatricians are sometimes hesitant to raise concerns. Here diagnostic uncertainty, the time-consuming nature of child protection documentation and procedures that such referrals generate, and negative experiences in court may contribute to such hesitation.

In contrast, the presence of a variety of adverse or negative characteristics within a family setting may influence referral decisions and lower the threshold for referral despite equivocal clinical features.

**Psychosocial contextual features as an indicator of likely abuse**

In the era of “discovery” of child abuse as a social and clinical entity, Kempe claimed: “child abuse is psycho-dynamically related and has nothing to do with race, colour, creed, sex, income, education, or anything else”. However this assertion was soon challenged. It is now clear that an adverse family environment and various social risk factors are associated with an increased risk of neglect and abuse. Reported risk factors for child maltreatment include young motherhood, the child living with unrelated adults, parental domestic violence or record of violent offending, parental alcohol or drug misuse, and parental mental and

---

299 Raising such concerns in small rural communities may be particularly problematic; see James Anderst, Nancy Kellogg and Inkyung Jung 'Is the diagnosis of physical abuse changed when Child Protective Services consults a Child Abuse Pediatrics subspecialty group as a second opinion?' (2009) 33 (8) Child Abuse Negl 481, citing J. Shapiro and R. Longenecker 'Country doctors in literature: Helping medical students understand what rural practice is all about' (2005) 80 (8) Acad Med 724. Additionally, social constructs and values in some societies make raising the issue of possible child abuse particularly problematic: see, for example H. S. Habib 'Pediatrician knowledge, perception, and experience on child abuse and neglect in Saudi Arabia' (2012) 32 (3) Ann Saudi Med 236.

300 A. W. Newton and A. M. Vandeven 'Update on child maltreatment' (2007) 19 Curr Opin Pediatr 223; A. D. Theodore and D. K. Runyan 'A survey of pediatricians' attitudes and experiences with court in cases of child maltreatment' (2006) 30 (12) Child Abuse Negl 1353. There is more confidence and less reluctance to report possible physical abuse compared with emotional abuse or neglect, possibly because physical injury is regarded as a more severe form of abuse, and that their concern can be justified by “hard” physical evidence.

301 C. H. Kempe 'The Battered child and the Hospital' (1969) 44 Hosp Pract 44.

physical illness and/or learning disability. And although I do not address it further here, racial origin has also been a particular focus for child maltreatment risk stratification, notably in the USA.

Given the robust empirical evidence of a strong association between general social disadvantage and risk of abuse, especially physical abuse and neglect, it is unsurprising that in the UK most of the children who come to the attention of social work services because of possible abuse live in poverty.

Thus it seems that social disadvantage and a variety of risk factors associated with child maltreatment may lower professionals’ thresholds for concern and increase referral rates for assessment. Indeed the current NICE guideline on when to suspect child maltreatment specifically recommends that a concerned healthcare professional considering the possibility


304 In the USA, there has been longstanding concern that in some settings negative racial stereotyping underlies referral patterns (Brett Drake, Sang Moo Lee and Melissa Jonson-Reid 'Race and child maltreatment reporting: Are Blacks overrepresented?' (2009) 31 (3) Child Youth Serv Rev 309, quoting Administration on Children, Youth and Families, US Department of Health and Human Services, 'Child Maltreatment 2007' (US Government Printing Office 2009)). Certainly there is evidence that African American children are particularly overrepresented within the child welfare system, are disproportionally more likely to be the subject of alleged maltreatment referrals (Chapin Hall Center for Children, 'Understanding Racial and Ethnic Disparity in Child Welfare and Juvenile Justice', Racial and Ethnic Disparity and Disproportionality in Child Welfare and Juvenile Justice: A Compendium (University of Chicago 2008) 9; John D. Fluke and others 'Disproportionate representation of race and ethnicity in child maltreatment: investigation and victimization' (2003) 25 (5–6) Child Youth Serv Rev 359 ) and are twice as likely as white children to have concerns substantiated (Drake et al. 2009 above).

305 For a comprehensive analysis of the link between poverty and child abuse and neglect from a UK perspective, see Bywaters et al, n. 293 above. See also: Julie Wilkinson and Susannah Bowyer, The impacts of abuse and neglect on children; and comparison of different placement options: evidence review (Dept. for Education 2017).


that a child has been maltreated should explore such “collateral information” in order to judge the likelihood of abuse and so move to suspect abuse if adverse social factors are present.

However, there is arguably an element of circularity here. Parental support-seeking behaviour in relation to various family difficulties may produce a “visibility bias” – contact with social services in relation to general support needs leading to hyper-surveillance, so that even minor injuries or other concerns may provoke formal investigation, and later affect substantiation decisions.

As I discuss later, such risk stereotyping may work against vulnerable children living in family settings conventionally perceived to be “low risk”, where abuse may be missed or recognition delayed because of such false reassurance, sometimes with tragic consequences.

Given the skewed psychosocial background and adverse contextual features that characterise the selection of children who are referred for paediatric forensic evaluation, the question that arises here is whether such negative contextual factors might bias a paediatric expert’s substantiation threshold for alleged abusive injury, the origin of an injury classified as “likely” rather than “possibly” inflicted on the basis of adverse social or other family characteristics.

The important issue here is to recognise the distinction between the various psychosocial factors associated with risk of abuse and the forensic interpretation of the cause of an evident injury.

311 C. Jenny and others 'Analysis of missed cases of abusive head trauma' (1999) 281 (7) JAMA 621.
312 NICE guideline 89, n. 309 above, at page 11.
injury in an individual case. At an individual level risk factors have a very low specificity for abuse, and while there is an established association between poverty and risk of abuse and neglect, most children from socioeconomically deprived backgrounds are not abused or neglected. Rather, accidental injuries are more common in children from deprived households.

Thus, this emphasises that environmental risk factors for abuse need to be treated with great caution in considering the origin of an injury in a child. Incorporating such risk factors into the forensic decision-making process of an individual case would represent a bias based on a base rate fallacy.

---


Paediatric assessment of possible inflicted injury

There are instances where the biasing information is so inherent and integrated into the actual evidence that it is practically impossible to separate them. In her summary of the paediatric diagnostic process in assessing alleged physical abuse, Dr Hilary Cass, then president of the RCPCH, noted that “corroborative detail” was an important element that the paediatrician should use in deciding whether an injury was abusive or not. She did not define this further, but RCPCH guidance specifies a number of specific features that should be considered and recorded.

Such information gathering generally involves three overlapping areas:

1. The history from the relevant professional of the concern that has led to the medical assessment, often provided by an accompanying social worker and/or police officer.
2. The background health and social history and family structure, including details of employment and health of parents/carers. This involves “specific questions … regarding domestic violence, parental drug or alcohol problems, parental mental health difficulties, [and] parental learning difficulties”.
3. A detailed verbatim record of the events leading to the injury, and the parents’ actions thereafter. This is expected to include consideration of any discrepancy in the accounts given, and the plausibility of these accounts in relation to the clinical findings.

Even when dealing with an apparently accidently injured child, there is a mandatory requirement for paediatricians to explore other previous hospital attendances, whether the

---

316 Dror, n. 290 above, page 542.
317 Cass (2014), n. 12 above.
319 ibid, at para. 6.7.2.3.
320 ibid, at para 6.7.1.4. The use of the term plausible here seems somewhat pejorative, bringing with it an intrinsic element of doubt.
child is recorded on the child protection register and, depending on the age of the child and
the nature of the injury, to make formal contact with primary care and social work services to
explore the family social and health background and other relevant issues.

Such information gathering and sharing is a core function of professionals involved in child
protection procedures, as reinforced by the conclusions of many serious case reviews
(SCRs), including not only that in relation to Declan Hainey, but many other high profile
Scottish cases, such as the deaths of Caleb Ness, Brandon Muir, and most recently, Liam Fee.

This process obviously exposes the paediatrician to potentially biasing but unavoidable
contextual case information, given such information sharing is intrinsic to the overall
safeguarding process such cases activate. Thus, potential blinding processes adopted in some
other forensic domains are not practicable in this setting.

Potential cognitive biases and the related intuitive factors in paediatric forensic assessment
are best appreciated in a real-world context, so before I address some of these in more detail,
let me provide brief details of a personal real-life case involving forensic diagnostic
uncertainty to illustrate some of the issues.

321 The Determination of the FAI in relation to Declan’s death is available at:
http://www.scotcourts.gov.uk/search-judgments/judgment?id=bc8147a6-8980-69d2-b500-
ff0000d74aa7 (accessed 11 May 2018).
11 May 2018).
325 Gary Edmond and others 'Contextual bias and cross-contamination in the forensic sciences: the
corrosive implications for investigations, plea bargains, trials and appeals' (2015) 14 (1) Law, Prob &
Ris 1.
Clinical vignette:

An 11-month-old girl attended the Emergency Department because of parental concern about an apparently painful swollen right lower leg. Her past medical history was unremarkable. She was the first child of unmarried middle-class parents, each in their mid 30s, who were both employed although the mother was on maternity leave. The parents had never lived together and the mother was the main carer, but the biological father was very involved and saw the child most days. There was no history of any injury.

When examined, the baby was systemically well without fever. She was evidently generally healthy, clean and thriving, and was very socially interactive. Examination showed a swollen and slightly bruised right lower leg and ankle that was evidently tender to touch. She would not weight bear.

Unexpectedly, X-rays showed recent fractures of both lower tibiae (shin bones), the right worse than the left. There were no signs of bone disease (nor a family history of such). She was examined by the orthopaedic specialist who suggested that the fractures were recent (within 10 days), and represented “compression fractures due to significant upward force from below. She has either been dropped or banged hard onto her feet from a height, or been forcefully struck on her feet, causing the bones to compress in a concertina fashion.” In the absence of any history of injury, a referral was made to me for child abuse assessment and I admitted the mother and baby to the hospital.

Her parents were anxious and distressed but co-operative fully with a subsequent child protection investigation. Over the following days the father spent many hours with his daughter and her mother at the hospital. He was attentive and clearly very involved in routine care, often feeding or changing the baby. No other injuries were found on skeletal
survey or brain imaging, and there was no radiological or laboratory evidence of bone fragility.

The multidisciplinary investigation did not reveal any professional concern. The Health Visitor was entirely happy with the care of the baby and neither relevant family doctor had any concerns about either parent. There had been no previous social work involvement and there were no financial or other worries. The Police interviewed each parent separately, noting a detailed timeline over the previous 10 days when one or both had been caring for the child. Neither parent had a criminal record. The parents co-operated fully with all agencies.

The fractures remained unexplained and a child protection case conference was planned.

I saw the parents on a number of occasions. In order to try to tie down the timing of the injury I asked about times when the baby might have been crying unexpectedly or for no apparent reason. Although the baby could be quite unsettled because of “colic”, they commented that on one recent day in particular the baby had begun crying persistently while in her baby seat in the mother’s car. They thought she had been frightened by a bang when the father, a tall man, had adjusted the driver’s seat position. I wondered if, sitting with extended legs, she might have been struck by the car seat. A Police mechanic inspected the car and found that the driver’s seat was faulty, allowing it to track back too far against the rear seat.

This accidental explanation was accepted at the subsequent case conference. So, an unusual case with a happy outcome. But let me modify the story a little, leaving the clinical features unchanged.

Now the details of the family are a little different. The parents are unmarried but live together. They are a little younger, the mother 18 and the father 24. The parents are both
currently unemployed. It is the mother’s first child but the father has a three-year-old son from a previous relationship, although there is little contact. When told about the concern and need for admission because of child protection concerns, the mother is very tearful. The father becomes angry and verbally aggressive.

In the ward, the mother is quiet and withdrawn but provides all the routine care for the child, the father visiting intermittently. The nurses report a parental argument outside the ward.

The Health Visitor has been seeing the family on a fairly regular basis because the mother has had moderate postnatal depression; she has had no concerns about the care of the child, although the house is usually untidy, and she has had to advise both parents to smoke outside. The Social Work department have had previous contact with the father in relation to child maintenance payments, and more recently with the family in relation to managing rent arrears. The Police report that the father has a previous conviction in relation to a drunken brawl in the street some six years previously.

So, we have the same clinical findings but a somewhat different atmosphere.

Let me now turn to dissect the potential biases lurking in the undergrowth of such contrasting scenarios in a little more detail. These could be considered to lie within three relatively distinct areas within an overall case narrative of a paediatric forensic assessment – the potential effect of negative social and other extraneous factors, the inherent risks of bias that may arise within the clinical process itself, and finally the likely impact of the family’s engagement and degree of co-operation with the forensic investigation.

As we will see, it is not uncommon for multiple cognitive biases to be at play in such situations, and many of these biases overlap.

326 In this immediate section I do not intend to reference individual cognitive and affective biases – they are too numerous. Rather I refer the reader to a recent paper listing more than 100 such cognitive biases: Milos Jenicek, Medical Error and Harm: Understanding, Prevention, and Control (Productivity Press New York 2011) 384.
**Priming and contextual biases from extraneous information or social characteristics**

As part of the referral or during the assessment, as detailed above, information becomes routinely available in relation to family structure and demographic details, and history of any previous social work involvement, together with “relevant” parental health history and any other concerns the referrer deems appropriate. As the second fictional vignette illustrates, such information may well contain a variety of negative or “concerning” features in the family history, here a history of minor violent offending, and professional concerns about parental mental health and parenting ability.

Such negative framing of case information is common, such family characteristics often considered indicators that a family is “vulnerable”, and thus associated with increased risk of child abuse. This effect may be reinforced by social workers comments, such as “we have been worried about this family for some time”. Similarly, the police, often involved initially in a parallel criminal enquiry, may offer information about family members’ criminal records, or use prejudicial terminology – a parent described as a “suspect” – that may potentially bias medical judgments.

Yet while much of this information may be relevant to a broader multidisciplinary safeguarding risk assessment, it should have no bearing on the clinical interpretation of a physical injury. On the other hand, it often means that even before a child is examined the paediatrician’s diagnostic objectivity may be undermined, primed with negative extraneous, contextual information that may be highly prejudicial to the subsequent examination and objective interpretation of injuries.

---

Such negative contextual information is well known to influence expert interpretation in other forensic settings. This negative stereotyping may mean that the paediatrician may become anchored to a pre-examination view that there is an increased likelihood of abusive injury, so that an equivocal injury may be interpreted as likely non-accidental because of a shift in clinical “certainty” stemming from a perception of ‘high risk’ based primarily on social and demographic factors.

There is good empirical evidence that paediatricians do take account of such family ecological factors, particularly race and socio-economic status, in their forensic assessment and subsequent diagnosis of non-accidental injury, and, more importantly, that such features impact significantly on their clinical decision-making. This may have counterintuitive consequences. Thus, in a report of the assessment of young children with head injuries, US doctors were more likely to fail to recognise abusive head injuries in white children compared to black children with the same presenting features, resulting in some potentially avoidable (white) child deaths.

Similarly, and somewhat incongruously, such reverse stereotyping underpinned the flawed expert testimony of Professor Sir Roy Meadow in Clark, when he argued that the absence of adverse socio-economic and other unfavourable family characteristics strongly mitigated...
against the possibility that the deaths of the Clark babies could be attributed to SIDS, a condition strongly associated with smoking and deprivation.

_Biases within the healthcare setting_

Once the face-to-face clinical assessment begins, other biasing factors may come into play.

As I have previously noted, before a parent and child are seen and the child examined, typically there is a discussion between the paediatrician and the professional making the referral, explaining the circumstances that led to the concern.

In the case vignette above, I was primed (i.e. explicitly directed) by another experienced paediatric professional – a senior orthopaedic specialist - to consider a diagnosis of abusive injury. Also known as _triage cuing_, such priming is well known to influence clinical diagnostic decisions more generally. And once a clinical team becomes _anchored_ to a patient’s diagnostic label it can be hard to remove it; indeed, it may become “stickier”, an effect known as _diagnostic momentum_, for as Professor Jo Delahunty QC, has observed:

> The hospital workplace is not immune from suspicions being aired between stressed colleagues as soon as Non-Accidental Injury is raised as a possible cause [of an injury]. Corridor discussions inevitably take place (unrecorded) where opinions are formed as evidence is emerging and facts are unclear.

This effect links to both _ascertainment_ and _confirmation biases_, where one’s thinking is influenced by prior expectations or other’s conclusions, so that one then sees what one expects to see, and selectively seeks other information to confirm that initial interpretation.

This overlaps with another potential bias - _selective attention_, where we narrow the focus of our interest. This was perhaps most famously demonstrated in Chabris and Simons’ famous

---

331 There is an established strong association between socio-economic disadvantage (and the various adverse factors within that term) and the risk of Sudden Infant Death Syndrome in those families; see N. Spencer and S. Logan ‘Sudden unexpected death in infancy and socioeconomic status: a systematic review’ (2004) 58 (5) J Epidemiol Community Health 366.


gorilla experiment. We do not see the gorilla because we are not looking for gorillas, unconsciously filtering out information we consider irrelevant. If you are a forensic paediatrician, your focus is child abuse – that is what you are searching for. And once you think you have found it, search satisficing — a form of premature closure — may take over; alternative explanations may be downplayed, or contradictory information simply ignored (confirmation bias again).

In the vignettes above, one explicit issue that emerges is when it was appropriate to stop looking for an alternative innocent explanation for the baby’s leg injuries and instead to conclude that, in the absence of an alternative explanation – medical explanation or natural injury - an abusive injury must be the likeliest option. The satisficing here may well be influenced by the clinically irrelevant social factors evident in the contrasting information I provided above.

A question that naturally arises here is how might the expert recognise and override such intuitive responses and avoid such biases. A variety of de-biasing strategies have been proposed to minimise such risks. I discuss these in Chapter 8.

335 Barbara Probst, 'Why This Book Is Needed' in Barbara Probst (ed), Critical Thinking in Clinical Assessment and Diagnosis (Springer International Publishing 2015) 1, page 31.
Of the possible and the plausible

As noted above, a crucial issue in the forensic assessment of a young child with physical injuries is for the paediatrician to consider any injury history provided by parents or carers and to judge how such proffered explanations might correlate with the clinical findings. Making sense of such explanations demands care. Apart from making a biomechanical judgment of the various forces involved in a particular injury and how they might correlate with the parental history, and considering how any described accidental event correlates with the developmental stage of the child, a broader element of expert judgment is to assess the overall plausibility of any explanation provided.

The accounts provided in such circumstances vary widely, from the absent – “I don’t know what happened”, via speculation – “I was in the other room, but I think his big brother hit him”, to a clearly witnessed account – “he rolled off the bed when I turned ’round to get a nappy. His grandmother was with us”. Generally, an absent history associated with clear evidence of trauma has been regarded as a strong indicator of inflicted injury, while inconsistent or evolving explanations are generally regarded with suspicion in the initial assessment phase or during judicial fact finding.

Yet while implausibility and inconsistency in parental histories are generally advanced as part of the diagnostic indicators of abuse, it is uncertain how firmly such interpretations may be justified. In the context of some particular injuries, such as alleged shaken baby cases (see


338 For example, multiple bruising is very rarely found in a non-ambulant child in the absence of an underlying blood disorder (Naomi Sugar F., James Taylor A. and Kenneth Feldman W. 'Bruises in Infants and Toddlers Those Who Don't Cruise Rarely Bruise' (1999) 153 Arch Pediatr Adolesc Med 399).

339 J. Hettler and D. S. Greene 'Can the initial history predict whether a child with a head injury has been abused?' (2003) 111 (3) Pediatrics 602.

340 See, for example, Re W (A Child) [2013] EWCA Civ 662, 2013 WL 1841748.
Chapter 7), such an interpretation raises the issue of “circularity” in reasoning, a theme I return to repeatedly in later chapters. The broader implication here is that a parent saying “I do not know” may be interpreted as really signifying “I am hiding something”.

And sometimes what a parent says may emerge as an apparent crucial piece of “evidence”: the “gentle shake” to stimulate a child found unresponsive and not breathing may morph into a “confession” of inflicted injury.

Obviously, putative published correlations between the spectrum of explanations and a likelihood of abuse need to be viewed with caution. It may be, for example, that an absent or changing history provokes a more aggressive search for other injuries, thus revealing additional covert trauma, while cases judged to have a plausible initial history (even if the child was factually abused) are not subjected to more extensive radiological or other investigation, and so abuse is missed.

Again, however, like the issue of psychosocial risk factors I discussed earlier, we are in difficult territory here, where the risks of subjective influences come into play. In judging the plausibility of the explanation for an injury, the paediatrician (unconsciously) applies the simulation heuristic, constructing a mental model of the scenario, the ease with which the

---

343 In a recent study of infantile forearm fractures from the USA, the authors reported a strong association between absent and changing history of injury and a final abuse diagnosis, although the original injuries were indistinguishable. Of note, there was a marked preponderance of black families in the “abuse” group, suggesting a degree of racial bias in their selection for further investigation; see Elizabeth Ryznar, Norell Rosado and Emalee G. Flaherty 'Understanding forearm fractures in young children: Abuse or not abuse?' (2015) 47 Child Abuse Negl 132.
model can be mentally constructed determining a subjective probability judgment for the event.

And, by common experience, when others tell us a story, we do not simply judge its veracity by the content, but by the way in which it is told. Inherent to the quality of communication in such interactions is a degree of rapport and mutual understanding on both sides of the exchange. Not surprisingly, in the context of forensic child abuse assessment, dialogues with accompanying carers may be problematic.

The potentially biasing effect of parental demeanour during the consultation

It seems trite to observe that for many parents a child protection investigation is an intimidating and highly stressful experience. Already labelled as “suspected child abuse” before their attendance, from the point of arrival they are treated differently from other parents accompanying an injured or sick child to hospital. If not accompanied by a social worker, then a nurse is usually assigned to stay with the child and family at all times. This is not just to provide support during the process; this individual also has a protection and surveillance role, part of which is to note and report any unusual interaction between the child and accompanying carer. Inevitably parents find the atmosphere of such assessments uncomfortable, and for some the process may be overwhelming.

As part of any paediatric medical consultation, it is also common to consider the parent or carer’s demeanour, mood and interaction with the child as part of an overall clinical evaluation. While this is infrequently formally recorded in routine child health consultations

---


except when such interactions are atypical, a record of parental demeanour and interaction is mandated as part of the standard child protection medical assessment documentation. However, I suggest that the utility of such observations in the context of a forensic assessment is uncertain and potentially misleading as an indicative marker of likely abuse.

Given the inevitably stressful and often frightening experience that parents go through, it is hardly surprising that during child protection forensic consultations such interactions are often characterised by parental distress, anger and defensiveness. Such responses might be expected irrespective of their possible involvement in inflicted injury, but parental reactions, and their ability to communicate effectively, may significantly influence the paediatrician’s ultimate judgment of the likely origin of injuries.

As Laskey has described in a realistic vignette, a “difficult” mother who accompanies her child to hospital following social work concern about bruising and who is increasingly hostile during interview and assessment may significantly influence the paediatrician’s perception of the likelihood that an injury is due to abuse. Assessment of the same injury in a child accompanied by a co-operative articulate middle-class mother who appears responsive to her child’s needs during the assessment may produce a very different conclusion, perhaps reflecting our tendency to associative bias.

346 It has been suggested that some specific features in parent-child observed during a consultation may be an indication of neglect (A. M. Naughton and others “Mind the Gap” - Observable features of parent-child interactions amongst neglected children aged 0 to 13 years’ (2014) 99 (Suppl 1) Arch Dis Child A72).
347 Child Protection Companion, n. 318 above, para 6.7.1.4.
349 A non-abusing parent may be frightened or ashamed that their child has been injured, and there is a risk that such a “guilt” reaction may be misinterpreted; see Safeguarding Children Toolkit for General Practice (RCGP/NSPCC), at section 4.2.
350 Laskey, n. 222 above.
351 A positive bias that develops when an individual recognises they share beliefs, interests, or experiences with another.
Thus, caution is required in interpreting parental behavioural features during forensic consultations. Can we really distinguish apparent remorse from grief? Does parental anger and verbal aggression indicate an individual with a short emotional fuse? Why is this parent quiet and apparently unconcerned? Here negative interpretations may reflect a *fundamental attribution error* (correspondence bias). The question surely is not whether but why would they not behave in this way, for in the context of the inevitably highly threatening process of a child protection investigation, the emotional characteristics of a “fight, flight, or freeze” reactions are arguably quite appropriate.

Despite these caveats, evidence of clinicians’ interpretation of parental reactions and interaction in the context of a seriously injured child has been adduced in prosecutions in the Scottish courts, as noted in *Hainey*, and subsequently in *Younas*.

And, of course, while there is explicit emphasis on the importance of the parental/carer interview in assessing the likely origins of an injury, the utility of such interactions may be undermined when the paediatrician holds (conscious or unconscious) *a priori* beliefs about the nature of the case that may shape the interview process and distort perceptions. Thus, a presumption that a parent inflicted the injury may cause the parent to appear more defensive during interview (therefore “guilty”), such perceptions then reflected in a failure to gather

---


354 Younas v HMA [2014] HCJAC 114. Faisal Younas, appealed his conviction for the culpable homicide of his infant daughter by shaking her. The High Court rejected his appeal that the trial judge had misdirected the jury in relation to expert medical evidence. Of more relevance here, at his trial treating medical staff had been asked detailed questions about the demeanour of the child’s parents at the time of her admission, and about other subtle interactions – glances and other communications - between the parents, features which the prosecution suggested indicated the guilt of the accused [paras 8 – 10]. The trial judge had directed the jury to ignore some of this evidence, but not all of it.

disconfirmatory information, or to a selective over-emphasis of “suspicious” features (confirmation bias). \(^{356}\)

What I hope I have illustrated above is that from the initial stages of a child protection case, the assessment environment is rich in cues that may provoke a number of intuitive responses in the paediatrician during clinical evaluation, \(^{357}\) inducing cognitive biases that may potentially influence the objectivity of the paediatric forensic opinion. Naturally, the impact of such influences will vary with the specific details of the individual case.

**Conclusion**

In this chapter I began by briefly reviewing the evolution of current understanding of human judgment and decision-making, and the recognition that human reasoning involves a blend of two cognitive processes - a rapid, efficient and unconscious process dependent on heuristics, and a much slower and cognitively demanding deliberative process that tends to be recruited when we encounter complex or unusual judgment tasks.

I also discussed current understanding of expert decision-making, in particular the diagnostic process applied by medical experts, again emphasising the dominant dependency on unconscious heuristics on which expert judgments are often made. I noted that, despite its efficiency, an extensive scholarship also shows that a wide range of factors may predictably influence (bias) the conclusions associated with the use of such heuristic-dependent processes.

I moved on to acknowledge that many of these same heuristic processes are also now known to be involved in the application of forensic expertise, and that it is generally accepted that

---

\(^{356}\) For a useful overview of confirmation bias in relation to forensic interviews, see Jillian R. Rivard. ‘Confirmation bias in witness interviewing: Can interviewers ignore their preconceptions?’ (PhD, Florida International University 2014).

\(^{357}\) See E J Reese, ‘Techniques for Mitigating Cognitive Biases in Fingerprint Identification’ (2012) 59 UCLA L Rev 1252, note 60, in relation the various terms that are generally used for such cognitive biases in forensic settings.
many resulting cognitive biases permeate most forensic domains. In particular, I noted that irrelevant contextual information is a particularly prominent source of forensic bias.

Recognising the burgeoning scholarship of cognitive forensics, I explored the possibility that paediatricians and associated expert practitioners working in the field of clinical child protection might also be vulnerable to such biases. I argued that, while largely unacknowledged by paediatricians, a multitude of potential cognitive biases might influence the diagnostic judgments involved in forensic clinical child abuse assessments and thus undermine the opinion evidence provided by paediatricians in such cases.

I noted in particular that the information-sharing and multidisciplinary nature of the safeguarding process and the environment within which such judgments are made are a rich source of contextual and other cues that may induce negative and prejudicial perceptions that influence objective interpretation of the causation of injuries, generally producing a risk of over-diagnosis of abuse in the face of diagnostic uncertainty. I also highlighted the risk that heuristic responses to the demeanour and co-operation of parents/carers during such investigations may additionally contaminate expert objectivity, and lead to subjectivity and bias in such professional judgments.

Clearly, I am not immune from such effects, so, in the context of these observations, let me return to consider my contrasting vignette of my own real and imagined cases.

I spent a considerable time with the couple in the real case; they were co-operative and understanding of the responsibilities I had to meet. Certainly, they were anxious and emotional at times, but never angry or obstructive. I now wonder whether my social identification with them and my empathy with their situation led me to persist with the search for an ultimately exculpatory explanation.

On the other hand, I admit I am uncertain whether I would have spent as much time with the fictitious second couple. “Difficult” patients, or their families, inevitably raise negative
feelings that may disrupt communication and distort objectivity and seriously inhibit the empathy that is said to be at the core of the therapeutic relationship (see next chapter).

Dealing with this father, in particular, is likely to have been an uncomfortable experience, and in the face of frank hostility I suspect I would not have obtained the clues that ultimately became available in the real interviews. Rather I imagine that I would have made unconscious and unwarranted inferences about the characteristics of that individual’s personality and likely behaviour – that an angry uncooperative parent is aggressive (so more likely to hit their child) or trying to hide something (so is likely to have caused the injury), whereas in reality his behaviour may well be determined by situational factors – a stress reaction and feelings of helplessness in the face of a frightening multiagency investigation.

However, this raises a separate important issue – the affective (emotional) responses that are induced in professionals by encounters with a very ill or difficult patient (or parent) and how those responses influence our judgment and decision-making. This brings in another potential bias that Laskey’s contrasting scenarios above also illustrate – the affect heuristic, where our visceral (psychological) reaction to another person may influence our decisions in relation to them.  

As I noted above, in recent years there has been a developing scholarship on the how doctors think and make diagnostic and other clinical judgments. This work acknowledges that a clinician’s negative or positive reactions to a patient may influence how that patient is treated. It has also begun to be recognised more broadly that how doctors feel – their

emotional state and other affective issues – also influences their clinical interpretations and judgments.\textsuperscript{361} We know now that when professional healthcare workers make decisions, the cognitive and affective aspects of such judgments are inseparable; indeed clinical decisions are less likely to be in the patient’s best interests if a purely cognitive approach is adopted.\textsuperscript{362}

Arguably in a child protection setting an additional potential source of bias is the emotionally charged nature of the circumstances inherent in such cases. This is an issue that I address in the next chapter – the potential biasing impact of emotion and empathy on the objectivity of paediatric forensic expert opinions.

\textsuperscript{361} Jane Heyhoe and others 'The role of emotion in patient safety: Are we brave enough to scratch beneath the surface?' (2016) 109 (2) J R Soc Med 52.
\textsuperscript{362} Jodi Halpern, From detached concern to empathy: humanizing medical practice (OUP New York 2001) 165.
Chapter 4: Empathy and emotion – a threat to objective paediatric forensic interpretation?

Introduction

In the spring of 2005 I attended the annual conference of the RCPCH at York University. Unusually, the Law and Ethics subspecialty session after lunch was highly oversubscribed. Despite the overflowing aisles, the atmosphere was strangely subdued and a tense silence descended over the room as the invited speaker arrived. In the aftermath of Clark, and the public vilification of Professor Sir Roy Meadow, Baroness Helena Kennedy had come to present her recently published guideline on the multi-agency protocol for the investigation of unexpected infant deaths.\(^{363}\)

The Kennedy report and related protocol in relation to SUDI investigation, recently updated,\(^{364}\) had followed the series of miscarriages of justice in 2003 involving mothers with multiple infant deaths – Clark, Cannings, Patel - and the subsequent media storm, parliamentary criticism and regulatory response that I discussed in Chapter 1.

While she was treated with respect, there was little warmth, and the questions from the delegates were polite but unusually challenging. Professor Terence Stephenson, later to become President of the RCPCH, and at the time of writing current chair of the GMC, asked a question that went to the heart of the mood among the delegates: “do you think that Sally Clark killed her babies?” Baroness Kennedy’s reply was brief and technical: “her conviction was quashed by the Court of Appeal”. Mutterings of disquiet permeated the room; paediatricians were angry.

That day in York paediatricians were not just feeling indignant and defensive about the humiliation of their past College President. Their anger and hurt had a broader basis. Some highlighted that in Clark none of the pathologists had suggested the boys’ deaths were due to

\(^{363}\) Baroness Kennedy, n. 131 above.

SIDS, while other informal delegate conversations focused around worrying autopsy findings that could be interpreted as indicators of injury. There was a sense of injustice on behalf of the dead infants, that the “system” had failed them.

In the fallout of these cases many felt that one of the core responsibilities of a paediatrician – a duty to advocate for and seek to protect children - was being undermined. This was compounded by an apparent lack of understanding by the GMC of these professional imperatives.

The collective mood and raw emotions evident at that meeting over a decade ago echoes personal reactions I have observed in individual clinical colleagues working in child protection. This has sometimes involved explicit distress and anger at apparent abusive injuries in a child during the initial assessment stage of a case. At other times there have been expressions of indignation and a sense of injustice that the outcome of subsequent legal proceedings has not reflected the severity of injuries observed.

My personal observations have been mirrored in the single relevant empirical study that I have been able to find. Here, paediatric expert witnesses involved in child abuse cases in the criminal and family courts in New Zealand acknowledged sometimes feeling angry when dealing with a severely injured child and a “strong duty to go along [to court] and try and make someone answer for what they did”. Yet despite these feelings, the witnesses in that study generally acknowledged that their duty is to provide impartial evidence to the court, and suggested that they were able to separate such emotional reactions from their objective assessment of the forensic evidence and the opinions they provided.

---

367 Henderson and Seymour, n. 62 above.
368 ibid, at page 53.
369 ibid, at page 60.
In this chapter I challenge that assertion and argue that reactive emotional factors are potentially very likely to have a negative epistemic effect on the forensic opinions of paediatricians in some child abuse cases. Rather, separate from the risk of potential biases from contextual factors I discussed previously in Chapter 3, I suggest that the emotional and empathic responses that may be evoked in paediatricians assessing injuries in allegedly abused children are very likely to undermine the objectivity that expert forensic opinion demands. This is an area of scholarship that is largely unexplored.

This chapter is presented in three sections.

In Section 1, I set my argument on the largely uncontroversial foundation that man is fundamentally an intrinsically emotional creature, with an advanced emotional repertoire that reflects the evolutionary pinnacle we have reached. Here, in particular, I provide a brief and selective overview of current understanding of the construct of empathy as a fundamental feature of being human. Empathy, albeit a multifaceted and complex construct, is a particular focus for the discussion because, in the context of paediatric forensic assessment, within that shared understanding and affective response to another’s plight may lie the potential for expert error, bias, and subjectivity.

Having provided that foundation, in Section 2, I consider the concept and application of empathy in the specific context of healthcare generally, and then in the particular setting of paediatric clinical practice.

Finally, in Section 3, set against an evolving scholarship examining the influence of emotions on human decision making, I argue that there is a plausible link between paediatricians’ empathic responses to a seriously injured child and the potential for induced bias in forensic assessment, tracing a likely path whereby a compassionate response to an

---

injured child may undermine a paediatrician’s objective judgments and the associated expert forensic opinions they provide.

So first I set such my argument in its broader neurobiological context.

**Section 1 – Empathy as an intrinsic aspect of being human**

The recognition that we humans are a highly social species is long established. Our capacity to communicate and share information is one of the fundamental features that sets us apart from other animals and explains our dominant evolutionary position. Much of our communication is non-verbal. Part of this interaction includes our ability to infer others’ emotional states and resonate with them - an empathic phenomenon: ‘…where another person’s mental state (which include intentions, desires, emotions and proprioceptive states) automatically activates in the observer a representation of the observed state’.

This effect explains why, for example, certain emotions such as crying and laughter are socially “infectious”.

In evolutionary terms, such an empathic response may occur at a very primitive level, typified by emotional contagion – a precursor of empathy. This feature, likely to have evolved from the survival value of primitive group alarm signalling, is seen, for example, when one young infant’s crying sets off a cascade of crying in other babies in the vicinity.

---

372 Here proprioceptive states refer the somatic and other physiological changes related to underlying emotional states, perhaps typified by body posture and facial expression.
a feature that is also evident in other species. While such primitive emotional responses do not completely disappear, mirroring the evolutionary journey we humans have made over millennia, as infancy progresses through childhood to adulthood, a much more subtle human-specific capacity to recognise and respond to the emotional state and behaviours of others develops. This capacity, conceptualised within the Theory of Mind (ToM) (social understanding) paradigm, is based on self-awareness, self-consciousness and ability to empathise with others. It refers to “the ability of an individual to make inferences about what others may be thinking or feeling and to predict what they may do in a given situation based on those inferences”. Some argue that this ability provides the basis for the collaboration and cooperation on which our morality is built.

Certainly understanding that a fundamental (and beneficial) feature of being human involves recognising and responding to another’s situation or emotional state can be traced back at least to the Enlightenment, and the Eighteenth Century writings of David Hume and Adam

---

378 For example, recognition of “negative” facial expressions of surprise, fear, disgust or anger are universally recognised and may be useful; see Rachael E. Jack, Oliver G. B. Garrod and Philippe G. Schyns 'Dynamic Facial Expressions of Emotion Transmit an Evolving Hierarchy of Signals over Time' (2014) 24 (2) Curr Biol 187. It is also likely that primitive emotional contagion may underlie certain crowd behaviours in mature individuals: M. Lhommet, D. Lourdeaux and J. P. Barthès. 'Never Alone in the Crowd: A Microscopic Crowd Model Based on Emotional Contagion' (2011 IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology 2011) 89.
379 Functional MRI and other studies have gradually exposed the neurobiological basis for such development; see, for example, J. Decety and K. J. Michalska 'Neurodevelopmental changes in the circuits underlying empathy and sympathy from childhood to adulthood' (2010) 13 (6) Dev Sci 886; Jean Decety, Kalina Michalska and Y. Akitsuki 'Who caused the pain? An fMRI investigation of empathy and intentionality in children' (2008) 46 Neuropsychologica 2607.
383 See, for example, Michael Tomasello 'The ultra-social animal' (2014) 44 (3) Eur J Soc Psychol 187. Tomasello argues that it is the shared intentionality intrinsic to the ultra-social nature of human society that separates us from other species and is the basis for human morality.
These philosophers, reflecting the moral sentimentalist perspective of the time, developed very similar (albeit individual) accounts of the natural principles that govern human morality, which recognised the prosocial utilitarian benefits that arise when we (unconsciously) identify and are affected by another’s attitudes and feelings. As Adam Smith observed in 1759, in his introduction to *The Theory of Moral Sentiments*:

> How selfish soever man may be supposed, there are evidently some principles in his nature, which interest him in the fortune of others, and render their happiness necessary to him, though he derives nothing from it except the pleasure of seeing it. Of this kind is pity or compassion, the emotion that we feel for the misery of others, when we either see it, or are made to conceive it in a very lively manner.

Both Hume and Smith used the term *sympathy* to describe the process underlying how we develop such fellow feeling. However, as Hume’s particular description of this affective response reveals, in identifying that we not only intuitively tune in to and are able to interpret the emotional state (sentiments) of others but react to them – he seems to describe what we would nowadays term *empathy*, a term that would only enter the English lexicon some one hundred and fifty years later.

---

384 It is not my intention here to provide a detailed overview of this topic. There has been a great deal of comment and analysis of Hume and Smith’s concepts of sympathy in relation to the development of their ideas that our affective reactions to others and how we interpret them are the foundation for the moral judgments we make. While they differed somewhat in their interpretations of the mechanisms involved, they were in fundamental agreement about the overall principle. For a brief and digestible overview for a non-specialist (like myself), see: Geoffrey Sayre-McCord 'Hume and Smith on Sympathy, Approbation, and Moral Judgment' (2013) 30 (1-2) Soc Philos Policy 208.


386 Adam Smith, *Theory of Moral Sentiments, or An Essay towards An Analysis of the Principles by which Men naturally judge concerning the Conduct and Character, first of their Neighbours, and afterwards of themselves, to which is added a Dissertation on the Origin of Languages.* (Norton and Company New York 1759/1986).

387 Later, their views would diverge, Hume abandoning his view of sympathy as the key concept morality in favour of utility, while Smith strongly adhered to it. See F. L. Van Holthoon 'Adam Smith and David Hume: with Sympathy' (1993) 5 (1) Utilitas 35.


389 The term was coined by the American psychologist Edward Titchner, when he (mis)translated the German word “*Einfühlung*” [“feeling into”], originally used in relation to aesthetic interpretation of art by Robert Vischer, then aesthetics more generally by Theodor Lipp; see G. Jahoda 'Theodor Lipps and the shift from "sympathy" to "empathy"’ (2005) 41 (2) J Hist Behav Sci 151.
Hume’s concept of sympathy also accommodated both a positive or negative valence – “to rejoice with the joy of others, and to feel grief when they are unhappy”, while our modern construct of the term sympathy is generally focused on the negative.

As I describe below, such definitional ambiguities represent a significant challenge to itinerant medics such as myself dipping a hesitant toe into this area of scholarship.

Definitions and fuzzy semantics

There are probably nearly as many definitions of empathy as people working on the topic.

Over the last century in particular, “empathy” has been considered in multiple knowledge domains. For the newcomer to this topic, this raises some challenges, for there are multiple variations in the construct of “empathy” and its various related terms — sympathy, pity and compassion - in wide-ranging literatures, with a plethora of associated definitions. Even within the healthcare setting, there is significant conceptual diversity evident, and, as in the other domains, the disparate constructs and use of the terms empathy, sympathy and compassion can be hard to disentangle; indeed in some circumstances, they seem to be used interchangeably.

---

390 See Van Holtoon, n. 387 above, at page 36.
Thus, I have struggled to navigate the semantic maze that characterises these various research domains. However, despite these semantic inconsistencies and evident inherent ambiguity, I suggest that it is possible to distinguish some common ground in these related concepts. So let me summarise the definitions/concepts that I adopt here that I will build on in relation to my argument later in this chapter.

**Empathy**

The construct of empathy has been generally differentiated into two distinct reactions of an individual - emotional (affective) and cognitive – to another’s observed experience or situation. These are viewed as related but distinct cognitive processes that have separate evolutionary origins and differentiated developmental profiles, affective empathy built on primitive protective responses, while cognitive empathy represents a later evolutionary product.

In recent years, significant progress has been made in understanding the developmental origins and neuroscience that underlie these empathic responses. The revelation that when we observe the physical actions of another individual, “mirror neurons” reciprocally activate in our own brains was the first of a number of empirical studies that exposed the close neurophysiological interconnectedness underpinning human interaction. Similarly

---

395 See, for example, William Ickes, *Everyday mind reading: Understanding what other people think and feel* (Prometheus Books Amherst, NY. 2003) 350. Recent neuropsychological evidence suggests that these two forms of empathy, although they may work in tandem, are phylogenetically distinct, and associated with two separate brain systems (Erin B. Tone and Erin C. Tully 'Empathy as a "risky strength": a multilevel examination of empathy and risk for internalizing disorders' (2014) 26 (4 Pt 2) Dev Psychopathol 1547). Two main theories have dominated the evolving scholarship of how we develop an empathic understanding of others – Simulation theory and Theory theory. For a useful review, see Anat Perry and Simone Shamay-Tsoory, 'Understanding emotional and cognitive empathy: A neuropsychological perspective'(3rd edn, Oxford University Press 2013).


398 For review, see: P. F. Ferrari and G. Rizzolatti 'Mirror neuron research: the past and the future' (2014) 369 Philos Trans R Soc Lond B: Biol Sci 20130169. It is now recognized that a particular
functional MRI studies have suggested that the expression “I feel your pain” is real in those with strong emotional bonds to another individual. Such imaging and other clinico-pathological studies in a number of domains have now begun to disentangle the neurobiological basis of various emotional reactions. Thus, it has recently been shown that empathy and compassion (see later) are biologically separate and are dependent on different neuronal networks.

**Conceptual distinctions – emotionally “hot” and “cold” empathies**

Empirical research supports a neurobiological distinction between the two constructs of empathy – emotional and cognitive - noted above. Emotional empathy tends to be associated with a blend of affective reactions that may include experiencing similar feeling as the other person, distress at their situation, and a compassionate response. Thus, the construct of emotional empathy reflects an “I feel what you feel” response, and is associated with a mirrored but usually attenuated emotional reaction in an observer to another’s emotional state or situation.

On the other hand, cognitive empathy is considered a more advanced (and emotionally detached) perspective-taking cognitive process that provides an intellectual understanding of another’s emotional state: “I understand what you feel”, the latter based on appreciating...
another’s situation from their perspective. This construct of cognitive empathy has some resonance with the broader concept of “Theory of Mind” - the general ability to distinguish the mental state of another individual – that I referred to earlier. Unlike emotional empathy, cognitive empathy is considered to be more akin to a skill that can, to an extent, be developed (learned) with experience.

While the distinctions between these different forms of empathy provide insights into the development and functional aspects of human interaction, in day-to-day human encounters this separation seems somewhat artificial, for emotional and cognitive empathy rarely occur independently of each other, the relative balance of the two elements dependent on the particular circumstances of the encounter.

And it is also reasonable to suppose that other factors specific to the observer must also come in to play. Doubtless, I will react differently if I see my own grandson falling off his bicycle, compared to seeing the same happen to someone else’s grandchild. Thus, the range and degree of our emotional reactions are conditioned by what is significant to us, what most concerns us, whether biological ties, or professional attitudes and values, for, as I discuss later, strong empathic reactions may come into play in paediatricians’ encounters with allegedly abused children, primed by the cognitive prerequisites of professional values and concerns.

For completeness, let me briefly mention empathy’s cognitive cousins: sympathy and compassion.

---

Sympathy

Sympathy in this context seems a particularly ambiguous concept, with notably inconsistent distinctions made between empathy and sympathy, depending on the scholar and the knowledge domain. However Hein and Singer have illustrated the conventional distinction between the two constructs by employing two contrasting phrases: “feeling as [empathy] and feeling for the other [sympathy]” (emphasis in original).

Thus, as noted above, empathy involves the recognition of another’s emotional state or situation but with a significant blurring of the self-other distinction and, most notably, in emotional empathy, a congruent sharing in that emotional state. In contrast, sympathy is generally regarded as a more detached, and often incongruent, vicarious emotional response to another’s situation, typified as sorrow or concern for the other individual’s plight. Here there would also seem to be superficially little distinction here between cognitive empathy and sympathy. However, sympathy seems to reflect a secondary phenomenon, since it must begin with an empathic recognition of another’s situation. Thus, both an emotional or cognitive empathic response may provoke a secondary sympathetic reaction.

Compassion

Let me finally address the third element in this trio of human affective responses to another’s situation - compassion. Again, it is difficult to disentangle clearly this concept from

sympathy and there seems some overlap. Both sympathy and compassion may be regarded loosely as moral emotions, and they seem primarily to be elicited by negative emotional valence in others, notably distress.

Again, like sympathy, compassion also appears to represent primarily a secondary phenomenon, predicated on an empathic understanding of another’s situation. Indeed, a compassionate response - “[a] feeling that arises in witnessing another’s suffering and that motivates a subsequent desire to help” - may be considered an expanded form of sympathy, the concern for another’s plight augmented by an altruistic impetus to improve their situation by some action, the observer moved in a motivational sense to respond in some way to relieve the other’s suffering.

Interim summary

As I release you from this brief foray into the semantic jungle of our reactive human emotional responses to the plight or situation of others, what I hope to have shown in Section 1 is that, irrespective of the subtle distinctions and apparent overlapping constructs, there is clear evidence that as humans we are innately conditioned to identify and react to the plight of others, and that such reactions will vary depending on how much we are emotionally connected to them.

I provided a necessarily brief overview of the in-built nature of human empathy, its fundamental role in how we recognise each other’s situations, and how such understanding represents an essential aspect of how we relate to each other as individuals. I offered a short (and necessarily simplified) description of current technical constructs of empathy, contrasting the concepts and mental processes involved in emotional and cognitive empathy.

---


Goetz et al, n. 408 above, at page 351.
I suggested that the empathic response of an individual, both in intensity and in the balance of its affective and cognitive components, is likely be significantly influenced by the personal perspectives of the observer and their relationship – familial or professional – with those with whom they interact.

In Section 2 I move on to discuss the particular place of empathy and compassion in clinical practice, a foundation on which I later build my argument that such emotional connections may have particularly significant implications in the context of forensic paediatric clinical assessment.
Section 2 - Empathy and compassion in the healthcare setting

The issue that I want to address in the remainder of this chapter is whether, through engagement of the group of empathy-related responses noted above, the forensic paediatric expert’s vicarious emotional responses to an injured or dead child may make them vulnerable to a loss of objectivity and thus undermine the opinions they provide.

I will argue that there is certainly a significant risk of that effect.

Part of my justification for that claim is based on the proposition that one of the prominent and professionally desirable characteristics associated with paediatricians – their ability to connect and empathise with their child patients – may in this context reflect a “risky strength”,411 a virtue that comes at a price, so that in the context of forensic interpretation in alleged child abuse there may be negative epistemic effects. Let me develop that argument further by first considering empathy and compassion in its broader healthcare context.

Empathy and compassion as a core component of effective healthcare

Compassion is strongly associated with vulnerability and undeserved suffering, so it is perhaps not surprising that empathy and compassion are considered to embody the core values of care, duty, and moral motivation seen as integral to effective healthcare delivery.412 This was not always the case. As a product of medical school in the 1970s, I well remember the atmosphere of caring but detached paternalism that characterised clinical teaching and the doctor-patient relationships at least in hospital practice well into the 1980s. Here emotional distance was maintained by a focus on the disease rather than the person, with

411 See Tone and Tully 2014, n. 395 above.
patients often dehumanised into clinical problems, labelled by their condition rather than their name: “the diabetic in cubicle 2”.

Explicit justifications for this depersonalised approach included concern that professional emotional involvement might interfere with objective clinical decision-making. However, dealing with deaths on a regular basis, the clinical uncertainties underlying much clinical practice, and concern about medical errors, many of which were never acknowledged to patients or their families, are all likely to have contributed to this emotional distance. In paediatric settings, where three-way discussions between doctor, child and parents complicate such interactions, communication difficulties, including the liberal use of medical jargon, reflected this paternalistic professional-patient separation.

In the last 50 years or so, there has been a paradigm shift to a more egalitarian, patient-centred model of care, with a strong emphasis on dialogue, empathy and shared decision-making. This change not only reflected concern, among other things, about the severely skewed power dynamic that had characterised the doctor-patient relationship, but empirical studies had showed that empathy plays an important role in the effectiveness of medical care, benefiting both the patient and the clinician.

Reflecting this changed perspective, compassionate care is now explicitly identified as a professional obligation of all clinicians, and a core element of clinical practice. With

414 Whether doctors acknowledge this perceived power imbalance is less clear: Laura Nimmon and Terese Stenfors-Hayes 'The "handling" of power in the physician-patient encounter: perceptions from experienced physicians' (2016) 16 BMC Med Educ 114.
this wide acceptance that the ability to see a situation “through someone else’s eyes” is essential for an effective therapeutic relationship,⁴¹⁹ “empathy training” has become an integral component of some medical undergraduate teaching programmes,⁴²⁰ although in reality it is acknowledged that empathy cannot be directly taught, but rather that its development can be facilitated by providing the right conditions and resources.⁴²¹

Conversely it is also recognised that in some circumstances emotional attachment and over-identification with patients may impair proper clinical care.⁴²² Thus cognitive empathy – “detached concern” – has been strongly promoted in clinical practice in order to maintain appropriate objectivity, while emotional engagement with patients has been discouraged.⁴²³

This dichotomy seems desirable, and perhaps might be typified by surgical practice. You want your surgeon to understand your perspective on things and come to a joint decision with you about the most appropriate surgical intervention. However, when that same surgeon has opened up your body and is interfering with your vital organs, you want detachment and


⁴²⁰ Johanna Shapiro, ‘The Paradox of Teaching Empathy in Medical Education’ in Jean Decety (ed), Empathy: From bench to bedside (Social Neuroscience Series, MIT Press 2011) 275. The idea that in reality empathy can be “taught” to medical students, or anyone else, is controversial, as a recent east-west disagreement across the Scottish academic divide reveals: D. Jeffrey and R. Downie 'Empathy - can it be taught?' (2016) 46 J R Coll Physicians Edinb 107.

⁴²¹ Angeliki Kerasidou and Ruth Horn 'Making space for empathy: supporting doctors in the emotional labour of clinical care' (2016) 17 (1) BMC Medical Ethics 8.

⁴²² Nightingale et al., n. 402 above. This risk is perhaps most obviously exemplified in the firm guidance from the General Medical Council that doctors should not treat family members or others with whom they have a strong emotional ties: see Good Medical Practice, GMC (2013), available at http://www.gmc-uk.org/guidance/10247.asp (accessed 30 June 2018).

technical skill, and a cool head if something unexpected happens, rather than the shaky hands of emotional overload.

But balancing empathy with detachment may not be easy. As I noted above, while it seems that these cognitive processes may function autonomously, it appears that when an empathic engagement occurs, generally both cognitive processes are involved to some extent. An accurate interpretation of another’s situation may depend on dual activation. Indeed, some consider that in doctor-patient interactions such a detached albeit “purposeful” intellectual exercise in the absence of an affective empathic response must limit the quality of shared decision-making and may not be in the patient’s best interests.

Finding the right balance between emotional involvement and detachment may be difficult, and when dealing with some types of patients and their families, appropriate emotional detachment may be particularly challenging.

Certain clinical specialties, by their very nature, are set in emotionally challenging professional environments, and paediatrics is recognised as one where particularly demanding emotional encounters may occur.

---

426 As Halpern has argued in the context of empathy in clinical care, while “emotional involvement introduces risks of error”, without a true emotional connection with a patient, their perspective cannot be fully understood (J Halpern, 'From detached concern to empathy: humanizing medical practice' (OUP 2001) 165, page xxiii).
Are paediatricians particularly vulnerable to empathic emotional contagion?

Paediatricians are reported to have some “unique” (primarily positive) personality characteristics compared to other medical specialties that make them particularly sensitive and responsive to the needs of children. These include particularly high levels of conscientiousness, agreeableness, and emotional stability. Aspiring paediatricians are also reported to have particularly high levels of empathy compared to physicians in some other specialties. It is unclear if this indicates specific characteristics intrinsic to the specialty that attract certain personality types, or whether it is an indirect reflection of the comparatively high proportion of women within the specialty, a gender claimed to have generally higher empathy quotients compared with their male counterparts.

As I discussed earlier, if some degree of emotional engagement is generally unavoidable in healthcare encounters, and indeed desirable, what of the effect of particularly emotive clinical situations, such as that often encountered in paediatric clinical care? Perhaps not surprisingly, in such situations a significant degree of emotional empathy may be provoked, and in certain circumstances, “over-arousal” may occur when exposed to another’s

---

427 Elizabeth Schott and others. 'Relationship of the Personality Traits of Early Career Pediatricians and Career Satisfaction and Burnout' (Pediatric Academic Societies Annual Meeting Baltimore, April 30, 2016). Naturally, such career choices are self-selected.


431 Mohammadreza Hojat, 'Empathy and Gender: Are Men and Women Complementary or Opposite Sexes?' in Mohammadreza Hojat (ed), Empathy in Health Professions Education and Patient Care (Springer International Publishing 2016) 169. Whether this is true is debatable; rather it may be a reflection of a traditional gender behavioural stereotype that women are more emotionally expressive than men, although men and women experience emotion equally: Y. Deng and others 'Gender Differences in Emotional Response: Inconsistency between Experience and Expressivity' (2016) 11 (6) PLoS ONE e0158666.

There is some evidence paediatricians’ high levels of professional commitment and personal engagement with child patients and their families may indeed take its toll. Thus, paediatricians working in intensive care settings may experience severe grief reactions and burnout,\footnote{Suzanne Crowe and others 'Grief and Burnout in the PICU' (2017) 139 (5) Pediatrics e20164041.} while those involved in the long-term care of children with life-limiting progressive conditions may find it especially difficult to make objective decisions about when to withdraw active care,\footnote{J. O. Beattie. 'Deciding about life-sustaining treatment for incompetent children: doctors, parents and the courts' (MML, School of Law, University of Glasgow 2010).} and suffer a sense of guilt or bereavement when their child patients die.\footnote{Jessica Plante and Claude Cyr 'Health care professionals' grief after the death of a child' (2010) 16 (4) J Paediatr Child Health 213.} These effects may not seem particularly surprising. However, more relevant to the argument I am pursuing in this chapter is the fact that paediatricians and other doctors whose work involves relatively brief encounters with unexpected infant deaths or severely ill or injured children may also suffer significant and persistent secondary psychological effects.\footnote{Clare E. Hollingsworth and others 'Impact of child death on paediatric trainees' (2018) 103 Arch Dis Child 14; S. R. Knazik and others 'The death of a child in the emergency department' (2003) 42 (4) Ann Emerg Med 519; J. Solberg and H. Range-Hudson 'Academic emergency physicians' experiences with patient death' (2011) 18 Acad Emerg Med 255; Francis J. Somville, Véronique De Gucht and Stan Maes 'The impact of occupational hazards and traumatic events among Belgian emergency physicians' (2016) 24 (1) Scand J Trauma Resusc Emerg Med 59.}

Perhaps not surprisingly, given the emotive and stressful nature of their work, a significant minority of paediatricians report “burnout”, underlying emotional exhaustion dominating such self-reports.\footnote{Tara F. McKinley, Kimberly A. Boland and John D. Mahan 'Burnout and interventions in pediatric residency: A literature review' (2017) 6 Burn Res 9; A. J. Starmer, M. P. Frintner and G. L. Freed}
Section 3 - Empathy, emotions and the implications for paediatric forensic judgments

If the combination of professional drive, emotional commitment and relatively high levels of burnout are reported in paediatricians across the range of paediatric sub-specialties, it is perhaps predictable that paediatricians working in child protection should exhibit particular evidence of psychological difficulties, for the responsibility for distinguishing between accidental and inflicted injury intrinsic to the forensic role is arguably just as stressful and demanding as that of the emotional demands of dealing with a sudden childhood death in an emergency department, or treating a severely injured or ill child in a PICU.\(^{438}\)

Although specific research in relation to paediatricians is sparse, there is evidence that doctors (and other professionals) involved in child protection work may experience considerable stress and emotional distress, with attendant risks of significant psychological harm.\(^{439}\) These are not the effects that would be expected in cognitively detached practitioners and must reflect the emotional engagement intrinsic to the role.

---

\(^{438}\) Work-Life Balance, Burnout, and Satisfaction of Early Career Pediatricians' (2016) 137 (4) Pediatrics e20153183. “Burnout syndrome” in doctors has generally been categorised by the joint prevalence of “high” emotional exhaustion and depersonalisation and a “low” sense of personal accomplishment, although the relative dominance of emotional exhaustion in such assessments is debated; see Jan Beckstrand, Nancy Yanchus and Katerine Osatuke 'Only One Burnout Estimator Is Consistently Associated with Health Care Providers Perceptions of Job Demand and Resource Problems' (2017) 8 (7) Psychology 23.

Such vicarious trauma - described as “a normal response to repeated exposure and empathetic engagement with traumatic material”\textsuperscript{440} - is a well-recognised and significant problem noted in social work child protection practitioners,\textsuperscript{441} and others working with abused clients in various professional settings, including lawyers.\textsuperscript{442} It is correlated with empathic engagement, where “[e]mpathy is the helper’s greatest asset and also possibly his/her greatest liability”\textsuperscript{443}.

Thus, it seems that the specific personality and other features considered particularly suited to a paediatric career might indeed represent “risky strengths” in the context of the applied expertise of paediatric forensic work.

And such empathic affective responses may elicit a number of secondary emotional reactions – including pity, compassion, and anger - so might such emotions also affect the paediatrician’s decision-making processes and the outcomes of those processes? There is a good theoretical foundation for such concern.

**The potential impact of emotions on diagnosis and forensic judgments**

While much about the subtle inter-relationships of emotions and cognition remain to be explained, outwith the clinical/forensic situation, emerging scholarship shows that that “[e]motions powerfully, predictably, and pervasively influence decision-making”\textsuperscript{444}.

\textsuperscript{440} Zoë Morrison, Australian Institute of Family Studies and Australian Centre for the Study of Sexual Assault, *Feeling heavy*: vicarious trauma and other issues facing those who work in the sexual assault field (Melbourne: Australian Institute of Family Studies 2007), page 1.

\textsuperscript{441} *Vicarious trauma: the consequences of working with abuse* (NSPCC 2013) 1. See also: Allison Clarke. 'Secondary traumatic stress in child welfare work' (Masters in Education, University of Lethbridge 2010).

\textsuperscript{442} A. P. Levin and others 'Secondary traumatic stress in attorneys and their administrative support staff working with trauma-exposed clients' (2011) 199 (12) J Nerv Ment Dis 946.


\textsuperscript{444} J. S. Lerner and others 'Emotion and decision making' (2015) 66 Annu Rev Psychol 799, p. 802. For reviews, see also G. Loewenstein and others 'Risk as feelings' (2001) 127 (2) Psychol Bull 267;
Some of these emotions may be incidental - generated by factors unrelated to the specific judgment to be made – perhaps fatigue or previous inter-professional disagreement. In other circumstances, like those in which the paediatric forensic specialist works, the emotions may be generated by circumstances integral to the particular judgment itself.\footnote{445}

Whether we are aware of it or not, such integral emotions are now known to be particularly likely to influence the conclusions we ultimately reach, tending to distort rational judgments. This may be partly explained by the effects that emotions have on the cognitive processes we use, for in the context of the dual-system of reasoning I discussed in Chapter 3, emotions associated with a particular task are thought to influence the relative extent to which intuitive/heuristic (type 1) or analytic, deliberative (type 2) processes contribute to the judgment.

Thus, for example, there is consistent empirical research suggesting that individuals in positive (e.g. optimistic or happy) mood states tend to use heuristic processing when making judgments, while, negative mood states, such as sadness, may induce a more deliberative reasoning process.\footnote{446}

However, moods are not the same as emotions, which are usually reactive and focused around or towards a specific event.\footnote{447} Again, while different emotions at the time of decision making are reported to be associated with different cognitive processing responses,\footnote{448} their

\begin{footnotesize}
\begin{enumerate}
\item Lerner, n. 444 above. See also P. Litvak and others, 'Fuel in the fire: How anger impacts judgment and decision making' in M. Potegal, G. Stemmler and C. Spielberger (eds), International Handbook of Anger (Springer; Springer 2010) 287.
\item See, for example: Galen V. Bodenhausen, Lori A. Sheppard and Geoffrey P. Kramer 'Negative Affect and Social Judgment: The Differential Impact of Anger and Sadness' (1994) 24 (1) Eur J of Soc
\end{enumerate}
\end{footnotesize}
effect on an individual’s judgment in a given situation may depend on the particular emotion(s) that were provoked. The results are discordant with those associated with positive and negative mood states.

Thus, although anger may be considered a “negative” emotion and might be expected to induce an analytical cognitive response (c.f. the effect of a negative mood state), anger appears to have unique effects that provoke particular cognitive responses. Concurrent anger often leads to stereotyping and biased judgments of others, an enhanced sense of confidence and certainty in one’s own judgment and is reported to induce heuristic rather than systematic/analytical processing. Such emotionally driven cognitive responses may potentially have an impact on the expert paediatric forensic evaluations in cases of alleged abuse.

As I noted earlier, paediatricians may have profound emotional reactions, including anger, when dealing with some child abuse cases. Given the evolving scholarship on the impact of emotions on judgment and decision-making, it seems reasonable to suggest that when dealing with a severely injured child, such emotional reactions may, through unconscious intuitive reasoning, undermine their objective interpretation of the clinical features and their subsequent forensic opinion of the likelihood of underlying abuse.

While awareness of these emotions and their impact may be partly ameliorated by anticipation and conscious recognition of the need to temper compassion while retaining objectivity, such automatic responses may be unmodifiable. For these reasons, true

---

449 Litvak et al, n. 446 above.
objectivity may be unattainable, even if conscious efforts are made to recognise subjective feelings.

Finally, I need to recognise that much of the content of this chapter is particularly relevant to expert paediatricians involved in primary child abuse forensic assessments, rather than external paediatric experts brought in later to provide secondary opinions by the prosecution authorities or defence legal teams. Having acknowledged that, it is important to emphasise that even “at a distance” such cases can be highly emotive, with, for example, exposure to graphic injury photographs the norm, so although external experts are somewhat protected from such influences, the emotional impact of such cases may still be significant.

It is also important to recognise that the impact of emotions on decision-making may not always be negative. Some emotions may have a positive epistemic effect – fear and anxiety in anticipation of a difficult time in the witness box may make an expert circumspect in their conclusions.

And of course, while it is important that I acknowledge that judgments and decisions are never made in an emotional vacuum, and that fully rational decision-making may need emotional input, the focus of this thesis is on potential frailties in paediatric expert forensic evidence. In that specific context, I suggest that there are potentially many more negative than positive influences of emotion on such forensic judgments.

As I close this chapter, it is appropriate to acknowledge that it is not only forensic clinicians that are at risk of bias due to the emotional context and content of forensic work, for there is evidence that jurors and judges are not immune from such cognitive contamination.

---


Conclusion

In the last two chapters I have been considering factors that may influence how we think and make decisions, and how these might relate to and adversely impact on the objectivity demanded of forensic interpretation in the context of alleged inflicted childhood injury.

In the preceding chapter I discussed the broad landscape of human judgment and decision-making, including clinical diagnosis and medical judgments. I described some of the current key concepts of cognition that underlie such responses, emphasising the fundamental role that unconscious intuition plays in such activities. I then moved on to examine how such processes apply in the context of forensic assessment, showing the wide range of possible sources for biased forensic interpretation, but in particular the risks from contextual bias. Finally, I considered that scholarship in the context of paediatric forensic interpretation, an area where the potential risks of such biases have received virtually no attention. I exposed the wide range of possible biases that may threaten the epistemic justification for paediatric forensic assessment.

In the present chapter, I have moved on to consider a separate potential source of bias in paediatric forensic assessment, questioning whether the highly emotive nature of child protection work may independently threaten a paediatric expert’s objectivity.

Having established that empathic understanding and feelings for others is integral to being human, I have argued that although this area of scholarship had not been addressed so far, there were good theoretical reasons to claim that such a potentially biasing effect was highly plausible.

And as I had previously noted, the broader professional imperative of the safeguarding/child protection ethos focused on child welfare and risk management intrinsic to the work of all paediatricians also provides an influential perspective against which both emotional reactions and forensic judgments are set.
Given that such biases may threaten the reliability of paediatric expert evidence and may increase the risk of flawed outcomes in criminal and civil cases involving alleged abusive injury or worse, how might such risks be mitigated?

One way that a forensic paediatrician might minimise the impact of the potential biases I have described in the last two chapters is to ensure that, as far as possible, their conclusions and related opinion evidence is securely grounded on a sound scientific foundation and thus reaches a level of reliability acceptable to the relevant legal fact finders.

Accordingly, I turn in the next chapter to examine the evidence base that purports to underpin such opinions.
Chapter 5 - Medicine-based evidence in alleged abuse: a secure epistemic basis for expert opinion?

Introduction

In the preceding chapters I highlighted a number of issues that both individually and collectively might undermine the objectivity of a paediatrician in formulating a forensic expert opinion in a case of alleged child abuse.

These included societal pressures and expectations, the possible impact of various ingrained professional roles and values intrinsic to the safeguarding ethos, and the potential for therapeutic carer/ forensic role conflicts. I also discussed relevant aspects of human cognition, in particular the risks posed by the largely intuitive nature of the expert clinical diagnostic process generally, and the potential impact of cognitive biases, in particular from unavoidable exposure to negative contextual information linked with parallel multi-agency investigations and the associated mandatory information-sharing process. I argued that the paediatrician’s interpretation and opinion is liable to be contaminated by circumstantial information, including family circumstances and others’ professional concerns about potential risks, co-operation and lifestyle.

I suggested that, in addition to these factors, the highly emotive nature of such cases might separately threaten the impartiality of paediatric experts faced with the challenge of deciding if a young child has been deliberately or accidently injured. Arguably, all of these threats to expert objectivity might be neutralised, or at least significantly mitigated, if the clinical and scientific evidence base on which such opinions were founded was sufficiently robust.

In this chapter, I turn to explore the epistemic foundation of the evidence base that expert paediatricians must rely on in coming to a conclusion that a child has or has not been deliberately injured, and from which they must justify any expert opinion they may provide to the family and criminal courts and associated fact-finding tribunals. The chapter is presented in three sections.
In Section 1, I provide a brief overview of the evolution of evidence-based medicine (EBM) and introduce the general approach and professional obligations that underlie the move to evidence-based practice. Here I highlight EBM’s focus on critical appraisal linked to a hierarchical evidence pyramid to reflect the quality of the evidence and discuss how such evidence might be embedded within clinical practice, particularly through graded recommendations within various forms of clinical guidelines.

In Section 2 I move on to describe the development of the evidence base for clinical child protection work and illustrate how critical appraisal exposed the dubious scientific foundation of significant elements of forensic interpretations on which paediatricians had previously based dogmatic expert opinions.

Finally, in Section 3, while recognising the significant efforts that continue to be made to develop a more secure evidential foundation to support paediatricians in the clinical assessment of possible inflicted injury, I question the reliability and validity of a number of key elements on which the new child protection evidence-based has been built and applied to a variety of clinical presentations. In particular, I argue that the use of the outcomes of judicial and quasi-judicial tribunals to provide a justification of the validity of a clinical forensic interpretation presented to that tribunal is fundamentally flawed.

I begin by setting the evidence base for forensic paediatric opinions of infant harm within the broader context of EBM.
Section 1- The birth of evidence-based medicine

Evidence-based medicine has been embraced wholeheartedly, and rightly so, as the best approach for reducing clinical uncertainty and ensuring that patients receive treatment and care that are efficacious (i.e. they work) and effective (i.e. they work in real life). The last century saw a revolution in healthcare, most notably in a reversal of the balance within the complementary dyad of art and science that characterises medical practice. Medical care during the first half of the 20th century had followed a traditional paradigm that was highly focused on clinical diagnosis and the prediction of prognosis. Available laboratory investigations were very restricted in scope and sophistication, while effective therapeutic options were highly limited.

The second half of that century saw a major shift (if not reversal) in the equilibrium of the dyad. Stimulated by the demands and innovations of World War Two, the dominance of clinical observation and passive empathic care gave way to care based on progressive technological innovation and the availability of science-based methodologies. A rapid evolution of novel therapeutic and investigative techniques galvanised clinicians who applied them enthusiastically.

Despite this progress, at this time much clinical research was of a descriptive, case-based nature, and there was little attempt to review the outcomes of novel interventions systematically. Clinical enthusiasm was not always tempered by caution. Many treatments

---

456 For example, although Fleming had discovered and identified the potential of penicillin in 1928, production was constrained by limited commercial chemistry development at the time. High volume manufacture of antibiotics developed during World War 2, and was initially limited to military use: R. Sykes 'Penicillin: from discovery to product' (2001) 79 (8) Bull World Health Organ 778.
457 Perhaps one of the best examples of such a spin off was the invention of diagnostic ultrasound scanning by Dr Ian Donald, Regius Professor of Midwifery at the University of Glasgow in the 1950s, based on his exposure to radar and sonar technologies used during the Second World War. See: http://www.universitystory.gla.ac.uk/biography/?id=WJ2489&type=P (accessed 14 May 2018).
doctors used had little if any evidence of efficacy, some were positively harmful, and there were widespread variations in the ordering of tests and treatments for a variety of conditions across many medical disciplines. Much clinical decision-making was based on custom and practice, personal experience and the influence of local opinion formers, a process later described as “eminence-based medicine”.

It was only in the latter part of the 20th century that clinicians began to objectively and systematically review the impact of their investigative and therapeutic choices. This change can be traced to a series of provocative lectures in 1972 by the Oxford epidemiologist Archie Cochrane. Cochrane challenged clinicians to systematically and objectively evaluate the treatments and other interventions they used. He emphasised the need to use unbiased tools such as the randomised controlled trial (RCT) and high quality observational studies to assess treatment benefits and argued that such an approach should be part of a continuous cycle of review and critical reappraisal of clinical knowledge and its application.

Some paid a price for this well-meaning but unbounded eagerness to embrace new technologies and ideas. One example was the practice of radiation therapy applied to chests and neck of infants in the 1940s and 50s to “shrink” their thymus glands. The rationale underlying this therapy was concern about the possible risk of “suffocation” or cot death. The legacy of the infant radiation therapy for this “non-disease” is seen now in a marked incidence in underactive thyroid disease in these individuals, and higher rates of thyroid and other cancers generally in these individuals in later life. See M. J. Adams and others 'Thyroid cancer risk 40+ years after irradiation for an enlarged thymus: an update of the Hempelmann cohort' (2010) 174 (6) Radiat Res 753.


A comparative study in which participants are randomly allocated to intervention and control groups and followed up to examine differences in outcomes between the groups.
While Cochrane had conceived the concepts underlying EBM, it was not until the early 1990s that Canadian researchers at McMaster University first used the term. Later defined as “the conscientious, explicit and judicious use of current best evidence in making decisions about the care of individual patients”, EBM was rapidly and enthusiastically embraced by a vigorous intellectual community inspired to ensure that clinical diagnosis and care was founded on valid empirical science.

**Critical appraisal and evidence hierarchies**

Critical appraisal and stratification of evidence are the core foundations that underpin the application of the EBM paradigm. While “best evidence” is rarely defined, published evidence is usually analysed and stratified against one of a number of evidence hierarchies. The relevant grades are “based on ideological or consensus judgments about which studies are most likely to provide estimates of ‘truth’ through reliable and valid data”.

While many evidence hierarchies have been developed, inconsistencies and other weaknesses have led to the development of an international consensus on rating quality of clinical evidence and strength of recommendations - the Grading of Recommendations

---

466 Evidence-Based Medicine Working Group 'Evidence-based medicine. A new approach to teaching the practice of medicine' (1992) 268 (17) JAMA 2420.
469 Kerridge, n. 467 above, 366.
Assessment, Development and Evaluation (GRADE)\(^{470}\) – an standard approach adopted or endorsed by a wide variety of relevant organisations. Nevertheless, the key relevant agencies in the UK, the Scottish Intercollegiate Guideline Network (SIGN) and the National Institute for Health and Care Excellence (NICE), use their own taxonomies for evidence appraisal and grading of recommendations. Box 1 shows the SIGN evidence grading system.\(^{471}\)

<table>
<thead>
<tr>
<th>Grade</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>++</td>
<td>High-quality meta-analyses, systematic reviews of RCTs, or RCTs with a very low risk of bias</td>
</tr>
<tr>
<td>+</td>
<td>Well-conducted meta-analyses, systematic reviews, or RCTs with a low risk of bias</td>
</tr>
<tr>
<td>-</td>
<td>Meta-analyses, systematic reviews, or RCTs with a high risk of bias</td>
</tr>
<tr>
<td>++</td>
<td>High-quality systematic reviews of case-control or cohort studies</td>
</tr>
<tr>
<td></td>
<td>High-quality case-control or cohort studies with a very low risk of confounding or bias and a high probability that the relationship is causal</td>
</tr>
<tr>
<td>+</td>
<td>Well-conducted case-control or cohort studies with a low risk of confounding or bias and a moderate probability that the relationship is causal</td>
</tr>
<tr>
<td>-</td>
<td>Case-control or cohort studies with a high risk of confounding or bias and a significant risk that the relationship is not causal</td>
</tr>
<tr>
<td></td>
<td>Non-analytic studies; for example, case reports, case series</td>
</tr>
<tr>
<td></td>
<td>Expert opinion</td>
</tr>
</tbody>
</table>

Box 1: SIGN classification for grading evidence.

As is evident in the SIGN classification, most evidence hierarchies consider that RCTs and meta-analyses provide the most reliable form of research data.\(^{472}\) Yet despite their iconic


\(^{472}\) Atkins et al., n. 468 above; Alessandro Liberati and others, 'Grading Quality of Evidence' in Virginia A. Moyer and Elizabeth J. Elliot (eds), *Evidence-based Paediatrics and Child Health* (BMJ Publishing Group 2004) 85.
status, even RCTs may have intrinsic problems. This illustrates a simple truth - that published clinical research must be read carefully and interpreted cautiously, lest clinicians change their practice on the basis of misleading claims.

And despite its high status, the RCT design is not appropriate for all types of clinical research. Different clinical questions demand different study or research designs and they themselves are dependent on the study population or other resource available for investigation. Clinical diagnostic or prognosis research questions are often best answered by high quality observation studies. Such studies are more prone to bias than RCTs, and thus careful consideration of the design and reporting of such studies is particularly important is assessing the validity of the study conclusions. Specific guidance on the analysis of such studies is available - the STROBE checklist - to ensure that any potential biases in such observational studies are more likely to be exposed.


476 There are three main analytical designs used in observational research: cohort, case-control, and cross-sectional studies. All present various difficulties in minimising bias. See S. Straus and others, *Evidence-based medicine: How to practice and teach EBM* (3rd edn, Churchill Livingstone Edinburgh 2005) 180-3.

477 See, for example, G. P. Hammer, J. B. du Prel and M. Blettner 'Avoiding bias in observational studies: part 8 in a series of articles on evaluation of scientific publications' (2009) 106 (41) Dtsch Arztebl Int 664. Such biases may occur from a variety of design difficulties. Thus, mothers of infants with congenital malformations are more likely to report potential adverse exposures during pregnancy than unaffected control mothers: M. M. Werler and others 'Reporting accuracy among mothers of malformed and non-malformed infants' (1989) 129 (2) Am J Epidemiol 415.

Over the last 20 years the concept of EBM and its applications have become embedded and intimately intermeshed within virtually all aspects of healthcare and across all clinical specialities. This is not to suggest that the concept of EBM or its practical applications have been without difficulties or critics. Journal policies resulting in the selective reporting of positive results, distortion of the research agenda by “big pharma”, the creation of new “disease” states by overreaction to subtle risk assessment evidence of dubious individual clinical benefit, on-going concern about a lack of focus on the personal circumstances and values of individual patients, and apparent poor appreciation of the contribution of professional experience and judgment in individualised clinical care are among key concerns that are fuelling a recent call for a reconsideration of the EBM agenda. Notwithstanding these concerns, EBM is here to stay.

---

479 Carl Heneghan 'Developing evidence-based medicine in everyday practice' (2008) 4 (5) Foundation Years J 207. Heneghan et al. emphasise the universality of the application of EBM in everyday clinical practice, whether newly qualified junior trainee doctors in hospital practice or long-established family doctors.

480 V. Moyer and E. Elliott 'How to practice evidence-based pediatrics' (2001) 174 (3) West J Med 158. Thus EBM has been incorporated into undergraduate and postgraduate medical curricula (Paul Glasziou, Amanda Burls and Ruth Gilbert 'Evidence based medicine and the medical curriculum' (2008) 337 BMJ a1253; R. Hatala and G. H. Guyatt 'Evaluating the teaching of evidence-based medicine' (2002) 288 (9) JAMA 1110), and has become established as a fundamental element of clinical care, where skills in searching electronic databases, asking structured clinical questions, and critically appraising the literature are now regarded as core elements of professional practice (M. Rudolf and others 'What's the use of Archimedes?' (2002) 87 (2) Arch Dis Child 168).


482 J. Sterne, M. Egger and D. Moher, 'Addressing reporting biases', in J. P. T. Higgins and S. Green (eds), Cochrane Handbook for Systematic Reviews of Interventions (The Cochrane Collaboration 2008) 297. One important effort has been the demand that all clinical trials should be registered: Catherine De Angelis and others 'Clinical Trial Registration: A Statement from the International Committee of Medical Journal Editors' (2004) 351 (12) N Engl J Med 1250). In 2010, a UK Clinical Trials Gateway (UKCTG) was established, offering a single entry point and environment for trials involving UK participants; see https://www.ukctg.nihr.ac.uk (accessed 4 May 2018).


484 For example, recent intense debate about the evidence of benefits and risks of the universal use of statin cholesterol-lowering drugs in the over 50s; see Margaret McCartney 'Statins for all?' (2012) 345 BMJ; F. Taylor and others 'Statins for the primary prevention of cardiovascular disease' (2013) 1 Cochrane Database Syst Rev CD004816.
**EBM in general paediatric practice**

Like their colleagues in other specialties, paediatricians adopted and have attempted to apply EBM approaches to their clinical care. Here they have faced significant hurdles.

Undertaking research in children, like all vulnerable groups, poses particular ethical and practical challenges, and so it is hardly surprising that, in comparative terms, both the quantity and quality of research in children lags far behind that of adults. Thus, in making therapeutic or other clinical decisions, paediatricians often have to extrapolate from adult studies, using interventions or treatments that may not have been formally evaluated in children. Children sometimes pay a significant price for this knowledge gap.

Despite these constraints, the development of an evidence-based approach to paediatric clinical care has gathered momentum, and while expert consensus, together with

---


487 J. W. Zylke, F. P. Rivara and H. Bauchner 'Challenges to excellence in child health research: Call for papers' (2012) 308 (10) JAMA 1040.

488 Maurizio Bonati, Evelyne Jacqz-Aigrain and Imti Choonara 'Licensed medicines, off-label use or evidence-based. Which is most important?' (2016) 102 (1) Arch Dis Child 53; Nicolino Ruperto and others 'Extrapolation or controlled trials in paediatrics: the current dilemma' (2017) 102 (10) Arch Dis Child 949.

extrapolation from adult research, remain prominent elements, EBM is now firmly embedded in paediatric professional practice both in the UK and internationally.

However, as I now discuss in Section 2, the development of an evidence base on which to base diagnostic formulation and forensic opinions in relation to child protection and alleged child maltreatment has presented particular challenges.


491 Responding to the drive for evidence-based care and consistency in practice, the RCPCH Clinical Standards Committee has overseen the development of a number of guidelines involving various paediatric medical conditions, some highly specialised. Additionally, the RCPCH reviews and, if appropriate, endorses clinical practice guidelines produced by other relevant guideline-development agencies and professional societies, most notably from the Scottish Intercollegiate Guideline Network (SIGN) and the National Institute for Health and Care Excellence (NICE). See: https://www.rcpch.ac.uk/rcpch-guidelines-and-standards-clinical-practice (accessed 10 May 2018).
Section 2 - A brief evolutionary history of the evidence base for child protection

While it was recognised over 150 years ago that children could suffer abuse and neglect,
society, politicians, and the medical profession showed little inclination to intervene in what
were often regarded as private, domestic matters. Accepting that historical background, it
is usual to consider that the modern era of concern about child abuse began in the second
half of the last century when it was “rediscovered” as a recognised clinical and social
problem. In 1962, coinciding with the emergence of paediatric medicine as a distinct
professional specialty, Henry Kempe and colleagues in Denver, Colorado produced a
seminal paper that defined a range of clinical features under the umbrella label of the
“battered child syndrome”, emphasising the duty of paediatricians to recognise and
intervene in such cases. However, despite this recognition, there was still ambiguity about
the thresholds for concern – for as Kempe noted: “child abuse is the difference between a
hand on the bottom and a fist in the face”. Today such a distinction clearly no longer
applies.

492 A. Tardieu ‘Etude médico-légale sur les services et mauvais traitements exercés sur les enfants’
493 Meadow 1997, n. 303 above. For a helpful historical overview of the recognition of child abuse
and neglect as a specific issue, and the evolution of social work interventions and professional practice
in relation to child protection, see Brian Corby, David Shemmings and David Wilkins, Child Abuse: An
494 Corby, n. 493 above.
495 C. H. Kempe and others ‘The battered-child syndrome’ (1962) 181 (1) JAMA 17. In 1958,
Professor Kempe established one of the first Child Protection Teams to identify and child abuse; see
496 Nigel Speight, ‘Non-accidental Injury’ in Roy Meadow (ed), ABC of Child Abuse (3rd edn, BMJ
497 A Member’s Bill that aims to remove the defence of “justifiable assault” from Scots law, which
can currently be used by parents who punish their children was recently brought before the Scottish
Parliament, and is supported by the Scottish Government. See:
Over the following 30 years or so, child abuse was a clinical problem dealt with within general paediatric practice. Until perhaps the early 1990s, many of the clinical issues in relation to suspected inflicted injury (with the exception of the emerging challenge of child sexual abuse and factitious illness) seemed relatively straightforward. When child abuse was suspected, children would be admitted to an informal “place of safety” – often represented by a hospital children’s ward. Local social work services would take the lead in coordinating inter-professional opinions and contributed their own assessment of family and other relevant circumstances. The police became variably involved depending on the case details and local circumstances. 498

The child would be retained in hospital until a case conference was convened, 499 where the case details were reviewed and decisions made about the child’s future. Here conclusions and statutory interventions were often linked to the perceived severity of the injury. This was a factor often left to social workers rather than medical opinion, using a simple informal injury severity scale. 500

For the paediatrician, while these could be uncomfortable and time-consuming cases to deal with, there was a professional confidence in the clinical judgments made, often reinforced by stereotypical perceptions of the families from which such cases arose. 501

498 In those early days, dedicated police officers specially trained in child protection issues were rare and the sometimes-unsubtle contribution of the police was not always helpful.

499 At this time, in the face of parental objections, local authority applications to the Sheriff Court for “place of safety” orders were rarely declined; such a low threshold for care orders would change later.

500 As noted by a Canadian report, “[c]ases of physical harm were defined by the investigating child welfare worker's judgment of the presence of an injury or health condition (visible for a minimum of 48 hours) because of maltreatment. The severity of harm was based on the investigating worker's judgment about whether medical treatment was necessary: moderate harm was observable but not thought to require medical care, whereas severe harm was defined as requiring medical treatment” (M. G. Ward and S. Bennett 'Studying child abuse and neglect in Canada: we are just at the beginning' (2003) 169 (9) CMAJ 919).

501 Early theoretical constructs of child maltreatment within families placed the blame firmly on psychological defects in the parent or carer that required understanding and treatment (Kempe, n. 495 above). Others acknowledged the impact of poverty and social disadvantage among many other external factors that interacted with the child and family to set up the conditions for abuse and neglect (David G. Gil, Violence against children: physical child abuse in the United States (Harvard University Press 1971) 220).
Observational studies, case reports and personal experience provided the foundation on which paediatricians based their diagnosis of child abuse in its various guises and informed the opinions they authoritatively provided to the justice system and statutory authorities.

**The emerging evidence base for child abuse diagnosis**

However, coinciding with the emergence and designation of child protection as a paediatric subspecialty with specific training needs and development of a curriculum required to support that, there was mounting concerns about the quality and reliability of the scientific evidence supporting paediatricians in dealing with child protection cases, in particular in relation to the forensic aspects of that work. Justifiably, the judiciary were becoming concerned about divergent expert opinions provided to the courts in child abuse cases, yet apparently founded on the same “evidence”. The highly publicised “shaken baby” murder trial of English nanny Louise Woodward in the United States in relation to the death of infant Matthew Eappen (see later, Chapter 7) also opened the public’s eyes to the fact that there could be major disagreements among expert witnesses in relation to such issues.

Such inconsistencies were hardly surprising. Although by the turn of the last century EBM had become an increasingly established and integral element within general paediatric clinical practice as it had across medicine more generally, an evidence-based approach to forensic diagnosis had largely bypassed the child protection community. As Cowley et al noted:

---


504 E. Butler-Sloss and A. Hall 'Expert witnesses, courts and the law' (2002) 95 (9) J R Soc Med 431; see also Lindberg et al., n. 54 above. Such variation in expert opinion evidence in relation to the likelihood that a child’s injuries are inflicted must reflect a number of factors, including the evidence base on which such opinions might be founded, but is also likely to reflect other subjective factors.

In 2002 there were no national guidelines on the recognition of abuse or neglect, nor was there any standardised evidence-based training for all paediatricians … [Child protection] training was offered on an *ad hoc* basis around the country …[while] no systematic reviews into the recognition or investigation of suspected physical abuse or neglect had been conducted.

**An evolving literature to inform child protection practice**

Reacting to these deficiencies and coinciding with the emerging needs of the new clinical specialty of child protection, an expanding international literature on all aspects of child abuse has gradually become available. This scholarship has evolved from isolated case reports and case series into sophisticated multisite observational studies and literature syntheses.  

Supported by the NSPCC and the RCPCH, a multidisciplinary team was established – the Welsh Child Protection Systematic Review Group (WCPSRG) based at Cardiff University - to provide a series of systematic reviews focused very specifically on the world literature in relation to the recognition, investigation and forensic interpretation of alleged child abuse.

Over the last 15 years or so the WCPSRG has produced a very impressive quantity of high quality analyses of the scientific and clinical evidence underlying many specific aspects of clinical child protection work. What their work has also clearly demonstrated is that, with a few notable exceptions, currently “…in many [topic] areas there is a paucity of high-quality, reliable, evidence-based research” on which to base forensic opinions in child protection.

---


cases, and that previous confident expert opinions on a number of important forensic issues were frankly wrong.

It became apparent that significant areas of established dogma in relation to the paediatric forensic interpretation of injury mechanisms, and the likely time frame for such injuries, had very little foundation and did not stand up to stringent scientific scrutiny. Thus, for example, the hitherto confident “ageing” of bruising based on the belief that bruises typically faded in a specific colour sequence, provides a good example where longstanding dogma was exposed as unreliable. Applying an EBM approach, review of the scientific literature showed that there was no sound basis on which to suggest the age of bruises either by direct clinical examination or from photographs. In contrast to previous practice, providing such an opinion in child protection proceedings, even within broad parameters, is now strongly discouraged.

---

510 See, for example, S. Maguire and others 'Diagnosing abuse: a systematic review of torn frenum and other intra-oral injuries' (2007) 92 (12) Arch Dis Child 1113.
511 Timing of death and injuries is a problematic area across all forensic practice; see, for example, Roger W. Byard 'Timing: the Achilles heel of forensic pathology' (2017) 13 (2) Forensic Sci Med Pathol 113.
512 Speight, n. 496 above; see also E. F. Wilson 'Estimation of the Age of Cutaneous Contusions in Child Abuse' (1977) 60 (5) Pediatrics 750.
514 Child Protection Companion, n. 509 above. Even such general terms as “old” or “fresh” bruising are strongly discouraged, although there are circumstances when “fresh” can be inferred from associated soft tissue swelling. The dating of fractures in children is also of obvious medicolegal importance but it is similarly now accepted that previous confident reports describing narrow potential timings of injury were unjustifiable, and that any opinion on the likely “age” of a fracture can only be suggested in very broad terms: I. Prosser and others 'A timetable for the radiologic features of fracture healing in young children' (2012) 198 (5) Am J Roentgenol 1014.
The exposure of these and many other uncertainties is unsurprising, since much of the “scientific” literature on various child abuse and neglect topics “largely consists of highly selective case studies and small retrospective case series”.515

One of the impacts of this general paucity of scientific evidence in relation to the forensic interpretation of childhood injury is that the various clinical guidelines that have been developed in relation to clinical forensic child protection have inevitably had to rely heavily on expert consensus views. This is often explicitly acknowledged in the various major relevant child abuse guidelines, primarily produced by professional groups in the UK (the RCPCH), 516 USA (the American Academy of Paediatrics), 517 Canada (the Canadian Pediatric Society), 518 and Australia. 519 Furthermore, a recent UK child abuse guideline by NICE focused on general aspects of the recognition of child maltreatment acknowledged that “an extensive review of the literature revealed major deficiencies with the evidence for many of the clinical features of child maltreatment to answer some of the key clinical questions”, 520 so that a Delphi consensus approach had to be used to address some key questions.


517 Nancy D. Kellogg and the Committee on Child Abuse and Neglect 'Evaluation of Suspected Child Physical Abuse' (2007) 119 (6) Pediatrics 1232; this guideline was subsequently reaffirmed by the American Academy of Pediatrics in May 2011.


519 Department of Health, 'Guidelines for Responding to Child Abuse, Neglect and the Impact of Family and Domestic Violence' (Government of Western Australia 2004); Paediatrics & Child Health Division, Genital examinations in girls and young women - a clinical practice guideline (Royal Australasian College of Physicians, August 2009 edn, Royal Australasian College of Physicians Sydney 2009).

Many of the other major guidelines in child protection have not even employed this formal consensus process. Rather, they have relied exclusively on expert round table discussions and agreement, a process highly vulnerable to a number of biases.

**Interim conclusion**

So far in this chapter I have provided a brief overview of the evolution of EBM generally and highlighted the particular challenges that research in children poses to the development of a research base to particularly support the application of EBM in paediatric clinical care. I then described some of the general issues of EBM in relation to child protection work, and the exposure of the relatively weak foundation on which some significant expert opinions had hitherto been based.

In **Section 3** I now turn to critically examine the response to these acknowledged limitations in the evidence base, in particular exploring the processes currently followed in the UK to develop diagnostic clinical guidelines that paediatric forensic practitioners are advised to use in order to justify their expert forensic opinions.

Here I question the validity of the specific evidence hierarchy into which the RCPCH systematic review group suggest published clinical evidence in relation to alleged child abuse should be graded, and which is meant to provide the foundations for relevant paediatric expert opinion evidence in criminal and civil tribunals dealing with alleged child abuse or neglect.

---

521 The most recent RCPCH guideline involved a two-year programme of revision 2011-13, updating the previous version using 38 contributing authors, reviewers and consultees; the process and final version was overseen by a six-person editorial board (Child Protection Companion, n. 509 above, 10-11).

522 A process originally developed in collaborative work by the NSPCC and Cardiff Child Protection Systematic Reviews at Cardiff University. Formerly known as The Cardiff Child PrOtection Systematic REviews project (CORE INFO), the RCPCH has taken over the systematic reviews, rebranding these as Child Protection Evidence systematic reviews; see [https://www.rcpch.ac.uk/improving-child-health/child-protection/child-protection-evidence/child-protection-evidence](https://www.rcpch.ac.uk/improving-child-health/child-protection/child-protection-evidence/child-protection-evidence) (accessed 8 May 2018).
Section 3 - A secure foundation for forensic expert opinion?

In the absence of a definitive gold standard test for abuse, one of the critical aspects and challenges of research in clinical child protection practice is to define specific clinical features that may reliably distinguish between those injuries that are highly likely to have been caused by abuse, and those which have not. Such information, if available, can then be used to inform the assessment of other children with similar features, and can be used to justify a forensic opinion provided to the various legal or quasi-legal tribunals that may be involved in the associated wider child protection process.

While there are certain clinical findings that are generally considered to be explicit and agreed indicators that an injury was inflicted, many other injuries are forensically non-specific and could be caused by accident or abuse. Thus, much of the published child protection literature that suggests that certain injuries may be attributed to assault or abuse are simply raising such a possibility. Here, rather than providing information that might support a definitive diagnostic feature, many of these publications are in reality endorsing a triage process, their intention to ensure that non-specialist clinicians consider the possibility of abuse when they encounter children with certain clinical features so that such cases are referred on for specialist comprehensive forensic assessment.

Specialist assessment of the probability that an injury is abusive

In order to make a considered forensic judgment when faced with a child with a “suspicious” injury, child protection specialists need to rely on an evidence base founded on published studies relevant to the feature(s) under consideration. Generally, current primary research in this area involves comparative case control studies where the incidence of a particular

---

523 Greeley, n. 507 above.
524 For example, patterned bruising where it is obvious that a child had been struck with an implement or hand, ligature marks, or where “natural” bruising does not occur; Speight, n. 496 above, and see also: K. W. Feldman 'Patterned abusive bruises of the buttocks and the pinnae' (1992) 90 (4) Pediatrics 633.
clinical feature – for example various patterns of bruising, types or location of bone
fractures, or particular clusters of cranial injuries in infant head trauma – are contrasted in a
group of children who were considered to have been abused with a similar matched group
whose injuries were considered to be accidental or reflect some other “innocent” cause.

As noted earlier, in relation to UK practice, the most dependable studies currently available
are those originally published by the CORE-INFO group. The majority of these are
retrospective systematic reviews of published case series and comparative studies, involving
detailed critical appraisal of individual studies, with meta-analysis performed where
possible. The group’s aspiration is to calculate the probability of abuse for specific (clusters
of) characteristics that may aid clinicians faced with similar cases, and to provide the
justification for forensic opinions that may be required in related hearings.

However, I suggest that there are significant concerns about the utility of the evidence such
systematic reviews provide in informing expert opinion evidence. Let me illustrate my
concerns by reference to a common and specific forensic question involving a typical review
focussed on childhood fractures.

525 A variety of studies are usually available focussed on various injuries that give rise to concern
about possible child abuse; these include population studies of childhood injuries, case series and
studies comparing features from accidental versus substantiated cases of abuse: (C. Skellern and T.
Health 771.

526 Certain group characteristics must be matched, including age and absence of systemic illness,
while other features may be excluded, perhaps certain types of injury mechanism (such as road
accidents), and (usually) post mortem cases.
“Is this child’s fracture caused by abuse?”

Here the Cardiff group wished to provide guidance on the specificity of certain fracture types and associated features to indicate the likelihood of inflicted injury. Thus they systematically reviewed the published world literature to answer the question “what features differentiate fractures resulting from abuse from those sustained from other causes?”

Their review included 32 comparative studies involving children (up to 18 years old) attending or admitted to hospital with fracture(s). Setting aside the wide age range they included - most abusive fractures involve pre-school children – and despite making efforts to exclude inappropriate studies, the review typifies the challenges faced by investigations in this difficult research area.

Thus, for example, they included studies that involved cases dating as far back as 1956, and many published 20 years or more before their review, thus at a time when many child abuse services were non-existent or in their infancy, and over a period when the clinical criteria for investigating possible child abuse have been evolving, as have statutory investigation and responses.

---

527 Alison M. Kemp and others 'Patterns of skeletal fractures in child abuse: systematic review' (2008) 337 BMJ 859,

528 Such studies now usually restrict cases to children up to four years old.


530 Their excluded review articles; expert opinion; consensus statements; studies of all age groups in which they could not separate data on children; studies judged to be methodologically weak owing to significant bias, confounding factors, case attrition, or incomplete ascertainment or in which the fracture pattern was the primary factor used to define abuse; and studies of outcome, management, or post mortem investigation.

531 An example of such an evolution is the current expectation that chest x-ray evaluation for possible fractured ribs in infants is now mandated to include oblique views of the chest, that may reveal “hidden” fractures not visible on standard views: Christine Weirich Paine and others 'Prevalence of Abuse Among Young Children with Rib Fractures: A Systematic Review' (2016) Pediatr Emerg Care 10.1097/PEC.0000000000000911.
Further most case series are retrospective, and there are inevitable inconsistencies between studies in the clinical criteria and definition of “abuse” used by various study authors to identify and classify cases, with a risk of “detection bias”.

However, the variable geographical source of included studies is arguably more explicitly problematic. Of the 32 studies finally evaluated in relation to childhood fractures, for example, the vast majority were from institutions in the USA (25 USA; 1 USA/ Canada combined), the remainder being made up from three UK studies, two from Australia and one from France. Such an inclusion pattern raises questions about the applicability of these study findings in other populations.

Given that the data are derived from populations with specific ethnicity, age or socio-economic characteristics particular to the study locus, and that the children have also been assessed within a variety of clinical settings with different healthcare arrangements, and in jurisdictions with diverse child protection procedures set within different legal frameworks, the results and interpretation may not be directly applicable to a different population without external validation.

This is problematic; such systematic reviews are critically dependent on the validity of the “abuse” conclusion made by the original authors, and thus the review group acknowledged that comparison of the studies they included was challenging because of the variety of case definitions and selection criteria for subject inclusion which the original authors had used, difficulties that are well recognised in research involving diagnostic systematic reviews more generally.

---

532 Joanne N. Wood and others 'Prevalence of abuse among young children with femur fractures: a systematic review' (2014) 14 BMC Pediatrics 169. Notably in this systematic review of childhood femur fractures, “confirmed or suspected abuse prevalence varied widely across the studies with point estimates ranging from 1.5% - 68.3%” (page 8).

533 S. Mallett and others 'Systematic reviews of diagnostic tests in cancer: review of methods and reporting' (2006) 333 (7565) BMJ 413. When systematic reviews and meta-analyses have to pool
Circular reasoning

A potentially more fundamental issue may undermine such systematic reviews – the possibility that the clinicians in these studies were applying circular reasoning in reaching their conclusions. This is a concern the Cardiff group themselves acknowledge:

‘[I]n the absence of a gold standard diagnostic test for child abuse or neglect, it is essential to minimise the risk of circularity, namely that the authors ‘diagnose’ abuse or neglect simply by the presence of the injury or feature under review’.

In responding to concerns about detection bias, the Cardiff group sought to provide additional selection criteria to ensure that they only analysed cases where abuse had been independently confirmed. Here they developed a specific evidence hierarchy posited to distinguish valid and invalid cases of abuse.

It is potential weaknesses in the foundations of this ranking system that I now seek to expose.

dissimilar subjects, various statistical strategies may be used to reduce the impact or expose the risk of such aggregations, but inevitably the strength of any conclusions must be undermined.

C M Sullivan, 'Child abuse and the legal system: the orthopaedic surgeon's role in diagnosis' (2011) 469(3) Clin Orthop Relat Res 768. The issue of circular reasoning in child abuse diagnosis has been of longstanding concern. Such a fallacious line of reasoning, where a conclusion is based on material already assumed in the argument, has been a prominent problem in published studies in many areas of child protection research, and has included issue of conflation of independent indicators. See, for example, Shady N. Hayek and others 'The Efficacy of Hair and Urine Toxicology Screening on the Detection of Child Abuse by Burning' (2009) 30 (4) J Burn Care Res 587, where the finding of traces of illicit drugs detected in hair samples were used to later reinterpret forensically non-specific injuries as abusive in origin.

The Cardiff ranking system – a valid epistemic warrant for abusive injury?

Given the issue of uncertainty about the validity of abuse diagnosis in the various studies they reviewed, the Cardiff group developed “a unique ranking system” to try to ensure that cases labelled as abusive or non-abusive had some level of independent confirmation of that status, and “[were] not based entirely on the clinical feature of interest”.

Accordingly, they ranked individual cases according to the recorded conclusion of any local associated child protection fact-finding tribunals, or conversely, whether abuse had been dependably excluded in the non-abused group (Table 2 below).

<table>
<thead>
<tr>
<th>Rank</th>
<th>Definitions of abuse</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Abuse confirmed at case conference or family, civil, or criminal court proceedings; admitted by perpetrator; or where the abuse was independently witnessed</td>
</tr>
<tr>
<td>2</td>
<td>Abuse confirmed by stated criteria including multidisciplinary assessment</td>
</tr>
<tr>
<td>3</td>
<td>Abuse defined by stated criteria</td>
</tr>
<tr>
<td>4</td>
<td>Abuse stated but no supporting detail given</td>
</tr>
<tr>
<td>5</td>
<td>Suspected abuse</td>
</tr>
</tbody>
</table>

Exclusion of abuse in non-abused group

<table>
<thead>
<tr>
<th>Rank</th>
<th>Definitions of abuse</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Abuse actively excluded by stated criteria; witnessed traumatic cause; or confirmed organic cause</td>
</tr>
<tr>
<td>2</td>
<td>Exclusion of abuse implicit in case definition used or stated criteria given</td>
</tr>
<tr>
<td>3</td>
<td>Aetiology of non-abuse group merely stated</td>
</tr>
<tr>
<td>4</td>
<td>Aetiology of comparison not stated</td>
</tr>
</tbody>
</table>

Table 2: Ranking system to assign cases to abused or not abused within included studies.

Yet while these authors have strongly promoted this grading system, and it has apparently been taken up or adapted by others, I suggest that there are some particular questions about its validity that require to be addressed. In doing so, for simplicity, I will focus on the criteria they posit as indicating the highest standard of confirmation of abuse. Thus, in rank 1, they accept that an injury was abusive if one of the criteria below is met:

- Abuse was the conclusion of a multidisciplinary assessment of the case (case conference), taking into account “social and historical factors beyond the presenting injury”
- Abuse was the conclusion of a formal judicial process - criminal, family or civil
- A perpetrator admitted to inflicting the injury
- If the abusive event was independently witnessed.

Is their ranking system fit for purpose?

The fundamental purpose of this ranking system must surely be to ensure that opinion evidence paediatric experts provide to the courts and other relevant tribunals is based on valid science that is evidentially reliable. And there is no doubt that the aim of reducing the risk of circular reasoning – “since we always say that this type of fracture is caused by abuse, this must be another case” – is a laudable goal, since such circularity inevitably erodes the validity and forensic applicability of such data. However, I suggest that not only is there an argument that it does not really achieve that aim, but more importantly, that there are more fundamental concerns about these criteria that may seriously undermine the reliability of expert opinion evidence based on this approach.

---

537 For example, Pierce, n. 531 above. See also A. M. Naughton and others 'Emotional, behavioral, and developmental features indicative of neglect or emotional abuse in preschool children: a systematic review' (2013) 167 (8) JAMA Pediatr 769; M. C. Pierce and others 'Bruising characteristics discriminating physical child abuse from accidental trauma' (2010) 125 (1) Pediatrics 67; S. J. Piteau and others 'Clinical and radiographic characteristics associated with abusive and nonabusive head trauma: a systematic review' (2012) 130 (2) Pediatrics 315.

538 Sabine Maguire and others 'What does the recent literature add to the identification and investigation of fractures in child abuse: an overview of review updates 2005-2013' (2013) 8 (5) Evid Based Child Health 2046.
I challenge the basis of the group’s assertion that the outcome of judicial hearings, whether criminal or civil, may be considered a proxy for the reliability of the science underlying the opinion evidence provided at such tribunals. This was exactly the basis on which Dr Colin Paterson justified his claims of the validity of the scientifically discredited “temporary brittle bone disease” diagnosis to explain infant fractures.\(^{539}\)

I described the case conference process in Chapter 1, emphasising the relatively undefined burden of proof involved, that the issues were primarily focused on the risk of future significant harm, and where many non-clinical factors, social and otherwise, may influence, and perhaps dominate, the ultimate conclusion.

My argument here in relation to child protection forensic evidence is that when certain clinical features are presented by medical experts to legal or quasi-legal decision-making tribunals as evidence that a child had sustained inflicted injury, that clinical evidence should be able to stand alone and be assessed on its own merits. Non-clinical “risk” or other factors cannot be used to inform or validate the expert’s interpretation and conclusions.

Such tribunals, whether at the criminal, family court or local authority level, are not in a position to determine independently the validity or otherwise of the scientific or clinical foundation of forensic opinion presented to them, for as Haack has pointed out, “there is usually no way of judging scientific evidence without substantive knowledge of the field in question”,\(^ {540}\) except in indirect terms.\(^ {541}\)

---

539 Paterson and Monk, n. 92 above.
540 Susan Haack, Defending science - within reason: between scientism and cynicism (Prometheus Books Amherst 2007) 236.
541 This may be by confirming that, among other criteria of reliability, any proffered forensic evidence is accepted within the relevant scientific community, has been subject to peer review, and stands up to logical analysis. These issues, among others, reflect the issue of judicial gatekeeping to ensure evidential reliability. I deal with this issue in detail in Chapter 7.
While it may be argued that, especially in a criminal trial context with its adversarial approach and high burden of proof,\(^{542}\) the court’s careful scrutiny of the forensic and other information adduced may provide an appropriately challenging setting in which to expose weak or unjustified scientific testimony, it is clear that in such settings exposure of unreliable scientific evidence is not guaranteed.\(^{543}\) Indeed, as I discuss later (Chapter 7), the judiciary have explicitly warned against the use of court judgments to validate the scientific reliability of expert evidence.

Many factors come in to play within both the adversarial system of a criminal trial and the more inquisitorial child protection civil hearing that may impact on and restrict forensic medical and other scientific testimony,\(^{544}\) and where, by definition, the evidence the parties choose to adduce is partial and incomplete. In all these circumstances, the outcomes of such hearings are very particular to the individual case before the court and are not generalisable to other cases, except in matters of law. As has been clearly stated in the context of scientific evidence in court:

> There are important differences between the quest for truth in the courtroom and the quest for truth in the laboratory… Scientific conclusions are subject to perpetual revision. Law, on the other hand, must resolve disputes finally and quickly.\(^{545}\)

---

\(^{542}\) I acknowledge that the simple polarised dichotomy of “adversarial” versus “inquisitorial” processes implied here does not reflect the reality, although the different burdens of proof and the focus of the tribunals differ significantly.

\(^{543}\) Gary Edmond and Mehera San Roque 'The Cool Crucible: Forensic Science and the Frailty of the Criminal Trial' (2012) 1 (1) Curr Iss Crim Just 51. Edmond and San Roque argue that much expert evidence adduced at criminal trials is either unreliable or of unknown reliability and that, despite the adversarial nature of the process, it is not adequately challenged. They suggest that the courts need a more cynical approach to scientific evidence and more robust safeguards to ensure that such trials are truly fair.

\(^{544}\) The availability and funding of expert witnesses (especially for the defence), the knowledge, impartiality and communication skills of the expert, the abilities of the legal actors to understand and deal with the scientific and clinical issues effectively, and the competence of the triers of fact to evaluate the scientific and clinical evidence are just some of the factors involved. For a useful overview of some of the issues at play here, see, generally, Emma Cunliffe 'Independence, reliability and expert testimony in criminal trials' (2013) 45 (3) Aust J Forensic Sci 284.

Legal fact-finding is about delivering justice (legal truth) and minimising error. This is not “truth” from a scientific perspective, and these considerations seriously undermine the Cardiff group’s implied suggestion that judicial settings are appropriate for open scientific discussion, debate and resolution of clinical or scientific controversy as implied within their ranking system.

Indeed, we know this approach cannot be relied on, for there is a history of judicial decisions that, under the Cardiff group’s criteria, have validated experienced expert witness testimony espousing confident opinions on the interpretation of scientific/ forensic features in child abuse cases where later such evidence has been shown to be unreliable.

From this, I suggest that, despite its superficial attraction, the use of judicial or multidisciplinary hearing decisions to validate the reliability of clinical forensic diagnosis is fundamentally flawed. Rather, I suggest that it must be the responsibility of the biomedical community to establish the validity of such clinical evidence and resolve competing interpretations and uncertainties.

In reality, even the proponents of the classification system themselves seem to acknowledge the relative vulnerability of the evidence base they are valiantly working hard to develop. In all their published systematic reviews, there is explicit advice that ‘if you have a specific clinical case, we strongly recommend you read all of the relevant references as cited and look for additional material published outside our search dates’.

I now turn to consider a separate key criterion that the Cardiff group have posited as validating the reliability of certain clinical features as secure markers of abusive injury – confession evidence or admissions of abuse.

---

Confession evidence as a validation of clinical features of abusive injury?

In this final part of Section 3, I want to examine another criterion posited by the Cardiff group as a secure independent indicator that a specific injury may be confidently interpreted as abusive in origin – a confession or admission by an alleged perpetrator. At first blush, this might seem like a non-issue because at an individual case level, a confession is likely to be extremely influential in determining the legal outcome of a case. But I suggest that as a criterion to validate forensic opinion as to the origin of a particular injury, it may not be so simple. There are two relevant potential problems with this criterion that I want to address here. The first takes me back to the questions of cognitive bias and circularity that I discussed earlier.

Confessions are uniquely potent in their ability to induce bias and exert a very powerful effect at all stages of a criminal case, influencing the process of investigation and evidence gathering, prosecution and conviction. This is no different in child protection cases, where a confession will almost always result in a criminal charge and frequently a subsequent conviction. As I already noted in Chapter 3, it is clear that priming information indicating that a suspect has admitted to a crime significantly corrupts objective interpretation within

---

547 In discussing this issue, while I recognise that in practice the two terms are usually used in somewhat different settings, here I will consider “confession” and “admission” as equivalent and use the term “confession” for both. Cobley has pointed out that, in relation to investigation of possible child abuse, parallel investigative processes by police and social work services are in train, albeit with somewhat different foci. These services employ different terminology, with “admissions or acknowledgments of responsibility [to social workers during the child protection investigation] becoming “confessions” in the criminal justice system”; see Cobley, n. 13 above.


549 Saul M. Kassin, Daniel Bogart and Jacqueline Kerner 'Confessions that Corrupt: Evidence from the DNA Exoneration Case Files' (2012) 23 (1) Psychol Sci 41.

550 As I already noted in Chapter 3, it is clear that priming information indicating that a suspect has admitted to a crime significantly corrupts objective interpretation within
Confessions or admissions create their own set of confirmatory and cross-contaminating biases, in particular confirmation bias. Such effects of confessions have been reported, for example, in forensic expert handwriting opinions, the interpretation of DNA mixtures, and may even alter the exculpatory evidence of alibi witnesses. Thus their use to validate the likelihood that certain clinical features suggest abuse must be treated with caution.

Let me illustrate this point by reverting to the reality of clinical forensic assessment.

Sometimes when faced with an injured child, forensic assessment is (relatively) straightforward. A young child presents with multiple injuries – let us say a combination of facial bruising, a skull fracture, and radiological evidence of an old healing fracture of the arm. There is no history of injury. In the absence of bleeding disorders and underlying bone fragility, the only logical opinion is that this child has been subjected to multiple assaults. Any associated confession is irrelevant to the forensic interpretation of such a case, because abusive injury is the only logical explanation for the constellation of injuries.

But, as I also noted earlier, many injuries encountered by paediatricians are forensically non-specific, and the situation becomes somewhat different when faced with such a “grey-case”. Here knowledge of a confession of assault by the carer may potentially have a

554 Dror and Hampikian, n. 551 above.
555 S. B. Marion and others 'Lost proof of innocence: The impact of confessions on alibi witnesses' (2016) 40 (1) Law Hum Behav 65.
major influence on the paediatrician’s forensic interpretation of the likely origin of the 
bruising, for example, irrespective of the specificity of the bruise characteristics as 
objectively indicative of inflicted injury.

There is very little published evidence to inform paediatric forensic opinion in relation to 
confession evidence, and generally such risks are not acknowledged within the paediatric 
forensic community. 557 With the exception of reported confessions in cases of alleged shaken 
baby syndrome (SBS), there have been no comparative studies of perpetrator admissions and 
associated medical evidence in other types of abuse and neglect. 558 In alleged SBS, the utility 
of confession evidence has become a highly contentious issue in the extremely polarised and 
apparently irreconcilable professional arguments whether or not violent shaking is the only 
plausible cause of serious brain and other injuries in small children (see later, Chapter 7). 
While some of the individual case descriptions of alleged perpetrator assaults are very 
detailed and appear convincing, 560 a variety of difficulties with the interpretation of such 
studies are evident, in particular the inherent challenges of retrospective case analysis, and 
the factual truthfulness of some of the confessions are uncertain. 561 This leads me on to a

557 In a recent study of bruising characteristics in children assessed because of suspected physical 
abuse, the clinical details of all those children where there was a confession of abuse, albeit a small 
number, were included in the analysis; see A M Kemp and others, 'Bruising in children who are 
assessed for suspected physical abuse' (2014) 99(2) Arch Dis Child 108.
558 Dean Biron 'Dialectics, Infant Shaking, and Perpetrator Statements in Child Maltreatment' (2014) 
29 (1) J Fam Violence 89. While Flaherty dealt with a range of abusive injuries in her study 
comparing initial histories with later confession details, her focus was on clues that might be gleaned 
from the initial history rather than clinical forensic features: Emalee G. Flaherty 'Analysis of caretaker 
histories in abuse: Comparing initial histories with subsequent confessions' (2006) 30 (7) Child Abuse 
Negl 789.
559 Catherine Adamsbaum and others 'Abusive Head Trauma: Judicial Admissions Highlight Violent 
and Repetitive Shaking' (2010) 126 (3) Pediatrics 546; Marc De Leeuw and others 'Confessed 
Abusive Blunt Head Trauma' (2013) 34 (2) Am J Forensic Med Pathol 130; Jan E. Leestma 'Case 
Forensic Med Pathol 199; S. P. Starling and others 'Analysis of perpetrator admissions to inflicted 
and others 'Confessed abuse versus witnessed accidents in infants: comparison of clinical, 
560 Erica Bell, Michelle Shouldice and Alex V. Levin 'Abusive head trauma: A perpetrator confesses' 
561 Leestma, n. 559 above.
separate question that I wish to raise in relation to the use of confessions as a strong indicator of the likelihood that particular injuries as abusive – plea bargaining.

**Plea-bargaining and false confessions**

One of the issues raised as a challenge to the validity of confession evidence is the spectre of false or coerced confessions. The frequency of coerced confessions in child abuse cases is unknown, although they undoubtedly happen. While the possibility of false confessions cannot be discounted, a separate issue that also comes into play here, especially, but not exclusively, in the USA, is plea-bargaining.

It is clear that in the face of the relatively punitive nature of the criminal justice systems in some US jurisdictions, some parents accused of murdering or seriously injuring their children confess in the face of an offer of a more lenient sentence, sometimes despite clear defence medical evidence of an alternative natural explanation for the clinical findings.

Plea-bargaining in the face of a severe sentence threat was also an issue noted in the Goudge

---


565 Jago Russell and Nancy Hollander 'The Disappearing Trial' (2017) 8 (3) NJECL 309. Information on informal plea negotiations in UK courts is unknown, although discussions between prosecution and defence in relation to guilty pleas to some or alternative charges undoubtedly occur; see Attorney General’s written answer on the incidence of plea bargaining in English Courts, House of Commons, 15 February 2017: http://www.parliament.uk/business/publications/written-questions-answers-statements/written-question/Commmons/2017-01-25/61736/ (accessed 13 May 2018).


567 As reported in the case of Tamara Sawyer. See Elaine Whitfield Sharp 'A matter of gravity' (2004) (Winter) The Warrior 34. See also Tuerkheimer, n. 110 above, at 99.
Inquiry into paediatric pathology in Ontario. ⁵⁶⁸ How such processes might influence the database of confession evidence that may be used within systematic reviews of relevant literature is uncertain, but the influence cannot be totally ignored, given the preponderance of US cases that have populated previous child abuse systematic review studies. ⁵⁶⁹

Let me return to the key question that this section considered, whether the proffered ranking system to validate a forensic conclusion that a particular injury pattern is likely to have been inflicted is sufficiently robust and stands up to scrutiny. Acknowledging the major contribution made by the Cardiff group in working hard to identify and interpret which sound clinical evidence can be relied on in child protection cases while exposing the false truths of the past, one has to conclude that their proffered ranking system to assign an indicative value to the likelihood of abuse in the face of various patterns of injury must be treated with significant caution as a foundation for expert opinion.

**Conclusion**

In the previous two chapters I dealt with the process of clinical forensic diagnosis, and explored the potential for various cognitive biases, emotion and other influences, conscious and unconscious, to affect objective interpretation and sound forensic conclusions. Putting it colloquially those chapters emphasised the reality that a fundamental aspect of being human was that the heart might rule the head irrespective of the kind of decisions to be made. I argued that paediatricians working in child protection are as susceptible to such effects as anybody else. I suggested that in clinical practice, application of a formal evidence-based approach might mitigate such threats to expert objectivity.

In this chapter I described the evolution of EBM, explored some of the key elements on which it depends, and emphasised that the philosophy and practice of EBM, albeit with

⁵⁶⁸ Goudge, n. 101 above.
⁵⁶⁹ For example, Kemp et al., n. 527 above.
caveats and concerns, is now inextricably embedded within all areas of clinical care, including paediatric practice.

What this chapter also acknowledged was that, for a number of reasons, there were delays in the adoption of an evidence-based approach to clinical child protection work, involving both the process of recognition of inflicted injury and the related provision of expert opinion evidence.

Over the last two decades, major efforts to respond to that deficiency have been made. This work has usefully exposed out-dated dogma, but also revealed that there is a relative paucity of robust objective research in child protection on which to base sound evidence-based forensic interpretations, so that there is great reliance on professional consensus.

It is clear that a very active and impressive programme of research is continuing, but, as I have argued here, there are questions about the evidence hierarchy that underpins the systematic review process on which such professional guidance is based. Rightly such data comes with an explicit caution that each case needs to be assessed individually and with a practitioner’s full understanding of all the relevant literature.

What such studies purport to offer forensic paediatric practitioners are probabilistic interpretations of the likelihood that specific injuries in children may be abusive in origin. But the issue for the paediatric expert is how such likelihoods should be applied to an individual case requiring forensic assessment, and when they stand in the witness box providing an expert opinion on that case.

If the systematic review data suggests, for example, that abuse was the origin of 80% of cases of a particular type of fracture, how does the practitioner use that information in formulating an opinion about the individual child they encounter with a suspicious injury? As Dawid has noted, this raises important questions about how we should collect and use such data to inform inferences and decisions about an individual case – what he terms the
“group-to-individual” problem. Like the question the defence QC asked the jury to consider in Hainey – whether Declan was an example of a rare but accepted case of SIDS in an older child - might that child with a “suspicious” fracture come from the 20% of children whose fractures have an “innocent” origin?

And what if the systematic review suggests that a particular fracture has an average or low specificity for abuse, but there is no explanation provided or the explanation offered seems implausible? How should such separate elements be integrated into a defensible expert opinion?

Given the nature of the evidence base in relation to the diagnosis of abusive injury that I have discussed here, and its strong dependency on professional consensus, I suggest that two key questions emerge from this analysis:

The first is that, given the typical uncertainties I have outlined above, how should paediatricians providing expert opinion evidence express their opinion and a level of confidence or certainty in the conclusions they provide in a particular case? This I address next, in Chapter 6.

Second, and from a different perspective, given the nature of this evidence base, how can the courts ensure that only reliable opinion evidence from skilled witnesses is put before them in cases of alleged abusive injury? I consider this issue later, in Chapter 7.

---

Chapter 6: Semantics and ambiguities in expressing expert opinions

Introduction

[If] opinions are not understood, the justice system operates on misinformation…The innocent should not be charged or convicted, or the guilty go free, on the basis of expert opinions that are misunderstood.\(^{571}\)

The above quotation from the Goudge report, although focused on paediatric forensic pathology, is no less appropriate for paediatricians providing forensic reports in cases of alleged child abuse, nor indeed any other form of expert scientific testimony, for the criminal justice system depends on clarity of explanation and dependable expressions of certainty in the expert guidance it receives.\(^{572}\)

Whether in their written reports or oral evidence, a key responsibility of a paediatric (or any other) expert providing opinion evidence in alleged child abuse is not only to explain the scientific basis for their opinion, but also to express the degree of confidence or certainty with which they have concluded that a child’s clinical features are, or are not, indicative of an abusive injury.

Here there are important linguistic and semantic issues to consider, for there is clear empirical evidence demonstrating that fact-finders’ interpretations of forensic expert evidence is significantly influenced by how that information is presented.\(^{573}\) It is this key aspect of paediatric forensic expertise – the formulation and phrasing of their opinions and conclusions – and the potential frailties therein, that I consider in this chapter.

There are broadly two interrelated key elements that I wish to focus on here. First, the language with which an expert expresses the substance and conclusion that makes up their opinion – in fact, their ‘forensic diagnosis’, and secondly, how they may express a level of

---

\(^{571}\) Goudge, n. 101 above, page 406.

\(^{572}\) É O'Brien et al., n. 69 above.

confidence or degree of certainty to qualify that conclusion, in effect matching the strength of the conclusion to the strength of the science.

I have subdivided this chapter into three sections.

First, in Section 1, in order to put the issues relevant to the expert paediatrician into its broader context, I discuss how forensic experts more generally express their confidence or degree of certainty in their evaluative conclusions, observing that as the scientific foundation of many long-established forensic techniques has come into question, this has been reflected in changed expectations of the current forensic community and its users as to how evaluative expert opinions are stated. Here I track the change from a categorical to a probabilistic approach that reflects the emergence and application of the Bayesian framework paradigm now considered most appropriate in the analysis and expression of uncertainties associated with much forensic work, before briefly discussing the response of the courts to that approach.

In Section 2 I consider the issue of the language or wording used by paediatricians in expert forensic reports, and illustrate some of the practical challenges involved.

Finally, in Section 3, I discuss the difficulties in applying a Bayesian or formal probabilistic approach to paediatric forensic evidence.

The key conclusion of this chapter is that, except in narrow circumstances, the epistemic mode of reasoning in paediatric forensic clinical assessment does not fit well with a statistical approach. Thus, in the absence of an ability to derive a numerical scale and linked ordinal phrases with which to express the certainty of their conclusion, in formulating and phrasing their opinions the paediatric expert must depend on subjective belief-type

---

574 For the sake of clarity, I note that a numerical or interval scale usually represents a graded sequence of specific values, for example, the scale on a thermometer, while an ordinal scale is usually a series more informal graded verbal equivalents roughly congruent with the numerical range. Thus we may describe an outdoor thermometer reading of 2°C as “very cold”, and that of 30°C as “hot”.
interpretations of probability, and find a way of expressing the strength of their opinions that the courts (and other experts) understand. The current absence of an agreed verbal scale and standards with which to consistently qualify such subjective interpretations represents a significant frailty intrinsic to paediatric expert forensic evidence.

**Section 1: Quantifying expressions of certainty in expert opinions**

As I noted above, this issue of forensic expert (mis)communication with the courts and the justice system more generally goes far beyond the paediatric realm, so in this section I briefly leave the relatively small, albeit challenging, domain of forensic child abuse assessment for the bigger world of forensic science more generally. Here difficulties around the communication of scientific truth, opinions and the impact on legal fact-finding have recently been particularly acutely exposed, and there is much of relevance for an expert paediatrician to observe and reflect on.

**Emerging cracks in the foundations of much forensic science**

In recent years, technological developments spanning many scientific domains have allowed novel techniques to be applied to the investigation of a variety of forensic problems. While the integration of new techniques may in themselves pose challenges for the criminal justice system, the application of novel scientific developments to forensic procedures have also raised questions about the scientific validity and reliability of some previously long-established forensic practices. Notably, the application of DNA profiling, while itself representing a revolutionary forensic tool, exposed a significant number of unsafe

575 For example, applying forensic anthropology techniques to examination of the digital images of abuser’s hands to identify perpetrators of online paedophile abuse (Graham Jackson and Sue Black ‘Use of data to inform expert evaluative opinion in the comparison of hand images: the importance of scars’ (2014) 128 (3) Int J Legal Med 555.

convictions across many jurisdictions, and raised major questions about the reliability of a number of previously trusted forensic techniques.

In the face of high-profile attribution blunders, even the dependability of traditional fingerprint analysis, the emblematic identification technique of criminal investigation, has been exposed as vulnerable to error. Concerns here focussed not only on the highly subjective nature of such interpretation and the potential for contextual bias that I previously discussed in Chapter 3, but also on examiner claims that they could categorically identify an individual, and that their “error rate for fingerprint comparison [was] ‘essentially zero’.”

While such confident assertions are now rejected as “not scientifically plausible”, these claims also reflected a systemic issue rather than individual failing: as late as 2010, US fingerprint examiners risked a professional misconduct investigation if they described a latent print identification as “possible, probable or likely”, rather than “certain”.


Notably the flawed attributions of fingerprints by the Scottish Criminal Records Office (SCRO) of Scottish police officer Shirley McKie, and by the FBI of Brandon Mayfield in relation to the Madrid train bombings: see, for example, the overview provided by Simon A. Cole 'The Prevalence and Potential Causes of Wrongful Conviction by Fingerprint Evidence' (2006) 37 Golden Gate U L Rev 39.

More correctly termed forensic friction-ridge identification. See, for example, R N Haber and L Haber 'Experimental results of fingerprint comparison validity and reliability: A review and critical analysis' (2014) 54 (5) Sci Jus 375; B T Ulery and others, 'Accuracy and reliability of forensic latent fingerprint decisions', vol 108 (2011) 7733. While the criminal justice system relies on the skill of latent print examiners as expert witnesses, currently there is no generally accepted objective measure to assess that skill.

See NAS report, n. 221 above, 103.

This expectation of certainty went much further than latent fingerprint analysis. Forensic scientists in many fields presented their results as certainties,583 and as I discuss later, until very recently, even medical experts providing forensic opinion evidence in the US courts were expected to confirm the (relative) certainty of their conclusions.

Thus, unsafe convictions have been based not only on unreliable expert evidence due to untrustworthy science, but also on how experts express evaluative opinions based on this science to the courts.584

**Expert forensic opinions – bridging the science - law interface**

There is often a communication gap (and even a philosophical difference) between scientists and forensic science providers who develop forensic methods, and judges and juries who deal with evidence based on such methods… In principle the scientific aspects of forensic evidence should be uncontroversial but, in practice, court time and effort can be spent disputing what are essentially scientific issues.585

Of course, it would be wrong to suggest that these were one-sided problems. They also relate to the expectations and understanding of the criminal (and civil) justice systems, and it seems clear that such cross-cultural communication is often problematic. As Black and Nic Daeid acknowledge, it is impossible for experts to provide opinion evidence that is “consistently and unerringly accurate, precise, informed, impartial, definitive and right”.586 Yet while uncertainty is accepted, even expected, by the science community, qualified opinions do not fit well with the binary environment of the legal world. And of course, whatever the expert’s

---

583 For example, providing a single point value for a blood alcohol level in a drink driving case, without providing the confidence intervals to demonstrate that all scientific analysis has a margin of error (Imwinkelried, n. 582 above).
585 Bernard Silverman, *Research and Development in Forensic Science: A Review* (UK Home Office 2011) 4.1 - 4.2, page 10. While the Silverman Review specifically excluded forensic medicine, this comment is no less relevant to that forensic area.
586 Black and Nic Daeid, n. 69 above, at 2.
original opinion, skilled courtroom questioning may sometimes leave experts feeling frustrated that their words are being twisted and their opinions distorted.\(^{587}\)

Here it is not just that scientific and legal fact-finding are somewhat incongruent. As recently highlighted by the English Law Commission,\(^{588}\) a lack of scientific training among the legal actors involved in criminal trials (and other fact-finding tribunals) also poses significant difficulties in courts’ assessments of the reliability, relevance and weight of expert forensic opinion evidence in order to ensure there can be a confident acceptance of that advice to support an accurate fact-finding process.\(^{589}\)

Given such limitations, it not surprising that historically the courts have accepted much forensic science evidence with little challenge. Yet such unquestioned reception of expert opinion may threaten the integrity of the legal process,\(^{590}\) and there is ample empirical evidence that many miscarriages of justice have resulted from flawed but unchallenged forensic science and expert medical evidence.\(^{591}\)

\(^{587}\) Henderson and Seymour, n. 62 above.

\(^{588}\) Law Commission (2011), n. 101 above, [1.43].

\(^{589}\) G Edmond and M San Roque, 'The Cool Crucible: Forensic Science and the Frailty of the Criminal Trial' (2012) 1(1) Curr Iss Crim Just 51. A separate but related issue that arises here is the interpretation of scientific reports by non-scientists - prosecution authorities and lawyers - without the benefit of detailed discussion and input from the forensic scientists themselves. There is concern, for example, that the recent introduction of Streamlined Forensic Reporting focussed on rapid availability of preliminary results of forensic DNA analysis to provide early information and accelerate case management may result in the unqualified acceptance of very sparse preliminary results that may be much less definitive if the true limitations of the results were known. See http://www.cps.gov.uk/legal/s_to_u/scientific_evidence/sfr_guidance_and_toolkit/ (accessed 28 April 2018).


\(^{591}\) This has been exposed by the US Innocence Projects to multiple areas of forensic testimony, although identification evidence in its various forms, has been a very prominent source of flawed forensic evidence: http://www.innocenceproject.org (accessed 28 April 2018). See also, generally: B Sangha, K Roach and R Moles, 'Forensic investigations and miscarriages of justice: the rhetoric meets the reality' (Irwin Law 2010) 447, especially chaps. 7, 8 and 9; Erica Beecher-Monas, Evaluating Scientific Evidence: An Interdisciplinary Framework for Intellectual Due Process (Law in Context Series, Cambridge University Press New York 2007) 254:“In numerous jurisdictions, courts are circumventing rigorous analysis and admitting forensic testimony that is shockingly unscientific…[A]s a result convictions continue to be based on evidence for which the scientific foundation is highly suspect”, at page 95.
These problems underscore the need to ensure that expert forensic scientific evidence put to the courts, including paediatric expert opinion, is both reliable and objective, and that any limitations of the evidence and any associated opinion are explicitly disclosed.\(^{592}\)

Recent initiatives have begun to address the need for judicial education in the various domains of forensic science,\(^ {593}\) and the forensic science community has also recognised the need to provide evaluative forensic opinions that more explicitly reflect the epistemic status of the science behind those opinions. Put simply, the science and the strength of opinion should match.

**From categorical to probabilistic opinions**

In recognising that conclusions of absolute certainty do not reflect the reality of science,\(^ {594}\) in many forensic domains, when providing an evaluative report to the courts, experts are now no longer expected to report categorical or didactic opinions.\(^ {595}\) Rather, acknowledging that all scientific evidence is essentially probabilistic in nature,\(^ {596}\) the forensic community has adopted a more permissive approach to the expression of expert reasoning under

---


594 See NAS report, n. 221 above.

595 While it might be reasonably argued that forensic scientists making such categorical statements should be severely criticised, in some forensic domains such as fingerprint identification, exaggerated claims and categorical statements of positive identification continue to be made in some jurisdictions; see Edmond (2014), n. 81 above.

uncertainty, and a variety of essentially statistical principles have been applied to expert forensic interpretation and evaluation of evidential material and how experts’ conclusions are presented. Thus, with a degree of support from the legal establishment (but see below), many forensic disciplines in the UK and in Europe more generally have adopted a probabilistic approach. Bayesian principles and the likelihood ratio (LR) now dominate forensic scientists’ interpretations and evaluative reporting of expert evidence.

Reflecting this change, a set of standards published by the Association of Forensic Science Providers (AFSP) advised members that their expert opinions should be based upon the estimation of an LR. At a practical level, the output of such evaluations generally provides an expression of the degree to which the forensic expert’s findings support one of two competing propositions, usually expressed in terms of the magnitude of the relevant LR. In a criminal prosecution context, the comparison would typically involve comparing the “prosecutor’s hypothesis” with an alternative “defence’s hypothesis”.

---


598 For a brief but readable overview of this approach, see Craig Adam, Forensic Evidence in Court: Evaluation and Scientific Opinion (John Wiley & Sons, Ltd Chichester 2016a) 315, chapters 7 & 8.


603 A likelihood ratio is a measure of the relative strength of support that the expert considers a particular finding gives to one proposition against a stated alternative.
While this claimed normative way for an expert to express degrees of belief regarding the claim in question has not been fully embraced in other common law jurisdictions, such as the US, and rejected by some commentators, this probabilistic approach is now regarded as the norm in UK forensic science practice. It seems clear that generally the days of subjective “when you’ve been doing the job as long as I have” based forensic opinions are over.

However, as I discuss below, this does not mean that subjectivist belief-type interpretations of probability and related opinions are unacceptable, for such an approach may well be the best option depending on the forensic setting and data available. Rather there is now a clear expectation that the basis for such opinions should be explicit.

Having briefly described the move to a relatively formal statistical approach to forensic science interpretation, I do not intend to provide any further review of the use of Bayesian forensic analysis or related approaches here, for which an extensive literature is available.

---

605 Where a “frequentist” methodology is usually adopted. For an overview of this approach, see, for example, J. M. Curran, 'The Frequentist Approach to Forensic Evidence Interpretation' in Jay A. Siegel, Pekka J. Saukko and Max M. Houck (eds), Encyclopedia of Forensic Sciences (Second Edition) (Academic Press 2013) 286
607 Evett, n. 596, above.
Verbal scales to reflect expert certainties in their conclusions

However, it is important to note that, albeit derived by a statistical process, there has been a strong consensus that LRs derived from such an approach should generally be expressed using a verbal scale rather than presented numerically. This partly reflected concern about the general public (and thus jurors) understanding of probability statistics, and the potential for their misinterpretation, concerns also reflected in judicial resistance, albeit with particular exceptions, to such numerical expressions of expert opinions (see below).

Reflecting this expectation, the AFSP developed an ordinal scale to guide the conversion of numerical LRs into verbal expressions. This scale has been widely adopted (Table 3), although alternative qualitative scales have been developed by other forensic agencies, and by specific disciplines such as forensic odontologists.

<table>
<thead>
<tr>
<th>Value of likelihood ratio</th>
<th>Verbal equivalent</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;1–10</td>
<td>Weak support for proposition</td>
</tr>
<tr>
<td>10–100</td>
<td>Moderate support</td>
</tr>
<tr>
<td>100–1,000</td>
<td>Moderately strong support</td>
</tr>
<tr>
<td>1,000–10,000</td>
<td>Strong support</td>
</tr>
<tr>
<td>10,000–1,000,000</td>
<td>Very strong</td>
</tr>
<tr>
<td>&gt;1,000,000</td>
<td>Extremely strong</td>
</tr>
</tbody>
</table>

Table 3: Standards for numerical and equivalent verbal expressions of likelihood ratios.

---

611 See, for example, Gerd Gigerenzer and others "A 30% Chance of Rain Tomorrow?" How Does the Public Understand Probabilistic Weather Forecasts? (2005) 25 (3) Risk Anal 623, revealing widespread public misunderstanding about probabilistic predictions. This is perhaps understandable, given that in the real world the concepts of “probability” go much wider than the competing Frequentist and Bayesian views of statistical science; see Philip Dawid 'On individual risk' (2015) Synthese 1.
612 For example, by the Swedish National Laboratory of Forensic Science; see Anders Nørdgaard and others 'Scale of conclusions for the value of evidence' (2012) 11 (1) Law Prob & Ris 1.
However, despite efforts to provide consistency in the verbal expressions of interpretative confidence that such scales are intended to convey, introducing such verbal scales are not without problems for practitioners, and it has become clear that there may be significant differences in what experts mean by these qualitative phrases, and how lawyers, judges and juries interpret their meaning. Thus, a layperson’s understanding of a technical term such as “match” – in reality, weak in the context of forensic evidential evaluation, but perceived as highly probative from a lay cultural perspective - typify the risk of misinterpretation.

What is clear here is that the appropriate and effective use of probabilistic evidence is challenging, the format chosen to present such evidence is important, and that despite attempts to standardise evaluative qualitative phrases, lay individuals may interpret such expressions very differently.

It is also important to recognise that many comparative (and other) forensic tasks involve a degree of subjective judgment. Thus it must be appreciated that when experts in such fields derive LRs and convert them into an associated verbal expression, it is not ‘the’ (objective) probability that the forensic expert is reporting, but rather, they are an expression of the expert’s certainty (or otherwise) in that belief. Thus, despite the objective gloss that numbers may give, subjective influences may undermine the “truth” of an opinion.

614 Raymond Marquis and others 'Discussion on how to implement a verbal scale in a forensic laboratory: Benefits, pitfalls and suggestions to avoid misunderstandings' (2016) 56 (5) Sci Jus 364.


616 That such an approach needs to be applied appropriately, by an expert from the appropriate knowledge domain, is clearly exposed by the successful appeals in Clark, Cannings, and Anthony.

617 A. Biedermann and others 'Reframing the debate: A question of probability, not of likelihood ratio' (2016) 56 (5) Sci Jus 392.
Judicial responses to Bayesian reasoning and LR framework evidence

As noted earlier, while the use of Bayesian reasoning and the associated LR framework has been strongly promoted and used by the forensic community for some time, the reception of such reasoning and the related expressions of evidential strength has not been accepted uncritically or without problems in the courts. The judicial response to such evidence has been ambivalent and inconsistent, if not explicitly negative. While it is not practicable here to provide a review of all the relevant case law in relation to the use of Bayesian reasoning and the application of the LR framework in forensic evidence, it is important that I highlight some key issues in the presentation of certainty in forensic interpretation in court since they are relevant to the key focus of this chapter.

One important issue I noted above is judicial tolerance, or otherwise, of explicit probabilistic numerical evidence and argument. Here, generally, the courts have rejected the presentation of detailed statistical calculations, apparently concerned that a jury, overwhelmed by the numerical detail, may convict on the basis of numbers alone, a risk compounded by risks such as the prosecutor’s fallacy. The case law here is primarily focused on DNA profiling evidence and jury interpretation of random match probabilities.

---

618 Legal practitioners evidently struggle with the mathematical concepts of Bayesianism, and a variety of probabilistic fallacies, perhaps, most notoriously, the “prosecutor’s fallacy”, have contaminated the perceived utility of such an approach. For a review of such fallacies, and some of the solutions, see, for example, N. E. Fenton and M. Neil 'Avoiding legal fallacies in practice using Bayesian networks' (2011) 36 Aust J Legal Phil 114.


620 For a useful and actively updated list of court cases involving Bayesian reasoning and probability see that maintained by Prof Norman Fenton, Queen Mary University of London: https://sites.google.com/site/bayeslegal/legal-cases-relevant-to-bayes (Accessed 24 July 2018).


622 See Fenton, n. 618 above.
In the early 1990s, there were several examples of serious errors in the presentation of DNA evidence in court. A series of English criminal appeals - ((R v Doheny & Adams; R v Adams (No 1); R v Adams (No 2)) - exposed the challenges in adducing expert DNA evidence, and interpretative fallacies and other problems were identified. Doheny & Adams provided some guidance on how such statistical evidence should be presented. Here the courts considered that there was a clear mathematical basis for calculating a LR, and forensic scientists were permitted to present numerical probability evidence to the jury without further background statistical explanation or additional qualitative verbal comment.

However, outwith cases involving DNA analysis, the courts have been more circumspect with the use of numerical expressions in other types of forensic evidence. As Ward has noted, there will be a spectrum of appropriate forensic analyses depending on the particular features of an individual case and so the justifiable basis for calculating a numerical LR will vary. Thus, when dealing with other forms of forensic evidence, the courts have generally rejected the expert evidence presented in an explicitly statistical or

---

627 Such random match probability evidence as used in DNA forensics is a special type of likelihood ratio; see G S Morrison, The Likelihood-Ratio Framework and Forensic Evidence in Court: A Response to R v T, vol 16 (2012) 27. How DNA profile evidence is presented will depend on the particular requirements of an individual case; developments in DNA profiling significantly influence the way such information may be expressed. In reality, there are also cases such as those involving Low Template DNA, where there is a questionable basis for deriving a LR, and no database on which to derive a sliding scale; see R v Dlugosz and Ors. [2013] EWCA Crim 2, [2013] 1 Cr App R 32, [2013] Crim LR 684. See also Crown Prosecution Service guidance in relation to High Sensitivity DNA Analysis: http://www.cps.gov.uk/legal/s_to_u/scientific_evidence/high_sensitivity_dna_analysis/ (accessed 24 July 2017).
628 However, in Doheny the court recognised that such numerical evidence could be misused and misinterpreted (in particular, warning of the risk of ‘Prosecutor’s Fallacy’), and emphasised that DNA profiling evidence alone was not necessarily probative, so that such DNA matching evidence should be interpreted in the context of all the evidence adduced.
numerical format. Rather they have demanded that the expert expresses their confidence in the certainty of their interpretation of the evidence using a hierarchical or verbal scale as discussed above, (albeit the point on the scale may relate to the value of the expert’s subjective LR). 630

With the exception of DNA evidence, the courts have been generally more resistant to the use of numerical probabilities in expressing forensic opinion evidence. In R v T, 631 which involved footwear-identification evidence, the court seemed to reject the general application of the LR framework to the presentation of forensic evidence. Here the expert had used an undeclared confirmatory statistical analysis and derived LRs to reassure himself of the validity of his subjective judgment and qualitative verbal conclusion of the likelihood that footwear marks had come from a suspect’s shoes. However, the CACD, in quashing T’s murder conviction, held that the apparent use of confirmatory statistical analysis and derived LRs without a sufficient database to support that particular forensic interpretation rendered his subjective evidence inadmissible.

The firestorm of adverse academic and forensic practitioner comment in response to the judgment in R v T reflected widespread concerns about the court’s interpretation of that expert’s approach. 632 However, the issue in R v T relevant to my discussion here seems to have been the expert’s apparent misuse of statistical analysis and derived quantitative LR in the absence of a background comparative dataset, rather than expressing an opinion.

---

630 C G G Aitken and others, 'A verbal scale for the interpretation of evidence', vol 38 (Elsevier 1998) 279. See also, generally, Aitken, Roberts and Jackson, n. 608 above.


632 In R v T some consider that the court fundamentally misinterpreted the Bayesian approach - consistently committing a basic logical error, the ‘transposition of the conditional’. The judgment generated a storm of critical comments from the forensic and statistical communities; see for example, Charles E. H. Berger and others 'Evidence evaluation: A response to the court of appeal judgment in R v T' (2011) 74 (3) Mod Law Rev 444; William C. Thompson 'Discussion Paper: Hard Cases Make Bad Law - Reactions to R v T' (2012) 11 Law, Prob & Ris 347. For a bibliography of responses to R v T and ensuing discussion, see: http://forensic-evaluation.net/rvt/ (accessed 15 May 2018).
exclusively based on the use of a qualitative LR founded on study and experience, an approach well-established in the US.

Finally, in contrast to R v T, there are cases such as R v Atkins, involving photographic comparison (“facial mapping”) evidence, where there was no basis to calculate an LR and the expert simply adopted a particular verbal scale to convey his subjective judgment. Here the court appeared happy to accept the expert’s opinion, albeit entirely founded on personal study and experience, apparently reassured by his transparency about areas of subjectivity and uncertainty. The Atkins judgment contrasts starkly with the response to expert facial mapping evidence in some other jurisdictions, where there have been particular concerns about the pseudo-scientific gloss that informal hierarchical scales may provide to otherwise highly subjective interpretations.

The court’s finding in Atkins also produced much adverse comment, and perhaps reflected a broader issue – that the English courts generally have had a low threshold to admit expert evidence of uncertain reliability - typified by voice comparison, facial mapping, ear print and lip reading expert opinion evidence - where the scientific underpinnings of the “specialised knowledge” is at best uncertain. I return to this issue of the English courts’ expert evidence admissibility thresholds in Chapter 7.

Setting aside the significant concern that current techniques employed by facial-mapping “experts” are insufficiently reliable to be admitted in court, and indeed may increase the risk

634 This was the Bromby scale - a specific hierarchical scale for developed for facial comparison work (M. Bromby ‘At Face Value? The Use of Facial Mapping and CCTV Image Analysis for Identification’ (2003) 153 New Law J: Expert Witness Supplement 302).
635 Notably the Australian case of R v Tang, which involved both facial and body mapping expert evidence, and where the court rejected the expert’s personally derived hierarchical scale to quantify her opinion: [2006] NSWCCA 167; (2006) 65 NSWLR 681; (2006) 161 A Crim R 377.
636 See, for example, Gary Edmond and others ‘Atkins v The Emperor: the cautious use of unreliable expert opinion’ (2010) 14 IJEP 146.
of wrongful convictions, 
what the facial-mapping literature exposes more broadly is that problem may arise when the application of expertise does not provide a numerical output that may be matched with a formal hierarchy, yet experts’ subjective judgments are expressed in hierarchical terms.

Yet as the brief overview above suggests, the courts have developed an explicit expectation that when forensic experts provide opinion evidence, they quantify the certainty of their conclusions in a way that is both consistent and reproducible within their expert domain so that legal fact-finders may make a valid interpretation of the probative value of their evidence.

It is the difficulties paediatricians face in expressing their degree of certainty in their conclusions that I consider in the next section.

---

Section 2 - “Doctor – when you wrote that, what did you mean?”

In this section I consider the challenges that paediatricians face in appropriately expressing expert forensic opinions, and in particular, in the context of the probabilistic approach that I discussed earlier, how they might quantify and express their degree of confidence in the conclusion they have reached. First let me make some general points about paediatric expert conclusions.

Diagnosis: Murder

I cannot provide any natural explanation for this baby’s death. Rather, based on all the reasons I have provided in this report, I consider that it is beyond reasonable doubt that your client murdered his daughter.

Anticipating a plot line from the American medical crime drama television series ‘Diagnosis: Murder’ of the early 1990s, the above quotation is provided by way of a personal mea culpa. It is a direct quotation from one of my earliest expert reports provided to a defence lawyer more than 30 years ago. His client was charged with the murder of his two infants, some years apart.

Not surprisingly, I was not called to give evidence at the trial, in which the father was convicted and given a life sentence. While I still cringe as I reproduce it, happily I can report that my naivety and understanding of how a paediatric expert should, or should not, express the conclusions of their opinion in reports to instructing solicitors and the courts rapidly improved.

For the expert paediatrician providing a forensic opinion, it is still generally considered both professionally, and legally, inappropriate to explicitly “diagnose” a case as an instance of child abuse (or murder), although the former phrase is commonly used by clinicians in day-to-day practice. In a forensic context, this seems to focus on the avoidance of “diagnosing

---

a crime”, and by implication the paediatrician expressing an opinion on intentionality. Such
would usurp the function of the court.

Having said that, the diagnostic labels that are often used – for example, “Shaken Baby
Syndrome”, now more recently rebranded as “abusive head trauma” (AHT) or “non-
accidental head injury” (NAHI) (see later, Chapter 7) - appear to indicate a criminal
causation. Similarly, the use of terms such “accidental”, “unintentional”, “abusive” or
“inflicted” all seem to imply that the expert has considered the question of intent, again
apparently encroaching into the decision-making role of the court.

Yet except in very
particular circumstances, a clinician has no sound basis to look at any injury and provide
an opinion on the intent of the person who may have caused it.

Such questions of intent - the mens rea - clearly sit most appropriately within the legal as
opposed to clinical domain.

However, as I noted previously, the prosecuting authorities
often use expert evidence to guide the charges or indictments they may consider applying in
a particular case, so the use of such terms may have a significant influence in such decisions.

What is clear is that the words and phrases used by paediatric experts in expressing their
conclusions matter.

640 Maher, n. 120 above. Clearly, in some circumstances, the courts seem to give some leeway, such as
a paediatric pathologist providing an opinion in a case of sudden unexpected infant death.
641 Skellern, n. 222 above.
642 Arguably it may be appropriate in relation to alleged Factitious or Induced Illness: Gwen Adshead
‘Evidence-Based Medicine and Medicine-Based Evidence: The Expert Witness in Cases of Factitious
643 David, n. 5 above. David highlights the distinctions between terms such as ‘non-accidental’ and
‘deliberate’, emphasising the risk that a layperson (jury member) may conclude an unjustifiable
element of premeditation depending on the label applied within the paediatrician’s opinion.
644 Although some paediatricians have previously proposed a new classification of child abuse based
on motive and premeditation, rather the current system where the severity of injury or other outcome
of abuse determines the response from the various statutory agencies. See D. Southall, M. Samuels
and M. Golden 'Classification of child abuse by motive and degree rather than type of injury' (2003)
88 (2) Arch Dis Child 101.
Semantics and expressions of certainty in paediatric opinion evidence

Physicians make a diagnosis with varying degrees of certainty. When the consequences of an incorrect diagnosis are serious, diagnostic certainty should be maximized; however, there are situations (e.g., potential child abuse) in which the risks are great, yet no gold standard test exists to maximize certainty. To avoid the serious consequences of incorrectly diagnosing or excluding child physical abuse, clinicians must understand and communicate accurately the degree of diagnostic certainty to patients, families, prosecutors, jurors, law enforcement, and social services, who in turn must be able to recognize when an acceptable level of diagnostic certainty exists and when it does not.

I defend the use of the long quotation above because it sums up very precisely the issue I want to address here – how should a paediatric expert effectively express their certainty of the basis for their forensic opinion? In doing so let me briefly return to the standards an expert must meet in providing a forensic report for the courts.

As I previously noted, in complying with an expert’s duties, it is expected that the conclusions and opinion that an expert provides in a written report to the courts are true and complete to the best of their knowledge and belief. Thus, in England and Wales, an expert's report must include a relevant “statement of truth” to that effect.

Similarly, when a paediatrician (or any other expert) appears in court to give evidence in civil or criminal procedures, it is routine that once in the witness box they are asked to swear or affirm that they will tell “the truth, the whole truth and nothing but the truth”. The perhaps

645 Lindberg et al., n. 54 above, at e945.
646 In some circumstances, it is obviously quite appropriate to provide a provisional opinion if a final conclusion is dependent on some as yet unresolved question, perhaps the results of further medical investigations.
647 For the form of the statement of truth verifying an expert's report (which differs from that set out below), see paragraph 9.1(j) of Practice Direction 25B (The Duties Of An Expert, The Expert’s Report and Arrangements For An Expert To Attend Court); available at: http://www.justice.gov.uk/courts/procedure-rules/family/practice_directions/practice-direction-25b-the-duties-of-an-expert,-the-experts-report-and-arrangements-for-an-expert-to-attend-court (accessed 15 May 2018). There are four versions of such statements of truth relevant to paediatricians providing expert reports to the English courts, depending on the relevant judicial process; these include those for Civil, Family and Criminal Proceedings, and more recently a declaration which should be appended to the end of a Joint Statement following the Discussions of Experts. The various statements are available from the Academy of Experts website: https://www.academyofexperts.org/guidance/expert-witnesses/experts-declarations (accessed 15 May 2018). In Scots law, a “soul and conscience” declaration is the somewhat more generic phrase used to affirm that any written statement a skilled witness provides is true.
routine incantation of these ten words in the context of the administrative court process disregards some underlying fundamental metaphysical concepts about truth and belief, and expert evidence. As I alluded to earlier, questions arise here about the difference between scientific truth(s), and factual truth as determined in the legal system by a trial or other fact-finding process, so this issue of truth and belief is a slippery one in the context of paediatric expert opinion of alleged inflicted injury or the cause of a suspicious death.

Setting aside these metaphysical distinctions, what the expert actually means by the phrases he or she uses in written reports or oral evidence is clearly crucial to a proper understanding of their opinion and its relevance and implications in an individual case.

Fundamentally, the opinions paediatricians provide must surely reflect both the reliability of the underlying science and other knowledge relied on in reaching their conclusion, together with their degree of confidence or certainty in the interpretation of that science, and its limitations, in the context of the individual case. Such expressions of confidence reflect their strength of belief about the accuracy of their individual forensic judgments, and, at least in theory, sit on a continuum ranging from total certainty to complete doubt.

As I discussed earlier, how lay fact finders interpret those same expressions of confidence is equally important, and directly impact on the weight they give to that evidence, for we already know that in the setting of expert forensic testimony, even amongst the professional actors involved, there is much potential for ambiguity and misinterpretation. Although we


649 While medical expert witnesses are urged to ensure that “what you say in Court is the same as the opinions you have written in your report (Richard G. Notley 'Expert witnesses criticised' (2011) 17 (1) Clin Risk 2), a number of lawyers have suggested to me that when giving oral evidence, they find experts may be more diffident in the strength of their assertions.


651 G Jackson, C G G Aitken and P Roberts, 4: Case assessment and interpretation of expert evidence: guidance for judges, lawyers, forensic scientists and expert witnesses', Communicating and...
are not privy to *post hoc* analyses of the thought processes of lay jurors, it seems safe to assume that the potential for misinterpretation of expert evidence is likely to be at least as problematic for them.

At present, there is no empirical data available exploring the taxonomy of expressions that paediatric experts in the UK use to describe their confidence or levels of certainty about the causation of an injury. In the absence of empirical data, my own experience over many years reading innumerable expert reports, principally from the UK, is that the way paediatricians, and other (medical) experts, choose to qualify their opinions to indicate their level of confidence in their conclusions is extremely variable, highly subjective, and certainly open to (mis)interpretation.

This variability contrasts with a longstanding approach evident in the US, where until recently mandatory expressions of certainty in expert conclusions have been the norm. The contrast between the two jurisdictions is striking and merits exploration. So let me highlight two particular expressions of confidence used by paediatric experts in the US and the UK that expose some of the interpretative issues involved, and reveals the inherent challenges for the criminal justice system (in each jurisdiction) in understanding how certain paediatric experts are in the opinions they provide. As I will show, this exploration unmasks another specific frailty in relation to paediatric expert evidence.

First, I briefly consider the US situation, before coming closer to home.

---

Expressing forensic paediatric opinions in the US - “reasonable medical certainty”

In the US courts, the term “a reasonable degree of medical certainty” is long established within both civil and criminal courts as the expected standard of confidence for diagnostic and other clinical opinions provided by medical expert witnesses. This same phrase is routinely modified and applied by other experts depending on the forensic domain – a “reasonable degree of [discipline] certainty”.

According to Lewin, the routine practice of asking medical expert witnesses to state that they held their opinions to a “reasonable medical certainty” evolved in the Chicago trial bar in the early part of the last century, and had become routine in the Illinois jurisdiction by the 1930s. While the initial application was focused on the avoidance of speculative predictions of the future impact of illness or injury in civil tort cases, this practice soon spread to other areas of expert evidence.

By the 1960s, the phrase had become firmly established in many other US jurisdictions and had begun to emerge as a standard for admissibility. Although never a mandatory requirement in state or federal courts, the phrase continues to be remains embedded in civil and criminal practice in many US states, and its inclusion in forensic clinical opinions

---

652 “Within the medical discipline, the traditional standard for ‘factfinding’ is a ‘reasonable medical certainty’”: Addington v. Texas, 441 U.S. 418 (1979), at [10].
654 United States v Two Elk. 536 F3d 890 (8th Cir 2008); United States v Mornan 413 F.3d 372, 381 (3d Cir. 2005).
has hitherto been regarded as an obligatory standard for US paediatricians providing expert opinions for malpractice claims and alleged child abuse evaluations. Yet despite its ubiquitous use and the apparent reassurance the phrase generates, the expression has rarely been defined, and its meaning is uncertain and open to interpretation. While there has been relatively sparse academic comment on the issue, emerging criticism of its continued use argues that it is a phrase without scientific meaning, a ritualistic incantation that experts only use in a forensic setting and not in conventional medical or scientific communications:

Scientists might refer to personal degrees of confidence in a finding or to the degree of controversy surrounding it, but there is no accepted or working definition of a “reasonable degree of certainty” in scientific discourse.

At first glance, perhaps because of the inclusion of the word “certainty”, the phrase appears to indicate a high level of confidence in the clinical and scientific foundation of the expert’s opinion and, in a child protection setting, could suggest that injuries are very likely abusive in origin.

---

658 Holbrook v Lykes Bros Steamship Co Inc. 80 F3d 777, 785 (3d Cir 1996), at [42].
659 The American Board of Forensic Odontology previously defined reasonable scientific certainty as “no reasonable probability of error”; see McQuiston-Surrett and Saks, n. 613 above, page 1162.
Unfortunately, irrespective of its widespread use in the US courts, it is clear that there is a significant potential for misunderstanding what individual experts may signify by it, and thus how lawyers, juries and the courts may interpret the phrase. Thus uniformity of terminology by experts does not necessarily avoid or minimise misinterpretation or misunderstanding by their audiences in the legal setting.

Recent empirical evidence from the US shows that expert medical witnesses in child abuse cases apply very variable thresholds of “reasonable medical certainty” when providing expert opinions whether or not child abuse had occurred. While the vast majority of the experts considered that child abuse had to be “the most likely” diagnosis to signify “reasonable medical certainty”, there was significant individual variation in the level of probability the experts considered met that standard. Thus, while around half defined “reasonable medical certainty” as a probability of abuse of 90% or more, almost a third of the experts considered it to be set as a probability of 50% or more, while a minority set the level even lower.

Such variability in what experts mean has real world consequences, for in cases that turn almost exclusively on expert medical testimony, it is important that judges and jurors understand the probability threshold that an individual expert is applying in reaching their conclusions. Thus courts may be faced with two experts each providing conclusions with “a reasonable degree of medical certainty” as to the cause of an infant’s injuries, but in reality with very different thresholds of certainty. It seems that unless US paediatric experts start to state explicitly the probability thresholds they are using, there is the potential for significant injustice here.

---

664 See, for example, Nelson Abbott and Landon Magnusson 'An enigmatic degree of medical certainty' (2008) (July/Aug) Utah B J 20. The authors point to the contradictory construction of the phrase itself, arguing that “certainty” reflects a binary construct – an expert is either certain or not – thus challenging the modifier “reasonable”. They provide real life examples from the Utah Bar of the confusion and misinterpretation of the phrase, and its impact on the legal process.

Another related issue with the use of the term *certainty* here is the potential for confusion with the *burden of proof* standard, for lay jurors may interpret an expert’s use of the word *certainty* to imply his findings relate to the *beyond reasonable doubt* standard of proof that must be met in criminal proceedings.  

Perhaps the use of the word “certain” is the problem here, given its absolutist connotation. Arguably, in most forensic child abuse work, there is usually little doubt about the clinical facts of a case – the actual injuries or other features found. Rather the issue is uncertainty about the validity of the inferences that can be drawn about the *cause* of the injuries, which is exactly why expert opinion evidence is required. The challenge is to find a (relatively) unambiguous expression of confidence in the inferences that the expert has made as to causation.

For this reason, there have been attempts to develop and agree various ordinal and verbal scales that might promote consistency among US paediatricians and related experts in expressing the likelihood that a child has been abused. These have met with mixed success.

What the US experience does reveal is that, even when there is an apparent demand from the courts that paediatricians must hold an opinion with a particular level of certainty before it is admitted before the court, unless such thresholds are clearly defined and are commonly understood by both practitioners and the relevant legal actors, such semantic formulations are more likely than not to impair rather than support the justice system. Thus in responding to the need for more explicit clarity in how paediatric experts should qualify their opinions, the US experience suggests that much more than a semantic gloss is needed.

---

666 John E. B. Myers 'What is the meaning of “reasonable medical certainty”? Your guess is as good as mine' (2015) 50 Child Abuse Negl 228.
667 Lindberg et al, n. 54 above.
Before discussing that further, let me return to consider the semantic formulations used in the UK child abuse field.

**The semantics of forensic paediatric opinion in the UK**

In the UK, unlike the US, in both a clinical or forensic context, paediatricians or other medical experts do not use any established or recommended expression to describe their level of confidence in either a clinical diagnosis or a forensic inference. As I noted earlier, my experience indicates that, in stark contrast to the narrow US approach, UK paediatricians employ a wide variety of expressions to express the certitude of their forensic conclusions.

In most cases, such expressions represent their overall impression of a case. It is rare in my experience, for example, for a paediatric expert to provide a range of certainties in a particular case predicated on whether or not the opinion from another specialist, on which the paediatrician has relied, is accepted by the court. And while this lexicon is not formalised among the UK paediatric forensic community, reading below the surface, it is usually possible to glean some measure of understanding of an expert’s apparent confidence in their interpretation across a spectrum of certainty.

Thus expert reports that contain qualitative terms such as “diagnostic”, “strongly indicative”, or “highly suggestive” generally appear to suggest a high level of expert confidence that injuries found are abusive in origin.

---

668 It is clear that the expert should acknowledge that there is controversy and a range of opinion about what inferences are warranted in particular circumstances, and why there is an expectation that the expert does not just provide an opinion in limbo, but justifies that opinion by explicit explanation of the rationale behind it.
In contrast, terms such as “consistent with” or “suggestive of” indicate less confidence, while “a possibility” or “concerning” seem to suggest significant diagnostic doubt, and certainly little approaching a confident or convincing diagnosis of inflicted injury.

Let me return to the use of the word ‘consistent’ - a particularly opaque term that I have encountered relatively frequently.

“Consistent with”: lost in translation?

In expert opinion reports in relation to child abuse in the UK, it has been a relatively common practice for paediatric experts to state that a child’s history, symptoms and presentation are (or are not) consistent with abuse. Yet, while this phrase features prominently in the current guidance on forensic paediatric report writing in the UK, in the context of alleged child abuse I suggest that the expression is a vague and relatively unhelpful one.

As part of a formalised hierarchical semantic framework of the United Nations Istanbul Protocol focused on the investigation and attribution of injuries in alleged torture victims, the term “consistent with” is defined as: “the lesion could have been caused by the trauma described, but it is non-specific and there are many other possible causes”.

Thus it is a neutral term suggesting that there may be alternative innocent explanations for injuries that cannot be excluded. Despite its relative prominence in UK child protection

---

669 Abstraction of such phrases from case records has been used to categorise and rank diagnostic certainty in child abuse research (Lindberg et al., n. 54 above; Lynn K. Sheets and others 'Sentinel injuries in infants evaluated for child physical abuse' (2013) 131 (4) Pediatrics 701).
673 See, for example, the judgment of Thorpe LJ in Re B (Non-Accidental injury) [2002] 2 FLR 1133, at [25].
documentation, it is recognised to be particularly problematic when used in forensic
reports, and a term that jurors may interpret very variably.

In the context of fact finding in possible non-accidental injury, such uncertainties are in no
one’s interests.

The Goudge Inquiry considered in some detail the appropriate language experts should use
in forensic pathology reports, and emphasised more broadly the need for relevant
professional groups to agree a common and consistent lexicon that was appropriate to use in
individual specialties, and whose meanings were clear. Yet at present there is no explicit
guidance available from relevant professional bodies such as the RCPCH on how such
forensic evaluative opinions should be expressed except in very general terms, a deficiency
that needs to be formally addressed.

However developing a taxonomy of expressions to signify the degree of confidence or
certainty that paediatric experts might use to reflect the correctness of their forensic
conclusion faces particular challenges.

The challenge of developing an appropriate taxonomy of forensic certainty

Perhaps the most obvious difficulty stems from the stark difference between the clinical
forensic assessment process and the often much more narrowly focused laboratory-based
evaluations performed by forensic science laboratories where numerical outputs, LRs, and
integrated verbal expressions may apply.

Indeed, the term was also specifically criticised in the Goudge Inquiry: Graham D. Glancy and Cheryl
Acad Psychiatry Law 81.
675 R. Ross, K. Kramer and K. A. Martire. 'Consistent with: what doctors say and jurors hear' (2017)
676 Stephen T. Goudge 'Victorian Institute of Forensic Medicine Oration: The Inquiry into Pediatric
677 Glancy and Regehr, n. 674 above.
As I discussed earlier, paediatric forensic evaluations are largely set in clinical experience. And since, in general any injury may be sustained by accident or assault, the clinical findings (with some exceptions) are not in themselves direct markers of abuse, it is the broader interpretation of the findings in the context of other factors that represent the forensic expertise. While the courts have generally accepted experience-based opinion evidence, as I have previously noted in preceding chapters, it is potentially highly susceptible to cognitive biases, and, depending on the specific clinical issue, the reliability of the underlying evidence base is variable.

As I also previously highlighted, such forensic opinions are further complicated by the fact that the expert paediatrician often depends heavily on other specialist colleagues’ expertise and conclusions on which they must partly found their own evaluative opinion, and the expert paediatrician often provides an overview of the whole case, drawing together disparate expert opinions and agreed facts.

How such dependencies and the sub-certainties of these experts might be encapsulated and expressed by the forensic paediatrician’s subjective but justifiable expression of confidence in their conclusion is clearly complex.

As I noted earlier, derivation of formal LRs and associated verbal scales are not generally appropriate when dealing with clinical interpretation. In situations where it is not possible to empirically derive numerical estimates to support the “truth” of a forensic conclusion, formulating some form of subjective warrant for such beliefs would seem the only alternative. But as McQuiston-Surret and Saks have noted: “[f]orensic expert witnesses...”

---

679 As seen, for example, in Re L (a child) [2013] EWHC 4646 (Fam).
680 See, for example: Martire et al., n. 81 above; Kristy A. Martire and others 'The Expression and Interpretation of Uncertain Forensic Science Evidence: Verbal Equivalence, Evidence Strength, and the Weak Evidence Effect' (2013) 37 (3) Law Hum Behav 197.
cannot simply adopt a term, define for themselves what they wish it to mean, and expect judges and juries to understand what they mean by it.”

Thus this is not a role for individuals. Rather, relevant UK paediatric professional bodies will have to derive appropriate (specialty) verbal scales based on professional consensus with which to make explicit for fact-finders the certainty or otherwise of their opinions. In the absence of such an agreed verbal scale, individual idiosyncratic expressions as noted above will remain open to misinterpretation.

The two illustrative terms that I have discussed above – “reasonable medical certainty” and “consistent with”- reflect back on a broader issue that I referred to previously, the fact that when paediatric and other medical experts provide expert opinions they are moving from science into the relatively foreign realm of law, and there are different perspectives and expectations to meet.

As Narang has indicated, in the dynamic and unstructured diagnostic processes applied in even the most challenging aspects of clinical care, when lives are truly at stake, clinicians rarely stop and ask themselves how “reasonably certain” they are as they apply a series of diagnostic judgments, plan treatments, and observe how the patient responds. Put simply, certainty in clinical practice is largely illusory.

His reminder should prompt paediatric expert witnesses to recognise that, despite all that may be at stake in alleged child abuse cases, as medical experts we need to resist when artificial expectations of certainty are being imposed, and that it is important to acknowledge when we simply cannot be sure. As I have learned in my many visits to the witness box, expert evidence should rarely involve “yes” or “no” answers – appropriate explanation of expert evidence almost always demands a conversation.

---

681 McQuiston-Surrett and Saks, n. 613 above, page 1163.
Given the evident challenges in findings appropriate and consistent expressions of confidence in the conclusions paediatric experts provide, in the final short section to this chapter I return to the question of whether or not a Bayesian probabilistic approach might be usefully applied to some areas of paediatric expert evidence.
Section 3 - A Bayesian approach to forensic child abuse assessment?

Behind every opinion rendered by a forensic scientist there is a statistical basis. We may not know what that basis is, and we may have no feasible means of developing an understanding of that basis, but it is futile to deny that one exists. 683

In this final section I consider whether the developing probabilistic approaches to the expression of certainty in general forensic science might be applied to forensic paediatric expert opinions.

As I noted earlier, one of the most obvious differences in perspectives between science and the law is that they examine particular situations from different standpoints. As Dawid et al have noted: “science studies individuals in order to make statements about populations, while the law considers populations in order to make statements about individuals”. 684 In highlighting that the law reasons from group data to individual cases, Dawid distinguishes the usual scientific endeavour - trying to determine the “effect of causes” (EoC), from the legal focus on “cause of effects” (CoE). 685 In the context of alleged child abuse, the courts (albeit advised by expert paediatricians) are concerned with resolving uncertainties in assigning the cause of a child’s injury, in particular seeking to resolve the specific CoE question: “was this child’s injury inflicted?”


685 I acknowledge here that some commentators have challenged this claimed CoE and EoC distinction; the details of this debate, summarized by Cuellar (M Cuellar, 'Shaken Baby Syndrome on Trial: A Statistical Analysis of Arguments Made by the Defense and Prosecution' in S. E. Fienberg, J. Kadane and A. Haaviland (eds), Carnegie Mellon University (Doctoral Research Presentation edn 2016) 1, at 3.2), do not concern us here.
While this has been characterised as a probabilistic inferential process characterised as “backcasting”,686 as will become evident below, I suggest that forensic expert opinions in alleged child abuse generally sits very uncomfortably within such probabilistic paradigms.

Perhaps the most obvious challenge to the adoption of such a process is the highly complex setting in which such judgments must be made. Forensic child protection work seems a very different and much more cluttered interpretive environment than that of the forensic laboratory bench, where the focus may be on a relatively specific technical comparison, and, setting aside the human factors I discussed earlier, probabilistic reasoning seems to fit quite well.

In contrast, the idiosyncratic constellation of individual features that have to be interpreted in cases of alleged abuse mean it is not generally possible to derive formal numerical probabilities of the likelihood of one proposition compared to another. Instead the clinical expert must express their degree of support for a particular hypothesis based on qualitative evaluation founded on experience and knowledge, a process that relies on a subjective probability assessment of all the available facts and the expert’s personal degree of belief in one proposition over others.687

Thus, except in narrow circumstances, the discipline’s epistemic mode of reasoning in forensic clinical assessment does not seem to fit well with a statistical approach.688

However, although the literature on the subject is extremely limited, some formal attempts have been made to examine the utility of a Bayesian approach in forensic child abuse evaluation.

686 Dawid et al., (2016), n 570 above.
688 Although the use of formally calculated LRs to support clinical diagnosis may be very helpful, often focussed around the interpretation of ancillary investigations. See, for example, David A. Grimes and Kenneth F. Schulz ‘Refining clinical diagnosis with likelihood ratios’ (2005) 365 (9469) The Lancet 1500.
Published studies applying probability analysis to child abuse interpretation

Almost 30 years ago, preliminary studies were reported of such an approach to the diagnosis of physical abuse, focused on children with various bruise patterns. The same group later developed a scoring system that purported to function as a forensic diagnostic tool, based on a calculation of the prior probability of abusive injury dependent on the location of examination, together with differential weighting for bruises on different body areas and the totality of bruising.

Early optimism that this approach had potential for widespread application has not been fulfilled. Perhaps the most obvious area of concern was how the authors determined the value to be assigned for the prior probability of abuse. Using outcome data from previous evaluations, they allocated a value based on the location where a child was first assessed – thus 0.4 for bruised children seen at the local child protection suite, but a much lower figure (0.01) if a child with the same bruise pattern was assessed in an emergency department.

A final diagnosis of abuse was determined by the outcome of a case conference, which, as I previously noted, represents a decision strongly focused on an assessment of future risk and takes account of many other factors apart from the medical forensic evaluation. As the authors themselves acknowledged, such a scoring system could not replace the complex qualitative analysis intrinsic to the diagnosis of abuse.

More recently, adopting the backcasting perspective I noted earlier, an investigation into the potential utility of a Bayesian statistical approach applied to a theoretical complex child

---

691 This figure was derived from the proportion of children (~40%) diagnosed as abused after multidisciplinary evaluation in their dedicated child protection examination suite.
A protection case has been reported. This investigation, a joint enterprise by the RCPCH and the Royal Statistical Society, was prompted by the outcry over the misapplication of statistics by Professor Meadow in Clark and Cannings, and linked to concerns expressed by Professor David Southall that infant nosebleeds as reported in one of the Clark infants before his death, was a significant marker of likely non-accidental injury.

Thus, the project focused around the theoretical scenario of a young infant who presents to hospital with an acute life threatening event (ALTE), where it becomes apparent that there is a background of nosebleeds, and a history that an older sibling had previously died, the sibling death classified as SIDS. The study suggested that the estimated probability of abuse using this formal Bayesian approach was significantly lower than that suggested by expert advice or “informal reading of the literature”.

However, also evident to a non-expert in statistics such as myself was that such an approach involved highly complex statistical analyses well beyond the understanding of the “ordinary” child abuse expert (or lawyer for that matter). Thus how such an approach might be applied in day-to-day forensic expert assessment is questionable.

Further, as the authors themselves acknowledged, forensic evaluations involving a large number of symptoms and signs would require a different approach, and the lack of a gold

---

693 Owen Dyer 'Doctor charged with misconduct over murder claim' (2004a) 328 BMJ 1393.
694 An ALTE is usually defined as an episode of major change in an infant’s condition that is frightening to the observer and that is characterized by some combination of apnoea (cessation of breathing), colour change, marked change in muscle tone (usually limpness), choking, or gagging.
695 In young infants, this has been considered highly suspicious for inflicted airway obstruction; see D. Becroft, J. Thompson and E. Mitchell 'Nasal and intrapulmonary haemorrhage in sudden infant death syndrome' (2001) 85 (2) Arch Dis Child 116; L. J. Walton and F. C. Davies 'Nasal bleeding and non-accidental injury in an infant' (2010) 95 (1) Arch Dis Child 53.
696 This had many similarities with some of the case details in Clark.
standard for child abuse diagnosis significantly hampers this, or other, options. As the same authors stated in a further paper extending the previous analysis:

[Integrating practitioners’ assessments with research-based probabilities of abuse is fraught. Research on suspected signs may be sparse or absent, while signs themselves may combine in a multitude of ways, which might not have been researched jointly, even if individual signs have been studied.

What these various studies seem to reveal very clearly is that, in contrast to many other areas of applied forensic expertise, the complex and multifaceted aspects inherent in child abuse forensic diagnosis, compounded by the patchy and questionable evidence base I discussed earlier, do not fit in at all easily with a Bayesian statistical approach to expert reasoning under uncertainty, or the derivation of quantifiable expressions of diagnostic certainty in related opinions. Thus at present such analysis remains largely an academic exercise.

In considering the application of probabilistic reasoning to cases of alleged child abuse, there is one final application of a probabilistic approach to the diagnosis of possible abuse that it is important that I acknowledge – focused on the assessment of infants with possible abusive head injuries.

**The Predicting Abusive Head Trauma (PredAHT) clinical prediction tool**

In attempting to support paediatricians’ decision making when dealing with infants with head injuries, Maguire and co-workers recently published a decision support tool that claims to provide probabilistic decision support on whether such injuries were caused by abuse. The model evolved from one that predicted inflicted infant head injury based on individual

---

697 Best, n. 692 above, at 76.
clinical features, to predictions based on various combined patient characteristics, the tool then validated on a new dataset.

Put simply, if a paediatrician treating an infant with a head injury is concerned that it may have been caused by abuse, the tool purports to provide probabilities (and range of certainties) that the injury is inflicted or not, depending on the presence of a variety of other associated clinical indicators, including, for example, the co-existent presence of fractures and/ or bruising.

Superficially, this appears to be a very useful aid to forensic diagnostic decision-making, and when first published I personally used these publications to justify my opinions in reports I provided to the Crown in relation to alleged shaken baby cases. Here, by referring to research that provided estimates of the association of single and pooled clinical variables with abusive head injury, I was trying to comply with explicit guidance from the Kennedy report that even experts’ confident conclusions of abuse should be supported by scientific evidence.

However, having initially included details of these published papers as a strong justification for my conclusions that injuries were abusive, as time progressed I started to reconsider its specific value in this and similar cases. Eventually, between the submission of my most recent report to the Crown Office and subsequent precognition with the prosecuting counsel, I decided to withdraw that section of my report. I notified the Crown, and ensured that the defence team were also aware of this retraction.


701 Such features include a history of apnoea, rib and/ or long bone fractures, retinal haemorrhages, and head/ neck bruising.

702 Baroness Kennedy, n. 131 above.

703 Once my report had been submitted to the Crown, as is mandated, a copy of my report had been supplied to the accused’s defence team.
So, what had changed? There were three reasons that prompted my withdrawal of this section of my report.

First, as I considered the general implications for that paper, and in particular, the security of the interpretation of the data that underlay it, I became conscious that there were some concerns about selection bias in the ascertainment of included cases that might exaggerate the purported discriminatory value of the various clinical features in differentiating abusive and non-abusive cases.\footnote{See comment on Maguire et al. 2011, n. 700 above: Lindberg, Daniel M., Beware Spectrum Bias in Estimating Abuse Probability; published in Pediatrics 12 October 2011; available at \url{http://pediatrics.aappublications.org/content/128/3/e550.comments} (accessed 25 April 2018).} I had also become concerned about the issue of circularity in the confirmation of abuse in individual cases, a general concern that I discussed previously in relation to the evidence base for child abuse (Chapter 5).

Second, and from a more personal perspective, I became concerned about my ability to discuss the interpretation of such figures in the witness box, particularly the statistical background to such an analysis.\footnote{Notably the publication had an extensive on-line explanatory appendix detailing the complex statistical procedures and analysis that had been followed in generating the various probability estimates that had featured prominently in my report.} This reflected a concern to explicitly remain within the boundaries of my expertise.

Third, I had a broader concern about how the authors expect the tool to be used. The authors originally asserted that the tool may “assist clinicians offering medical testimony in civil or criminal proceedings, in demonstrating why certain combinations are more or less predictive of an abusive etiology”.\footnote{Maguire et al. 2011, n. 700 above, at e558.} However, later they seemed to deny that the purpose of the clinical tool is to provide a foundation upon which to base an opinion that will be used in a criminal prosecution.\footnote{Laura E. Cowley and others 'Letter to Editor' (2018) Law, Prob & Risk mgy002.} Rather it seemed that they now wish to restrict the tool to a triage role for use by non-expert clinicians, and that once a certain (unspecified) threshold of
concern that an infant’s head injury may be abusive is reached, such cases should be referred on to forensic practitioners for further evaluation.

Despite the apparent early promise that the tool suggested, I agree with that suggestion. While the various clinical features the authors highlight are clearly “red flags” that should raise serious concern about the possibility of abuse, for the forensic expert required to stand in the crucible of the witness box and defend an opinion in an individual case, I suggest great caution must be used in applying such probabilistic information in anything other than a general supportive sense, and those that do must be prepared to explain in detail the statistical basis behind their conclusions. Here the spectre of the unfortunate Sally Clark and the travails of Professor Roy Meadow loom large.

These examples illustrate that a Bayesian paradigm and formal probabilistic approach do not fit well with the particular milieu of paediatric forensic interpretation and associated evaluative opinion evidence.

**Conclusion**

This thesis examines the possibility that there are specific vulnerabilities in paediatric expert opinion evidence that might explain the relative prominence of flawed convictions in cases of alleged murder or inflicted injury involving infants and young children.

In previous chapters I argued that a variety of human factors, the emotive nature of such cases, and the influence of various professional values might enhance the risk of bias in paediatric forensic evaluations of alleged abuse, and questioned the foundation of some of the evidence base on which such opinions should be based.

What I have shown in this chapter is that despite an expectation that an expert witness should provide some measure of certainty or confidence in the conclusions they offer to the courts, this is a particularly problematic issue for the paediatric expert offering opinion evidence in cases of alleged abuse. There is no agreed lexicon, or hierarchical taxonomy, with which to
qualify such opinions, and the use of probabilistic conclusions are largely theoretical. This difficulty is just one of many challenges for the forensic paediatric expert that I have identified in the preceding chapters.

What I have shown is that in many core aspects of paediatric forensic interpretation and associated expert opinion evidence there are significant potential vulnerabilities that, both individually and collectively, may contribute to the risk of unjustified findings in the criminal and civil courts. It is notable that almost none of issues that I have discussed have been previously explicitly acknowledged or empirically explored by the paediatric community.

Having examined paediatric expert evidence from the standpoint of the expert paediatric practitioner, and with these vulnerabilities in mind, in the final substantive chapter of this thesis I now turn to consider the challenge for the courts in ensuring that only reliable paediatric expert evidence is admitted before them.
Chapter 7 – “Shaken baby” cases – a diagnosis on trial

Introduction

So far in this thesis my focus has been on exposing problems in the provision of expert opinion evidence in cases of alleged murder or serious inflicted injury in children from the professional perspective of the expert paediatric practitioner. Thus in the preceding chapters, following the natural sequence from case referral and information exchange, to parental interview and clinical assessment, and then diagnostic formulation and opinion generation, I have explored potential frailties in the key components that together constitute an expert paediatric forensic assessment.

In these various chapters I have highlighted relevant aspects of human cognition, sources of bias, uncertainties in the evidence base, and other potential threats that may individually and collectively undermine the reliability of the opinion evidence that paediatric expert witnesses may provide in supporting the fact-finding process of relevant tribunals in cases of alleged child abuse. As Clark and Hainey show, these risks are more than theoretical, and the relevance and reliability of paediatric expert evidence are not guaranteed by the eminence of the expert, or their performance in court.

Such cases illustrate more broadly the need to manage the interface between expert clinical knowledge and its forensic application, and in particular to ensure that any proffered expert evidence and its use complies with the normative criteria of procedure and rules of evidence, in order to meet the overriding objective of criminal (and civil) adjudication – that such cases are dealt with justly.\textsuperscript{708}

While the obligation to ensure that only evidentially reliable opinion or material is put before the courts rests with all the actors involved in such cases, as Hainey illustrates, it is on the trial judge that such a responsibility ultimately falls.

\textsuperscript{708} In English Law this is now an explicit normative obligation. See Criminal Procedural Rules 2015, Rule 1.1: The overriding objective.
Recognising the importance of this alternative perspective on paediatric expert evidence, in this final substantive chapter I turn to discuss this key challenge for the courts – the need to assess and resolve disputed uncertainties about the scientific reliability of any proffered expert opinion evidence. Here I focus on the difficulties that the various potential problems with paediatric expert evidence I discussed earlier may pose for those charged with managing such evidence from the other side of the witness box.

However, rather than take a broad-brush theoretical approach, I intend to explore this issue within the focused reality of the criminal prosecution of alleged “shaken baby” cases. Here the criminal proceedings depend almost exclusively on expert diagnostic opinion evidence, a situation that, in the most serious cases, has been likened to allowing experts to provide a “medical diagnosis of murder”. As I will discuss, it is in such “hard cases” that many of the frailties in paediatric evidence I highlighted in previous chapters are particularly evident, and the challenges for judges and other relevant legal actors become particularly explicit.

Setting aside this general justification, there are specific reasons that support my decision to concentrate on this particular topic.

First, as I discuss presently, the professional interpretation of the scientific evidence underlying a specific triad of clinical findings—subdural haemorrhage (SDH), retinal haemorrhage (RH), and brain damage (encephalopathy)—hitherto generally accepted within the paediatric community as “diagnostic” of a shaking injury in infants, has recently been challenged and the established professional consensus on which such forensic opinions have previously been based is now arguably less secure.

In considering the reliability of such evidence, the courts must resolve conflicting expert perspectives, with significant on-going institutional support for the traditional diagnostic paradigm underlying SBS challenged by novel claims of alternative exculpatory

explanations, while a significant professional minority simply assert that since the pathological processes underlying “the triad” in isolation cannot be currently explained, a forensic opinion firmly suggesting shaking as the cause of the clinical findings must be considered unreliable.

Disentangling the various scientific inferences and associated “certainties” with which discordant expert opinions on these disputed issues may be expressed goes to the heart of the challenge for the courts in trying to ensure that only opinions founded on reliable science are put before jurors by appropriately skilled witnesses.

Second, this is not a theoretical problem but a current real-world challenge for our courts. The Scottish judiciary encounter such cases on a reasonably regular basis. Most months in Scotland, at least one child is diagnosed with a non-accidental head injury, and despite the emerging scientific debate about the validity of the triad alone as indicative of a shaking injury, many of these prosecutions continue to rely almost exclusively on the presence of the triad—for example, HMA v John Dobbie (2014), HMA v Steven Davidson (2015), HMA v David Sinclair (2016), and HMA v Paul Burgess (2016), (all unreported).

Thus, an exploration of the reliability and thus admissibility of expert evidence in such cases may arguably expose some of the difficulties judges may encounter in applying such a responsibility more generally. And as we will see, many of the key concerns about paediatric

710 No official figures are available for prosecutions in Scotland in relation to NAHI. Clearly for a variety of reasons, criminal prosecutions in such cases may not be pursued, although the vast majority will trigger formal child protection procedures. The strength of medical forensic opinion is usually the determining factor in criminal prosecution decisions; see Cathy Cobley and Tom Sanders, Non-Accidental Head Injury in Young Children: Medical, Legal and Social Responses (Jessica Kingsley London 2006) 192, at 50.
expert evidence that I highlighted previously – various biases and the emotive impact of such cases, uncertainties about the evidence base, and how the certainties with which such opinions are expressed are particularly exposed in SBS cases.

This chapter is divided into four sections.

In Section 1, in order to allow the reader a sufficient understanding of the clinical context for my later discussion, I provide an account of the clinical and forensic issues involved in SBS cases. On this task, I describe the evolution of SBS as a “diagnosis”, highlight the key contentious issues associated with the forensic interpretation of such cases, and identify the challenges to and (relative) erosion of some of the scientific underpinnings of the related expert opinion evidence in relation to SBS.

Having provided the clinical background to the forensic assessment of alleged SBS, in Section 2 I reflect back on my preceding chapters and show that many of the potential problems with paediatric expert evidence that I identified earlier coalesce around the forensic issues in alleged SBS. In particular I discuss these in the context of paediatric institutional positions and the guidance they provide.

Having set out the clinical and professional perspectives and controversies around SBS, in Section 3 I move to the other side of the witness box and discuss SBS from the perspective of the courts. First, in order to set what follows in its appropriate legal framework, I provide an overview of the general principles of the admissibility of expert evidence in the English and Scottish courts.

Then, acknowledging the relative absence of the application of reliability thresholds to skilled witness evidence in Scots Law, I describe the changing approaches to the admissibility of expert evidence in criminal trials in England & Wales, noting in particular

713 See Lady Scott’s submission to Lord Cullen, Chair, Scottish Law Commission, proposing reform of the admissibility and use of expert evidence in criminal trials as a topic for inclusion in the Commission’s Ninth Programme of Law Reform, 16 August 2014; available at: https://www.scotlawcom.gov.uk/files/7114/3161/1713/Lady_Scott.pdf (accessed 12 May 2018).
the impact the English Law Commission’s reform proposals on expert evidence in criminal proceedings have had in that jurisdiction.

I then return to some of the key forensic challenges of alleged SBS. First, in the face of the highly polarised SBS professional dispute, I illustrate how the English courts have responded to claimed uncertainties about the evidential reliability of the triad as a marker of inflicted head injury.

Building on the foundation of the two preceding sections, I argue here that what the SBS dispute exposes very starkly is the special nature of paediatric expert evidence, and illustrate how many of the potential frailties in paediatric expert evidence I highlighted in earlier chapters come into particular focus in the criminal prosecution of alleged SBS.

Second, I argue that even in the face of efforts by the English courts to actively manage expert evidence, the reality is that except in relatively extreme circumstances, the courts are very limited in their ability to assess the true scientific reliability of the evidence experts provide, and are dependent on indirect proxies of uncertain validity.

Acknowledging that, as non-scientists, it is difficult to envisage how lawyers, judges, and jurors are expected to appropriately resolve such diverging opinions, I nevertheless question the justification of the court’s traditional contrasting asymmetric admissibility thresholds between “established” and “novel” science.

I argue here that not only is there is an inherent danger in deferring to the institutional consensus view in such circumstances, since professional institutions are often slow to recognise that science has moved on, but that in the particular context of paediatric institutions, the professional values and perspectives that such institutions espouse may separately inhibit acknowledgment of the weakening of the science that may underpin their establishment views.

In drawing this chapter to a conclusion, in the light of the issues that I have discussed in previous sections, in **Section 4** I raise some specific practical issues for the paediatric expert that arise from the preceding discussion, before ending by challenging the justification for criminal prosecution of cases based on the presence of the triad alone that are currently applied in Scotland.

Rather, I suggest that if such a prosecution is pursued, a defence challenge to the admissibility of proffered expert opinion evidence that the triad (in isolation) is a marker of a shaking injury should succeed on the grounds that it is insufficiently reliable to be put before the jury. Otherwise, I argue that, given the emotive nature of such cases, the prejudicial effect of even modest expressions of confidence by an expert that this is the likely cause of the injuries may outweigh the true probative value of such an opinion, especially if an accused’s lifestyle and other adverse social factors are emphasised within the prosecution narrative.

Thus, in terms of the overall original contribution of this thesis, this chapter brings together the various concerns about paediatric expert evidence I raised in previous chapters and exposes them within the stark reality of our criminal courts.

With that chapter structure in mind, I now move to **Section 1** where I set the scene with an overview of the clinical issues and scientific disputes in relation to the forensic “diagnosis” of SBS.
Section 1 - “Shaken babies” – sound science or an unproven hypothesis?

The history of SBS in the courts, however, reveals a weak point in our adversarial justice system. Physicians and prosecutors working with a widely accepted but unproven model of SBS have blurred the distinction between medical opinion and legal fact. The different standards and practices of the two disciplines—medicine and law—have allowed a theory to become institutionalized before it was proven. The science of paediatric head injury is now catching up, to reveal that what doctors believed in the late 1970s is not necessarily true. This cycle is common in medicine, and in most other fields as well, though not always with such tragic side-effects.

In this section, I describe the evolution of SBS as a clinical entity, and highlight some of current contentious forensic clinical issues in relation to the reliability of the triad as a marker of inflicted head injury in infants.

The evolution of Shaken Baby Syndrome as a forensic medical diagnosis

Although the possibility that infantile subdural haemorrhages may be traumatic in origin had been raised early in the 20th century, it was not until 1946 that John Caffey, an American paediatric radiologist, reported an association between limb fractures and SDH in a number of infants, concluding that although none of the infants had a history of trauma, this was the likely cause for the SDH. Later in the 1960s, animal experiments on monkeys exploring motor vehicle whiplash injuries chimed with observations by James Weston, a medical examiner (forensic pathologist) affiliated to the University of New Mexico, that shaking might cause some child deaths. Subsequently, seminal articles in the 1970s by A. Norman

---

716 For a useful but succinct summary of the evolution of SBS as a clinical entity for the non-specialist, see Cobley, n. 710 above, in particular Chapter 2.
717 John Caffey 'Multiple fractures in the long bones of infants suffering from chronic subdural hematoma' (1946) 56 AJR 163.
718 A. K. Ommaya, F. Faas and P. Yarnell 'Whiplash injury and brain damage' (1968) 204 JAMA 75.
Guthkelch, a Hull neurosurgeon, and Caffey reinforced this mechanism. Caffey later coined the term “whiplash shaken infant syndrome” to describe such cases and their cause. From the 1980s, the medical profession had developed and accepted a hypothesis concerning SBS. The perceived mechanism was that, usually in a fit of anger in response to persistent crying, a carer might forcefully grip a small child around the chest and violently shake it, causing the head to flop back and forward with great force. The theory was that this forceful shaking ruptured bridging veins over the surface of the brain and blood vessels at the back of the eye(s), while shearing forces tore the nerve fibres within the brain, leading to the triad of intracranial injuries that I listed earlier.

For years conventional expert medical opinion was to the effect that even in the absence of other signs or symptoms of trauma such as bruises or fractures, if known natural causes of the triad’s components were excluded, this triad was diagnostic of abusive head injury. Numerous criminal prosecutions and civil child protection actions were successfully pursued based exclusively on this medical evidence.

---

720 A. N. Guthkelch 'Infantile subdural haematoma and its relationship to whiplash injuries' (1971) 2 BMJ 430. Guthkelch proposed that whiplash injury caused subdural haemorrhages in infants by tearing the veins in the subdural space.


However, the televised 1997 murder trial of “Boston nanny” Louise Woodward following the death of 8-month-old Matthew Eappen, the son of two doctors, marked the beginning of a shift in this universal consensus.

Woodward, then a 19-year-old English au pair, was caring for Matthew when he collapsed. He was rushed to hospital and was found to have the SBS triad, together with a healing (i.e. old) wrist fracture, and a skull fracture. The hospital team suggested his head injuries were the result of violent shaking, the other evidence of trauma reinforcing that view. Matthew died a few days later. While Woodward admitted that she shook Matthew “lightly” when she could not wake him after a nap, she denied any violent shaking assault. She was charged with murder.

Multiple expert witnesses were called for the prosecution and defence. Although both sides agreed that Matthew died from massive intracranial haemorrhage, the case turned on their diametrically opposing theories to explain the bleeding. The defence case posited re-bleeding into a birth-related (natural) SDH.

The jury found Woodward guilty of second-degree murder and imposed a mandatory life sentence. Subsequently trial judge Hiller Zobel, applying his discretionary post-conviction authority, reduced the conviction to involuntary manslaughter and released her for time served. He held that while she may have handled Matthew roughly, the evidence did not support the level of assault claimed by the prosecution.

In the aftermath of Woodward, a joint response from 72 US paediatric child abuse specialists uncompromisingly ridiculed the defence explanation for the intracranial bleeding as a

725 Because of specific healing mechanisms, it is not possible to assign an age estimate to skull fractures.
726 A number of defence experts supported the view that Matthew may have had a pre-existing subdural haemorrhage that a minor injury caused to re-bleed. At the time, this was a controversial concept that is now fully accepted by the clinical community. See, for example, Lori D. Frasier and others ‘International issues in abusive head trauma ’ (2014) 44 (Suppl 4) Pediatr Radiol S647, page 81.
“courtroom diagnosis”, their retort emphasising that the triad was pathognomonic of infant shaking:

The shaken baby syndrome (with or without evidence of impact) is now a well characterized clinical and pathological entity with diagnostic features in severe cases virtually unique [my emphasis] to this type of injury—swelling of the brain (cerebral edema) secondary to severe brain injury, bleeding within the head (subdural hemorrhage), and bleeding in the interior linings of the eyes (retinal hemorrhages). Let those who would challenge the specificity of these diagnostic features first do so in the peer-reviewed literature, before speculating on other causes in court. 728

As the Woodward trial revealed, by the late 1990s the science underlying the expert evidence on alleged shaken babies was contentious, and expert paediatricians such as myself began to couch our forensic conclusions more carefully. Rather than suggest that shaking was the only explanation for the triad, we would now acknowledge that alternative causes had been considered and excluded, 729 before concluding that, in the absence of any known alternative, SBS was the likeliest explanation. 730

In the 20 years since Woodward, controversy about the forensic significance of the triad has continued and expert views on the diagnosis of SBS are arguably even more contentious.

**Controversies in the diagnosis of SBS**

While there is a general agreement that shaking a baby is potentially very dangerous, and most paediatricians and associated professional bodies still consider that the SBS triad usually represents an inflicted injury, 731 in recent years a vociferous professional minority of

---

clinicians, scientists and legal scholars have contested this established consensus, resulting in an on-going polarised, acrimonious and often highly personal dispute that straddles a number of knowledge domains, and where both institutions and individuals defend their positions or attack alternative theories with an evangelical zeal that seems at odds with the normative expectation that scientific inquiry is objective and grounded on falsification.

The key scientific controversies are focused on a variety of alternative mechanisms purported to explain some or all of the clinical findings that make up the SBS triad, or suggest that there may be other natural but as yet unknown causes for the triad’s individual elements. While some proffered alternative explanations have no accepted scientific validity, others provide more scientifically plausible challenges to the conventional paradigm of the triad’s causation.

The mechanisms that claim to provide exculpatory explanations for some or all of the elements of the triad can be grouped into two broad but potentially overlapping areas, focused around the biomechanics of injury in such cases, or on the possibility of alternative natural physiological or pathological causes for some or all of the SBS findings. It is

---


733 For example, for example, that paroxysmal coughing in an infant might cause SDH and RHs: J. F. Geddes and D. G. Talbert 'Paroxysmal coughing, subdural and retinal bleeding: a computer modelling approach' (2006) 32 (6) Neuropathol Appl Neurobiol 625; a hypothesis recently disproved: Naz Raoof and others 'Retinal haemorrhage in infants with pertussis' (2017) 102 (12) Arch Dis Child 1158.

734 For example, that vaccines may induce vitamin C deficiency leading to intracranial bleeding: M. D. Innis 'Vaccines, apparent life threatening events, Barlow’s disease, and questions about “shaken baby syndrome”' (2006) 11 J Am Phys Surg 17; but see also R. A. C. Bilo and others 'Using the table in the Swedish review on shaken baby syndrome will not help courts deliver justice' (2017) 106 (7) Acta Paediatr 1043.


neither appropriate nor practical to review the various putative alternatives here, but by way of illustration, let me highlight one of these alternative putative explanations, and a separate important forensic issue.

**Low-level falls and SBS causation**

There is general agreement that although normal domestic handling does not cause such injuries, the minimum amount of force required to injure an infant by shaking is unknown. Given this lack of an agreed threshold force to cause injury, some claim that low-level falls – for example, being dropped from a carer’s arms - could cause the same pattern of injuries.

While various case series show that it is very uncommon for infants to be seriously injured after short falls, it is now evident that very rarely short falls in complex circumstances may produce the triad, and exceptionally may prove lethal.

This shift in understanding has occasionally led to the quashing of convictions that were originally based on the rejection of that explanation (see later). Such cases highlight that expert forensic assessment must take into account the unique features of each case. Thus, as I noted in Chapter 5, any proffered injury history must be individually considered from a biomechanical perspective.

---

741 See Chadwick et al, n. 739 above.
743 Jonathon Hughes and others 'Biomechanical characteristics of head injuries from falls in children younger than 48 months' (2016) 101 (4) Arch Dis Child 310.
Timing of injuries – the “lucid interval”

One controversial but important forensic issues in relation to alleged inflicted head injury in infants is whether or not a young child may behave reasonably normally for a period after a significant head injury. The conventional view has been that such intracranial injuries are usually so severe that there would be obvious symptoms immediately.

The forensic issue here is that, in the absence of a lucid interval, the last person with an infant when it became unwell has usually been considered responsible for abuse. However, there are rare case reports of children with accidental head injuries who apparently appear unaffected before deterioration, and even death. Further, it is now accepted that the mechanism of brain injury in shaking is more likely to represent a hypoxic-ischaemic insult rather than a direct traumatic one, and the onset of signs of injury may be gradual, and depend on the severity of the injury. Evolving understanding of the difficulties in pinpointing the timing of the injury event has been an important factor in reconsidering the safety of a small number of infant murder convictions.

Thus, as the brief review above illustrates, the reliability of the established evidence base for the conventional SBS paradigm has been aggressively questioned, defended, and remains an area of significant controversy.

---

744 See Chadwick et al, n. 739 above. However, as Gilliland has noted: “in all of the cases [of SBS] where information was supplied by someone other than the perpetrator, the child was not normal during the interval”, M. G. Gilliland 'Interval duration between injury and severe symptoms in non-accidental head trauma in infants and young children' (1998) 43 (3) J Forensic Sci 723, at 724.


Section 2 - Expert evidence in criminal trials of SBS – are paediatricians and their institutions the problem rather than the solution?

In this section, I argue that many of the potential frailties in paediatric expert evidence I identified earlier are notably evident in the context of alleged SBS cases. And while I restrict my comments to criminal prosecutions in such cases, it is important to note that the issues I discuss are equally relevant to paediatric expert evidence provided in family court proceedings, albeit that the two legal frameworks have different functions and require different standards of proof.

Here, I focus initially on the influence of the various professional paediatric institutions on the expert evidence provided by individual paediatric experts to the courts, and how, by their guidance and other professional consensus statements, such institutions provide a powerful imprimatur of evidential reliability to the courts.

I argue that several of the factors I identified earlier in this thesis – the potential contaminating effect of paediatric institutions’ professional values which put children’s interests at the core of their purpose (Chapter 1), the delay often exhibited by institutions in acknowledging slippage in the validity of established scientific dogma within their knowledge domains (Chapter 5), and the emotional impact of the heated SBS debate in which they actively engage (Chapter 4), may all impact on the objectivity and content of the professional interpretation and guidance on SBS they provide for their members.

Second, having discussed these issues from an institutional perspective, I then turn to explore how such factors impact on the individual paediatric expert dealing with SBS cases.

Thereafter, I consider how the courts have responded to the particular issues around paediatric expert evidence in SBS cases.

First, I turn to the institutional perspective.

**Institutional denial, bias, and value-based preferences**

As I noted in Chapter 1, one of the key demands of experts is their impartiality. This issue is much broader than simple prejudice towards the instructing side, and as I discussed earlier in this thesis, particularly Chapters 2 and 3, in the paediatric sphere there is a spectrum of biases that may threaten expert objectivity.

While many of the generic biases that I discussed in Chapter 3 may coalesce in an individual expert’s forensic interpretation of a child with a suspicious head injury, because of their broader professional impact, arguably the most significant biases in the context of SBS may occur at the institutional level. The notable denials by the various international paediatric institutions that a scientific controversy exists about the forensic significance of the isolated triad in SBS may reflect such as bias.

Thus, for example, while RCPCH child protection representatives acknowledge that cases of alleged SBS involving the triad alone “cause us a little uncertainty”, the paediatric institutions in most common law jurisdictions continue to maintain an on-going and aggressive position on the high specificity of the isolated triad as a marker of inflicted injury, and dismiss claims of a scientific controversy about the evidence base to support such assertions or the possibility of alternative potential causes of the triad as contrived.

---


And this is not a dispassionate scientific argument; it is emotional and highly charged, with a flavour of religious zealotry. Here we see a staunch defence of the holy trinity of the triad, contrasting claims that Galileo-like truths are being blindly dismissed as “scientific blasphemy”, and institutional assertions that there is a small cabal of SBS heretics peddling pseudoscience to the courts and putting vulnerable children at risk. As I discussed in Chapter 4, such emotional overlay, even at institutional level, must inevitably threaten the objectivity with which individuals interpret the scientific evidence on behalf of these institutions, and impact on the evidential reliability of expert evidence dependent on institutional guidance.

And where does that emotion come from? From an institutional perspective, the answer is perhaps straightforward. Here I suggest we see the professional values of the “child’s interests” that I discussed in Chapter 1 coming to the fore, for there is explicit institutional concern that by acknowledging uncertainties about the forensic status of the SBS triad, and thus presenting less certain expert opinions to the courts, experts may place vulnerable children at risk by “providing lawyers with new ammunition to question valid scientific data”.

And of course, while it can be reasoned that an equally important value-based argument is that a non-abused child should not be unjustifiably removed from their family or their innocent parent wrongly convicted, it is the preceding perspective that seems to dominate.

---

756 Niels Lynøe and others 'Easier to see the speck in your critical peers' eyes than the log in your own? Response to Debelle et al' (2018) 103 (7) Arch Dis Child 714.
This defensive institutional standpoint seems to reflect the concern I raised earlier in this thesis about the potential for confusion between the child advocacy role of the paediatric institutions and child protection paediatricians, and the scientific objectivity expected of the forensic expert witness, a “noble cause corruption” effect that has been highlighted by a number of judges in relation to paediatric expert evidence.  

*The pangs of falsification?*

And of course there may be other reasons why paediatric institutions such as the RCPCH, and thus many individual child protection paediatricians, still firmly hold to the traditional paradigm and defensively assert that the isolated triad is strongly indicative of inflicted injury. Here I return to an issue I highlighted in Chapter 5 – that it takes time for revised understanding of long-held beliefs to become established, and that the professional bodies that represent scientific communities are often the last to support such reversals.

Yet while such falsification may be uncomfortable, in the case of forensic science, such institutional dogmatism must inevitably mean that some individuals within the justice system are poorly served during such transitions. This is an issue to which I will return later in this chapter.

And what about the forensic child protection foot soldier? In the face of such influential institutional tenacity, it is perhaps unsurprising that the majority of paediatric experts working at the frontline of child protection would generally default to the relative comfort of

---

757 For a Canadian perspective of paediatricians’ objectivity undermined by advocacy, see Gruspier, n. 763 above, at 64-65. See also the judicial criticism of the contaminating effect of advocacy on objectivity of the evidence of renowned expert paediatrician Dr Camille San Lazaro, in Lillie & Reed v Newcastle City Council [2002] EWHC 1600 (QB).

758 Personal communication, Dr Geoff DeBelle, Officer for Child Protection, RCPCH, 5 December 2017); see also Narang, n. 731 above.

759 Prasad, n. 714 above.

760 Typified by the FBI's initial rejection of claims of fingerprint misidentification in Brandon Mayfield case; see United States Department of Justice, Office of the Inspector General: *A Review of the FBI's Handling of the Brandon Mayfield Case* (US Department of Justice, Office of the Inspector General, Oversight and Review Division 2006).
the institutional view. Thus in trying to bolster the status of the traditional paradigm on SBS causation, I suggest that a claim that since “most paediatric experts agree with the institutional view on the causation of SBS” the institutions’ view must be correct is a specious one.

**Uncertain science, selection bias and circular reasoning**

Separate from the potential influence of ideology, earlier in this chapter I noted that a recent Swedish systematic review has raised concern that the evidence base purported to support paediatric institutional assertions that the triad alone is a secure marker of traumatic shaking is in reality relatively weak, and primarily based on a modest database of confessions of uncertain provenance.

These concerns take me back to Chapter 5, where I raised fundamental questions about the foundations of the evidential hierarchy underlying the evidence base for the diagnosis of abusive injury, a concern that others have also raised. I do not intend to reprise my criticisms of those standards here, but it is worth simply re-emphasising two key points I argued earlier. First, that confession “evidence” does not represent “scientific evidence” in any conventional sense. And second, in a similar vein, that the use of the outcome of previous criminal or civil cases surely cannot be used to substantiate the clinical interpretations that underpin expert opinion evidence provided in other cases, for as Edmond has clearly expressed from a broader forensic evidence perspective, “[u]nsound opinions and

---

761 Narang, n. 731 above.
762 N. Lynøe and others 'Insufficient evidence for 'shaken baby syndrome' - a systematic review' (2017a) 106 (7) Acta Paediatr 1021.
techniques are not validated because someone is convicted of a crime."\textsuperscript{765}

Despite such disquiet, such a rationale remains integral to the institutionally supported evidence base with which UK paediatricians are expected to justify their conclusions.

And of course, as I also previously discussed, underlying these two inter-twined putative validation criteria for forensic clinical interpretations of the triad lies a fundamental concern about circular reasoning in the various case studies allegedly supporting such assertions.\textsuperscript{766}

Yet in the face of such direct criticisms of the selective use of the literature and the application of circular arguments to justify their standpoint, rather than directly address such critiques the RCPCH and other paediatric institutions response has been to attack those raising the concerns, rather than address the substantive criticisms.\textsuperscript{767}

\textit{Glimmers of change?}

Despite this apparent institutional intransigence, there may be some evidence of a less entrenched position emerging. As I already noted in Chapter 6, there is now some backtracking evident in clarification of the purported utility of probabilistic clinical predictive tools to support a forensic diagnostic conclusion in babies with possible SBS.\textsuperscript{768}

There is also now some evidence that the institutional perspective is softening in relation to the specificity of the triad as a marker of abuse, with recent acknowledgment by the RCPCH child protection representatives that while the features of the triad may be \textit{associated} with

\textsuperscript{765} G Edmond, 'Is reliability sufficient? The Law Commission and expert evidence in international and interdisciplinary perspective (Part 1)' (2012b) 16(1) IJEP 30, at 57.

\textsuperscript{766} Lynøe et al., (2017a), n. 762 above. For a clear overview of concern about circular reasoning as it applies to the validity of shaken baby science, see Niels Lynøe and others 'Is accepting circular reasoning in shaken baby studies bad science or misconduct?' (2017b) 106 (9) Acta Paediatr 1445.

\textsuperscript{767} Lynøe et al., (2018), n. 756 above.

\textsuperscript{768} Helena Pfeiffer and others 'Clinical prediction rules for abusive head trauma: a systematic review' (2018) 103 (8) Arch Dis Child 776.
SBS, they cannot be regarded as diagnostic of that causation. Later in this chapter I discuss how the courts might interpret expert opinion of such an association.

So far in this chapter, my focus has been on clinical and paediatric professional issues in relation to alleged SBS cases. Having provided an overview of the topic for the non-specialist, I then illustrated how many of the potential frailties in paediatric expert evidence that I discussed in previous chapters coalesce within such cases. In particular I showed how these concerns about paediatric expert evidence are particularly explicit in cases involving only the triad of cranial injuries.

In the remainder of this chapter, I change perspective, and explore the particular difficulties that cases of alleged SBS present to the courts, and explore how the courts have responded to those challenges.

---

769 Alison Mary Kemp, Sabine Ann Maguire and Geoff D. Debelle 'Response to comments from Professor Niels Lynøe et al and Dr Nicholas R Binney et al' (2018) 103 (7) Arch Dis Child 715.
Section 3 – Admissibility of expert evidence - a brief overview

There is a wealth of academic scholarship and comment available relating to the historical origins of the use of expert evidence in a variety of jurisdictions and legal settings, the wide range of problems associated with such evidence, and the evolution of the rules that govern it. Here, I will only concern myself with those relatively recent and current developments involving expert evidence that directly relate to the specific intertwined issues of admissibility and reliability that are relevant to my later discussion.

Further, although I acknowledge that the mixed heritage of Scots Law means that the law and practice in relation to expert evidence is different in Scotland compared with England & Wales, for the purposes of this thesis I will consider them together because, setting aside procedural issues, the general principles in relation to the admission of expert evidence in both jurisdictions are similar.

Although this part of the chapter is primarily concerned with evidentiary reliability and the admissibility, or exclusion, of (paediatric) expert testimony, it is important that I also acknowledge that rules on expert evidence admissibility sit within a broader constellation of both the general common law principles that apply to the admissibility of all kinds of evidence, together with specific rules of evidence law and various procedural rules – for example, in relation to hearsay evidence, vulnerable witnesses, or alleged sexual offenses. Such rules vary with the relevant jurisdiction. Some of these general rules and procedural constraints impact indirectly on aspects of expert evidence, while others apply specifically to it.

Having acknowledged this broader context in which the rules of evidence related to expert evidence sits, I now briefly set out the general approach of the courts to the admissibility of

---

expert evidence, before discussing, in the context of that background, how they have dealt with expert evidence in SBS cases.

**General principles – the common law admissibility test**

Until recently, the courts in England and Scotland have generally exhibited a *laissez-faire* approach to the admission of expert evidence.\(^{771}\) Assuming opinion evidence was relevant to the disputed matter, such evidence simply had to be considered “sufficiently reliable” to be admitted, although no specific guidance was available for the judiciary as to how they might assess the degree of sufficiency.\(^{772}\) Once admitted, questions of reliability went to the weight of the evidence, and it was left to the adversarial process—involving cross-examination and sometimes the rebuttal testimony of other experts—to expose any scientific or other weaknesses.

While these principles of admissibility of expert evidence generally apply across all fact-finding settings, many of the key cases that have informed the development and clarification of legal criteria for admissibility have involved civil rather than criminal disputes, notably medical litigation or health and safety issues.\(^{773}\) Although there are differences between civil and criminal litigation and caution is needed in applying lessons in case law from one setting to the other,\(^{774}\) such distinctions are less relevant when it is the principles that are being summarised.

As I already noted, it is not my intention to provide a detailed review of the common law principles that govern the admissibility of skilled witness evidence under Scots (and English)

---


\(^{772}\) R v Reed, Reed, & Garmson [2009] EWCA Crim 2698, [2010] 1 Cr App R 23, at [111]. While the instant case was focussed around the admissibility of expert evidence in relation to the interpretation of low-copy DNA evidence, the general principles of the admissibility of expert evidence in relation to both the reliability of underlying science and the determination of the appropriate expertise of individual experts were reviewed.


\(^{774}\) See Redmayne (2001), n. 288 above, at 101.
law. Overviews of these principles are widely available elsewhere. However, it is appropriate to record that the recent 2016 case of *Kennedy v Cordia (Services) LLP* (hereinafter *Kennedy*), an appeal from the Inner House of the Court of Session heard in the UK Supreme Court (SC), provided a useful summary of the current common law principles that govern the admissibility of skilled witness evidence under Scots (and English) law.

Indeed, while the dicta in *Kennedy* particularly acknowledged Scots case law, the range of international case law cited emphasised that these principles reflect the general approach to the admissibility of expert opinion evidence across common law jurisdictions. *Kennedy* has since been applied in the Scottish criminal and family courts, and in the English civil courts.

Briefly, in *Kennedy*, building on the South Australian case of *R v Bonython*, the SC identified the key principles that determine the admissibility of expert evidence, and provided four key questions that must be considered:

---


778 ibid, at [39] - [41].

779 Jones v HMA [2016] HCJAC 65. Here the Appeal Court upheld a Sheriff’s ruling that a police officer was sufficiently skilled to provide opinion evidence about drugs, and that his professional links with the prosecuting authorities did not undermine his independence and impartiality.

780 In *Q v P* [2016] Fam LR 54, a complex hearing involving the impact on children of a mother’s proposed relocation within the UK, the court did not admit a child psychologist’s report as it did not assist the court in resolving the issues before it.

781 Deloitte LLP v HMRC Commissioners [2016] UKFTT 0479 (TC), where parts of an expert’s report were ruled inadmissible, since they were matters for the court to decide.

782 *Kennedy*, n. 776 above, at [44].
1) Will the proposed expert evidence assist the court?

a) Is the expert evidence needed for the fact-finder to reach a decision?

Referring to *R v Turner*, 783 and *Wilson v HMA*, 784 both incidentally involving expert psychological opinion evidence, the SC emphasised the key threshold test of admissibility was the necessity for specialised knowledge or experience beyond that of a lay juror or judge to assist the Court to reach a safe conclusion on the facts. 785

b) However, although the knowledge focus may be outwith common understanding, per *Davie v Edinburgh Magistrates*, 786 the SC reiterated that an expert’s opinion should not simply be an *ipse dixit*. The opinion must be explicitly justified by being set in an appropriate scientific or other evidence base, for “what carries weight is the reasoning, not the conclusions”. 787

c) However, notwithstanding the application of such expertise, such opinion evidence was not determinative: the fact-finder may choose to accept or reject all or part of the expert’s opinion; the expert cannot usurp the functions of the judge or jury.

2) Has the witness the necessary knowledge and experience?

The witness must demonstrate to the court that he or she has relevant personal knowledge and experience before they may refer to, or draw on, a general body of knowledge and understanding from a relevant field of expertise. 788 While the prime responsibility for ensuring that a proffered expert had the appropriate knowledge and experience to provide relevant and reliable evidence fell to those instructing the

---

783 (1975) QB 834, (1975) 2 WLR 56.
784 2009 JC 336.
785 However, in some circumstances, defining the boundaries of “common knowledge” may problematic, and the common knowledge rule has been abolished in some jurisdictions; as in Section 80 (b) of the Australian Uniform Evidence Act 1995.
786 1953 SC 34.
787 Davie, n. 786, at [40].
788 Kennedy, n. 776 above, at [50].
expert, the SC also acknowledged, per McTear v Imperial Tobacco, the judge’s responsibility in that regard. However there was little guidance provided as to how to validate this necessary expertise.

3) Is the witness impartial in his/her assessment of the evidence?

Independence and impartiality is a question of admissibility; it does not just go to the weight to attach to the evidence. Here the SC emphasised the adoption, in both civil and criminal matters in Scotland, of the duties of an expert witness as defined in The Ikarian Reefer (see Chapter 1), highlighting that a skilled witness’ duty was to the court and not the party instructing them, that any opinion must be based on an impartial approach to interpretation of the case evidence and any relevant published literature. While such an a provision would seem to deal clearly with what might broadly be called the “hired gun” effect, and would seem to address individual bias, it seems less focused on the problem of what might be called institutional or system bias, arising from what has been termed the “role effect”, encompassed within Dror’s taxonomy of bias pyramid (Figure 7), and originating in the case of paediatricians from their background professional values and broader caring role.

This expectation leads to the final and perhaps most challenging question the SC raised, and the issue particularly relevant to this chapter.

---

789 ibid, at [57].
790 McTear v Imperial Tobacco 2005 2 SC 1.
791 Despite this lack of clarity here, as I discuss later, specific guidance is available in relation to sudden infant deaths and shaken baby cases on criteria with which to assess a proposed expert’s competence to provide expert evidence; Baroness Kennedy (2004), n. 131 above. See also: R v Henderson (Keran Louise) [2010] EWCA Crim 1269, [2010] 2 Cr App R 24.
792 Per Lord Caplan in Elf Caledonia Ltd v London Bridge Engineering Ltd September 2, 1997 (unreported) at pp 225–227, and Wilson, n. 784 above, paras [59-60].
794 Guidance: Cognitive Bias Effects Relevant to Forensic Science Examinations, FSR-G-217 (Forensic Science Regulator 2015) 1, 4.2.7.
4) Is there a reliable body of knowledge or experience to underpin the expert’s evidence?

Here the SC noted that:

In many cases where the subject matter of the proposed expert evidence is within a recognised scientific discipline, it will be easy for the court to be satisfied about the reliability of the relevant body of knowledge. There is more difficulty where the science or body of knowledge is not widely recognised. Of course, as I discuss later, underneath this somewhat bland statement lurks one of the most problematic challenges posed by forensic science and other expert witness testimony - ensuring that only expert opinion evidence founded on reliable science is admitted before the courts.

The SC used the recent Scottish High Court’s rejection of the grounds of the appeal in Young v HMA to illustrate the application of such an exclusionary rule based on uncertainties about the evidential reliability, where the Court refused to admit novel scientific evidence based on “case linkage analysis”. But while Young may provide a good illustration of the appropriate exclusion of unreliable expert testimony—a significant error rate evident in the methodology arguably made the determination relatively straightforward—in many cases such distinctions may be much less clear-cut, and the absence of a defined reliability threshold (or guidance on assessing it) makes such determinations problematic.

---

795 Kennedy, n. 776 above, at [55].
798 The Court was concerned that expert evidence was founded on relatively recent and limited academic research using a methodology of as yet uncertain reliability.
Of course it would be wrong to suggest that the courts *routinely* deny the admission of novel expert forensic science evidence. However, while the courts have justified such admissions on the grounds that it “would be entirely wrong to deny to the law of evidence the advantages to be gained from new techniques and new advances in science”, as the wrongful conviction of Mark Dallagher linked to flawed ear print expert evidence illustrates, such a *laissez faire* approach may come at a price.

Thus, it seems reasonable that, before admitting expert evidence based on relatively novel methodologies such as case linkage analysis or gait analysis, the courts should carefully scrutinise the reliability of the forensic application of these techniques.

It is trite to observe that epistemic confidence in the reliability of a technique or scientific interpretation builds with time. What might later prove to be sound science takes time to become “an established body of knowledge” within the appropriate knowledge domain. Similarly, as the difficulties associated with the admissibility of Low Template DNA (LTDNA) evidence illustrate, new developments derived from older techniques may also struggle to secure a firm foothold in legal fact finding and for a related interpretive jurisprudence to emerge. Clearly a balance has to be struck.

---


800 Clarke, n. 799 above, per Steyn LJ, at [430].


802 For example, Professor Meadow’s flawed interpretation of novel evidence on the epidemiology of recurrent SIDS in Clark, n. 77 above.


805 Kennedy, n. 776 above, at [56].

However, what Young does suggest is that the Scottish courts may be beginning to reflect a recent and significant shift in the England and Wales courts, whereby a hitherto _laissez-faire_ attitude to the admission of expert evidence has moved to an explicitly more cautious and restrictive approach. Here flawed paediatric expert evidence played a prominent role in exposing the need for reform.

**Evolving approaches to admissibility of expert evidence in England and Wales**

In England, cases such as _Dallagher_ and _Clark_ illustrate the potentially negative impact if a liberal approach to the admissibility of forensic science and medical expert evidence is adopted, “…without sufficient regard to whether or not [such evidence] is sufficiently reliable to be considered by a jury”, and justify a more active process to validate the evidential reliability and therefore admissibility of proposed forensic and medical expert evidence. As the appellate court noted in relation to Meadow’s flawed statistical evidence in _Clark:_

> We are quite sure that the evidence should never have been before the jury in the way that it was when they considered their verdicts. If there had been a challenge to the admissibility of the evidence we would have thought that the wisest course would have been to exclude it altogether.

In the aftermath of _Clark_ and associated flawed convictions, a major review by the House of Commons Science and Technology Committee called for some significant changes to the rules and practices in relation to expert forensic and medical evidence in the English courts. One of their key proposals was that a pre-trial “gate-keeping” test should be developed for the admissibility of expert evidence. Responding to this concern, the English Law Commission (2011), n. 101 above, at para 1.8. G Edmond, 'Legal versus non-legal approaches to forensic science evidence' (2016) (20) IJEP 3. Clark, n. 77 at [177]. See House of Commons Science and Technology Committee, 'Forensic Science On Trial' report, n. 100 above. In March 2005, the House of Commons Science and Technology Select Committee examined the use of scientific evidence in court. Noting that when miscarriages such as _Canning_ and _Clark_ occurred, expert witnesses were often publically vilified, the Committee considered that such
Commission embarked on its own major review, producing a consultation paper on expert evidence in criminal trials in 2009, followed by a final report to Parliament in 2011. Much has been written about the Law Commission’s proposals and I do not intent to discuss them in any detail. However, it is worthwhile noting that flawed convictions involving paediatric expert evidence were a noticeable feature in their justification for reform, and paediatric topics such as SBS and recurrent sudden infant deaths were prominent in the examples the Commission chose to illustrate how their proposed reforms would work in practice.

The Commission proposed that the admissibility of expert evidence in criminal proceedings in England should be set within a statutory framework, the key proposal involving a new statutory reliability test. This test would be applied in combination with refinement and codification of the existing common law principles of “assistance”, “expertise” and “impartiality” noted above. It would be supported by a variety of other changes: modification to pre-trial disclosure rules, court-appointed experts and amendments to the Criminal Procedure Rules. There were also proposals for judges to receive scientific training to support their decision-making.

These proposals were justified on the grounds that the adversarial process and other trial procedures were unable to adequately expose flaws in expert evidence, leaving juries floundering to decide on the probative value of various forms of incriminating expert evidence.

---

811 Law Commission (2009), n. 101 above.
812 Other recommendations related to pre-trial disclosure, court-appointed experts, and amendments to the Criminal Procedure Rules (CrimPR), together with proposals for expert witness regulation.
In the event, the Government declined to enact the Law Commission’s draft Bill, principally because of concern that the cost and other claimed benefits of the inevitable pre-trial hearings associated with some cases would not be sufficiently offset by financial and system savings elsewhere.  

Instead, the Government invited the Criminal Procedure Rule Committee to consider developing relevant modifications to the Criminal Procedure Rules based on the Commission’s proposals, in order to “increase the likelihood of the trial judge and the opposing party, where appropriate, challenging expert evidence”, thus “go[ing] some way towards reducing the risk of unsafe convictions as a result of unchallenged inappropriate or unreliable expert evidence”. It has been suggested that these subsequent amendments to Part 33 of the Criminal Procedure Rules, parallel changes in associated Practice Directions, together with a variety of related legal training initiatives represented “a novel way of implementing [most of] an excellent Report”, albeit from a different route.  

What seems clear is that, albeit the common law continues to provide the criteria by reference to which the court must assess the admissibility and weight of expert evidence, these various initiatives indicate that the previous laissez faire attitude to the admission of expert evidence in the criminal courts in England and Wales has changed.


Now set within CrimPR Part 19, and combined with revised Practice Direction CrimPD 19A. For an assessment of the impact of these relevant amendments, see Michael Stockdale and Adam Jackson ‘Expert Evidence in Criminal Proceedings: Current Challenges and Opportunities’ (2016) 80 (5) J Crim Law 344.

Not only do the relevant Criminal Procedure Rules now provide detailed guidance on the content of expert reports so that the court may assess the admissibility of such evidence, but the judiciary have responded to the longstanding concerns with a more interventionist attitude, typified by the approach taken on the admissibility of Low Template DNA evidence. It is evident that the Law Commission’s reports represent a highly influential backdrop to such determinations, and seem to provide a template for judicial admissibility assessment, as noted in the relevant section of the Criminal Practice Directions:

Nothing at common law precludes assessment by the court of the reliability of an expert opinion by reference to substantially similar factors to those the Law Commission recommended as conditions of admissibility, and courts are encouraged actively to enquire into such factors.

As I will now discuss, this changing approach of the English courts to expert evidence has started to emerge in alleged child abuse cases, including SBS criminal trials, with significant implications for expert witnesses in that jurisdiction. And this involves not only enhanced scrutiny of the underlying science, but also how experts present their related opinions.

**Controversy, expert polemic, and obscurantism in SBS – the courts' responses**

Returning to the focus of this chapter, as the alternative hypotheses and professional tensions around the SBS interpretive paradigm emerged into the adversarial legal environment where alleged SBS cases are decided, the courts have had to consider and manage these conflicting expert explanations. In this regard, two appellate cases in the English courts have been particularly significant.

---

818 Reed, n. 772 above.
819 ibid, at [111].
Conjoined SBS appellate cases: Harris and Henderson

In the aftermath of the Cannings judgment in relation to alleged infanticide linked to multiple sudden infant deaths, a review of 297 infant death convictions over the preceding decade included 97 cases related to apparent SBS. In the face of concern about the safety of the convictions in two cases (Harris and Cherry), and emerging doubts more generally about the reliability of expert evidence in SBS, in 2005 the English Court of Appeal (Criminal Division) (CACD) heard conjoined applications involving appeals from four individuals against their respective convictions (manslaughter, murder, manslaughter and inflicting grievous bodily harm (GBH)) for shaking infants in their care. (R v Harris, Rock, Cherry and Faulder, hereinafter Harris).

The conjoined approach was founded on a common thread that fresh evidence—the so-called Geddes unified hypothesis—cast doubt on the safety of these convictions. An additional focus in two appeals centred on the claim that a short fall or minor injury may have caused the injuries found. While the rejecting the Geddes theories as primarily speculative, two of the appeals were successful and the respective convictions for Harris for manslaughter and Faulder for GBH were quashed. Cherry’s appeal was dismissed, while in Rock’s case, the court set aside his murder conviction, and substituted one of manslaughter, that being the appropriate offence in the circumstances.

Lorraine Harris’ successful appeal against her conviction in 2000 for the manslaughter of her four-month-old son Patrick is the most relevant to the focus of this chapter because the case involved the presence of the triad alone and the prosecution case had been entirely dependent.

---

822 Goldsmith, n. 550 above.
823 Harris, n. 799 above. Rock had already lodged an appeal while Faulder’s appeal arose following referral from the Criminal Cases Review Commission.
824 The concept that brain hypoxia (lack of oxygen) and resultant neurovascular and other responses may explain the triad; see J. F. Geddes and others 'Neuropathology of inflicted head injury in children: I. Patterns of brain damage' (2001) 124 Brain 1290; J. F. Geddes and others 'Neuropathology of inflicted head injury in children: II. Microscopic brain injury in infants' (2001) 124 Brain 1299.
on expert opinion evidence that only shaking could have caused Patrick’s injuries. In all the other cases, there were additional injuries as well as at least some elements of the triad.  

Significantly, the CACD in *Harris* held that the presence of the triad could not be regarded as a proven fact regarding the causation of SBS; rather it considered that it represented a medical hypothesis. Thus while the CACD held that the classic triad may be considered as strongly indicative of a shaking injury, “the mere presence of the triad on its own cannot automatically or necessarily” lead to the conclusion that an infant had been shaken.  

Further, while there was general agreement that simple “rough handling” was very unlikely to produce the various elements of the triad, it was also unknown what force was required to cause any or all of the elements of the triad, and there were examples where a small degree of force or minor falls had caused serious injury. In Lorraine Harris’ particular case, she had described shaking Patrick gently to revive him after he became unwell. Given the lack of a known threshold force to cause the triad, it could not be said that this action represented unlawful force.  

The CACD also acknowledged that there might be other, as yet unknown, natural causes of the triad. Reflecting on the many uncertainties in relation to infant death (and injury), the CACD highlighted the dicta in *Cannings*:

“[A] great deal about death in infancy, and its causes, remain as yet unknown and undiscovered…What may be unexplained today may be perfectly well understood tomorrow. Until then, any tendency to dogmatise should be met with an answering challenge.”

---

825 The Court held that Rock had unlawfully killed his partner’s 13-month-old daughter, but the evidence did not support intentional killing, and a conviction for manslaughter was substituted, while Cherry’s appeal based on the alternative explanation of a low-level fall was rejected. The Court upheld Faulder’s appeal, accepting the possibility that the father’s explanation of an accidental fall from his arms was a possible explanation for the non-fatal impact injuries suffered by his seven-week-old son, so the conviction was unsafe.  
826 *Harris*, n. 799 above, at [152].  
827 *Cannings*, n. 84 above, at [22].
The impact of Harris

While *Harris* raised a number of significant issues in relation to the management and admissibility of expert evidence in alleged SBS cases, including how the courts might manage expert evidence generally, a particularly notable issue that *Harris* exposed was how the courts should deal with novel scientific evidence. The two issues in particular here were focussed around the Geddes “unified hypothesis”, and the “short fall” explanation for the triad that I discussed earlier.

Reflecting a general, albeit inconsistent, tendency to admit and consider opinion evidence based on novel science, the Court emphasised, per *Clarke*,\(^828\) that new scientific ideas and techniques should not be kept from the courts, even if, like Dr Geddes’ theories, they were based on a hypothesis rather than established knowledge. However the Court stressed that, as was clear in both *The Ikarian Reefer*\(^829\) and *Clarke* judgments, the expert had a duty to be explicit about the status or general acceptance of the scientific theory that they were advancing, to be open about any controversies around it, and to provide any necessary information that contradicts those ideas.

*R v Henderson – three further conjoined SBS appeals*

These themes of novel science admissibility, the management of complex evidence from paediatric experts, and whether an as yet unknown cause might explain SBS findings were revisited in 2010 when the CACD considered three further conjoined appeals (*R v Henderson, R v Butler, and R v Oyediran* (hereinafter *Henderson*)\(^831\)) that turned almost exclusively on complex and often conflicting expert evidence.

\(^{828}\) Clarke, n. 799 above, involving facial mapping by way of video superimposition.
\(^{829}\) The Ikarian Reefer, n. 793 above, at [80], specifically criterion 5.
\(^{830}\) Harris, n. 799 above, at [272], citing Wall J in Re AB (Child Abuse: Expert Witnesses) [1995] 1 FLR 181.
\(^{831}\) Henderson, n. 791 above.
In *Henderson*, the CACD now provided detailed guidance on case management - particularly the management of expert evidence, the admissibility of novel medical theories, and on the appropriate content of judicial summing up in such cases. As I noted in Chapter 2, it was the approach commended by the CACD in *Henderson*, particularly, under the provisions of the then Criminal Procedure Rules 2010, the potential utility of pre-trial meetings of experts and the production of a joint minute summarising areas of agreement and disagreement that the Scottish HCJ referred to positively in the *Hainey* appeal.

In relation to the selection of experts and admissibility of their evidence, the CACD in *Henderson* also endorsed the recommendation of Baroness Kennedy QC, in her report in relation to Sudden Unexpected Death in Infancy, concerning judicial assessment of an expert witness’ professional status and areas of expertise, and the need to clarify whether an expert’s views reflected mainstream opinion.

More broadly, the Court also recognised that medical science, like all areas of science, is constantly evolving, and that “…today’s orthodoxy may become tomorrow’s outdated learning”. *Henderson* makes clear that juries must be in no doubt as to whether there are uncertainties in the relevant medical or other forensic science on which experts are providing opinions, and that in the face of these uncertainties special caution was needed where such expert opinions were fundamental to the prosecution.

In considering the status of the triad in SBS cases, the Court noted that, per *Cannings*, “…even where on examination of all the evidence, every possible known cause has been

---


833 Baroness Kennedy, n. 131 above.

834 *Henderson*, n. 791 above, at [217], quoting Holdsworth, a previous successful SBS murder conviction appeal (R v Holdsworth [2008] EWCA Crim 971, at [57]).
excluded, the cause may still remain unknown”.

Given the possibility that in the future as yet unknown natural explanations for part or all of the triad’s elements may emerge, in “appropriate” cases (presumably involving the triad alone) the jury should be directed that “unless the evidence leads them to exclude any realistic possibility of an unknown cause they cannot convict”.

As subsequent cases demonstrate, the dicta in Henderson do not mean that the possibility of an unknown cause inevitably means that juries cannot convict in such cases. Rather, they must consider the possibility of an unknown cause, before reaching a verdict on all of the evidence before them.

The CACD in Henderson also specifically considered one of the core issues I raised in Chapter 5 – the claim that the outcome of legal fact-finding tribunals may be used to provide a validation for medical experts’ opinions on similar cases. In Henderson the court explicitly rejected that proposition, for in controversial and complex cases such as alleged SBS that turn almost exclusively on expert evidence, the court “feared that the medical profession may have looked to the courts to resolve medical controversy”; the court emphasised that “[n]o appellate jurisprudence could provide authority for a medical proposition”. The court emphasised that the legal outcome of each case turned on its own details, and thus such outcomes could not inform clinician’s judgments in other cases.

As we can see, in Harris and Henderson, the courts critically explored in some depth the SBS scientific controversies around the causation of the triad and its individual components, and recognised that on some issues, expert disagreements remained unresolved.
It followed from this that it was no longer acceptable for an expert to suggest that if known alternative natural causes of the triad were excluded, the infant must have been shaken; a more restrained conclusion was necessary. The possibility of an as yet unknown cause had to be acknowledged.

It is thus evident in Henderson that slippage in judicial acceptance of the traditional SBS diagnostic paradigm had begun, a change in perspective that reflects back on some of the core issues I raised in previous chapters, notably the foundations of the evidence base (Chapter 5), and the need for experts to match the certainties expressed in their opinions with the status of the science underlying them (Chapter 6).

One of the significant outcomes of the Harris and Henderson judgments was that, in light of the evident unresolved expert views, in 2011 the Crown Prosecution Service (CPS) in England provided new and specific guidance (still current) for prosecutions of cases of Non-Accidental Head Injury (an alternative broader term for SBS). This guidance suggests caution in the prosecution of such cases based on the presence of the triad alone, and indicates that decisions to prosecute should generally require other supporting evidence in addition to the triad of injuries.  

The impact of the CPS guideline on prosecution trends is unknown. While there is no empirical information available about secular trends related to prosecutions and conviction rates in alleged SBS cases in England or Scotland, there is some evidence to suggest that these scientific uncertainties may be impacting on prosecution decisions and conviction rates in the Southern Hemisphere.

842 See, for example, reports of falling SBS prosecution rates in New Zealand and Australia, Dr Amanda Stephens, Paediatrician, Sydney Children’s Hospital has noted that only half of apparent cases in New South Wales result in a prosecution, mostly founded on confession evidence; see
Section 4 - From problem description to problem solution

So far in this chapter I have provided the technical background to the forensic assessment of possible infant shaking injuries, set out the general legal framework against which judicial assessment of the reliability of expert evidence may be assessed, and then illustrated how the prosecuting authorities and courts have begun to react to emerging challenges to the traditional diagnostic paradigm that holds that the presence of the triad alone is strongly suggestive of an inflicted shaking injury.

In relation to the latter issue, Harris and Henderson are good examples of the positive efforts made by the legal system in attempting to understand and manage controversial and apparently evolving complex forensic evidence.

But of course, the courts cannot resolve disputed scientific issues. Only members of the relevant scientific community – in this case a variety of paediatric specialists and related experts - can resolve such technical controversies. And while a scientific community may not be able to reach a consensus, it should at least be possible to define areas of agreement, disagreement and knowledge gaps so that triers of fact can take account of these uncertainties in the judgments they must make, and where new evidence may fit in on appeal as science evolves. Such an approach of course assumes that the relevant scientific community recognises that significant scientific controversy exists.

In the last part of this chapter, in the light of the uncertain status of some of the SBS science, and the prosecuting authorities and judicial responses to alleged SBS criminal cases in England and Wales, I explore some further practical issues that arise for experts and for the courts that relate to some of the concerns about paediatric expert evidence I previously


raised, before briefly considering, in the light of the foregoing, whether SBS cases in Scotland should be dealt with differently.

**The truth, the whole truth…and SBS**

From the perspective of this thesis, and in particular in relation to the issues already highlighted in this chapter, a number of practical questions arise for both experts and the courts in relation to SBS cases.

The first question is a normative one that relates to the issues I discussed in Chapter 6: how should a paediatric expert acknowledge and take account of the controversy in relation to the forensic significance of the isolated triad when providing their opinion?

Reflecting back to Chapter 2, and Professor Black’s failure, in *Hainey*, to acknowledge contrary published evidence in relation to Harris lines that she considered to represent “bad science”, the courts make clear in both *Henderson* and *Kennedy* that experts have an explicit duty to disclose any scientific controversies or uncertainties underlying the opinion evidence they provide, and to explain the basis of their conclusions in the context of that science.

Despite institutional claims dismissing the scientific validity of the disputed evidence in SBS, this would seem to involve a review of all the possible causes for the clinical findings, and provision of a range of opinions based on these possibilities, taking account of the specific details of the instant case. So far, so apparently straightforward. But of course it is in the detail that the challenge for the expert emerges, for at some point it must surely become appropriate for an expert to choose to ignore what they consider putative explanations based on “junk” science. Perhaps a more challenging related question is whether experts should proactively raise the possibility of an “unknown cause” due to lack of available research, or wait to be asked this “Harris” question in court.

Thus, the causation controversies in alleged SBS particularly highlight the difficulties that paediatric experts may face in formulating their conclusions.
There is, however, a separate more general question here: if an expert’s opinion is to be founded on a reliable body of knowledge and expertise, in the context of SBS, where and with whom does that reside? Is it with neurologists, neurosurgeons, paediatric imaging specialists, child protection specialists, ophthalmologists, or even biomechanical engineers? All of these experts, and many others, may become involved in such cases. This takes me back to the potentially problematic issue of overlapping boundaries of expertise that I discussed in Chapter 1, and the particular role of the paediatric expert in bringing together other experts’ opinions to provide a composite opinion.

The courts seem to recognise that in the specific situation of SBS cases, such boundaries are particularly blurred. In Squier v GMC,844 an appeal by a paediatric neuropathologist, an SBS sceptic, against her erasure from the medical register, the court notably focused, amongst other issues, on whether or not she had provided opinion evidence outwith her area of expertise. Here the High Court held that:

[W]hen the triad fell to be considered, any specialist…who supported or doubted the triad as indicative of NAHI, would be bound to be expressing an opinion outside his specialism. There can be no proper criticism of a [specialist] for explaining why he supports or doubts the majority view and in doing so, expresses a view about symptoms or pathological findings outside his own discipline. It is neither improper nor professional misconduct for an expert in one specialism to do so. The boundary line between a proper explanation of support or doubt and trespassing impermissibly outside the expertise of the witness is imprecise and difficult to identify in any particular case.845

This suggests that the courts are prepared, in particular circumstances, to be flexible in their interpretation of expert boundaries, recognising not only areas of overlapping expertise, but also when particular disputed topics straddle a variety of knowledge domains where various experts have a legitimate claim to provide opinions. Such an approach would seem to support the role of the paediatrician in providing a summative opinion in such cases.

---

845 ibid, at [17].
As I draw this chapter to a close, I conclude by raising some specific issues for the courts that lurk within the somewhat tangled undergrowth of the SBS landscape. I raised three of these—the issue of differential admissibility standards of established and novel scientific evidence, the challenge for the court’s in managing expert claims based on rare but possible exculpatory explanations, and the impact of the emotive nature of paediatric cases on fact-finder decision-making—at various points earlier in this thesis. The fourth is an important novel final question that arises from the cumulative concerns this thesis has exposed - the implications for the management of SBS cases in the Scottish courts.

Some on-going challenges for the courts

Discordant admissibility thresholds for “established” and “novel” science?

[S]cientific knowledge is fluid, and consequently…rigid rules on the admissibility of a particular category of expert evidence may persist long after their ‘sell-by date’. As I noted above, the Supreme Court (SC) in Kennedy seemed to make an explicit distinction between the admissibility thresholds of “established” and “novel” scientific evidence, seemingly suggesting that the former should automatically be admitted, while implying that the latter should generally be considered suspect (and therefore more likely to be challenged and excluded).

This SC view reflects a feature evident across all common law jurisdictions - that the courts have generally tended to accept the trustworthiness, and therefore the admissibility, of many longstanding forms of expert evidence if it has been regularly admitted in the past, and that challenges to the reliability of such evidence may be resisted. This “Trojan horse of precedent” approach, as Edmond has describes it, also appeared to be endorsed by the English Law Commission in order to minimise the need for detailed and repeated judicial

---

848 Edmond, n. 765 above.
scrutiny when admissibility decisions about expert evidence may be required in similar
future cases.

Of course, as I argued earlier, the consequence of such a “grandfathering” approach may be
that, even if there is significant slippage in the reliability of the science underlying it, such
evidence is likely to be admitted without challenge long after its probative strength has been
eroded.

Thus I suggest that in the context of SBS, and in many other forensic areas, the courts
should be more circumspect in their passive acceptance of the established consensus views
on the causation of the triad.

But, of course, lawyers and judges are not in a position to resolve scientific controversies. It
is the experts who provide such evidence and it is surely up to them, both as professional
groups, and as individuals, to recognise when the science has moved on or is no longer so
secure. Science is not about finding truth, but finding falsehood, reflecting the Popperian
concept that knowledge develops by using hypotheses that expose (falsify) established
knowledge. During such transitions, the courts inevitably face problems in reconciling
discordant expert opinions on the same disputed facts.

The stark reality of the challenges the courts face in relation to disputed SBS evidence was
recently illustrated at an event hosted by the Royal Society of Edinburgh when two judicial
primers on expert evidence were launched. It had been noted that one of the priority topics
for judicial education was to have been a primer on expert evidence in SBS. However this

\[850\] Smit et al., (2018), n. 584 above. See also: Marika L. Henneberg 'Admissibility frameworks and
scientific evidence: controversies in relation to shaken baby syndrome / abusive head trauma' (2015) 4
(2) Br J Am Leg Studies 555.
\[851\] Black and Nic Daeid, n. 69 above.
has been shelved meantime, until, as Professor Dame Sue Black noted: “the medics sort themselves out and decide what the evidence is”.  

*Alternative explanations – of outliers and the ecological fallacy*

A challenge that arises both for the paediatric expert witness and the courts concerns the possibility that a rare but known cause may provide an alternative exculpatory explanation for a child’s injury or death. As I noted earlier, one of the controversial forensic issues in SBS is the suggestion that, rarely, minor trauma – perhaps a simple fall - might cause the triad. This was part of the argument behind the successful appeal of Lorraine Harris.

Such an outlier issue was also the basis of the defence case in *Hainey*, raising the rare (albeit highly unlikely) possibility that Declan had died naturally – “was he the one child each year of that age in Scotland who dies of SIDS?”

While such cases are rare, it would be wrong to use statistical rarity to suggest that such explanations should be automatically rejected as inadmissible – such an error representing the ecological fallacy. But it does raise the question of when such explanations achieve sufficient evidential weight to counter alternative explanations for forensic findings.

So, rarity alone is not a defensible reason to rule such evidence inadmissible, but it is clearly incumbent on the expert supporting such an explanation (or rejecting it) to provide a very detailed rationale for their conclusion, taking into account both the rarity of such events and the possibility of other explanations. Of course, whether rarity creates sufficient doubt to

---

852 Personal oral communication, Professor Dame Sue Black, Royal Society of Edinburgh, 22 November 2017.
853 See Plunkett, n. 738, and Chadwick et al, n. 739, above.
855 See, for example, Yehuda Neumark ‘What can ecological studies tell us about death?’ (2017) 6 Isr J Health Policy Res 52.
raise questions about whether a conviction would be safe or not is not for the expert to consider, but the court.

_Might emotion exaggerate the apparent reliability of the opinion?_

As I noted in relation to _Hainey_, cases involving dead or severely injured infants are highly emotive. A separate relevant issue here is that in cases of alleged SBS, admission of incriminating opinion evidence of moderate or uncertain reliability might have a disproportionally high impact on lay fact finders, where “[j]uries, fraught with emotion, [may] seek to cast blame on someone in the desperate attempt to obtain justice for the innocent child”. In the highly emotive context of such cases, the prejudicial effect of even modest expressions of confidence by an expert that shaking is a likely cause of the injuries may outweigh the true probative weight of such an opinion. And, like _Hainey_, what about the potential impact of adverse social feature or a defendant’s lifestyle on the interpretation of expert evidence of uncertain reliability? Should the admissibility standard for expert forensic evidence of uncertain reliability change depending on the presence of other unrelated inculpatory evidence?

Rejecting this transformative effect, Edmond and Roberts note that “the reliability of an expert’s opinion is unrelated to the strength of the case [overall] or the value of other evidence,” [thus] “admissibility decision-making should be independent of other evidence and considerations”. Thus, they claim that the reliability of expert evidence should be independent of other evidence in a case, and should stand or fall on its own merits.

In turn, I argue that, in the absence of other injuries, given the uncertainties about its cause, the presence of the triad can only be considered in any circumstance to provide modest rather than definitive support for a claim of inflicted injury.

---


857 Edmond and Roberts, n. 590 above, at 856.
**SBS prosecutions – a Scottish perspective**

I noted earlier that it was my recent experience that cases of alleged SBS are prosecuted in Scotland where the triad represents the *only* forensic finding. Here the prosecution usually relies on paediatric expert opinion evidence claiming that the triad *alone* is a very strong indication that an infant has been violently shaken, often with no defence expert evidence to the contrary. Many are associated with late guilty pleas to reduced charges.

Earlier in this chapter I discussed current scientific uncertainty in relation to the origins of the SBS triad, and suggested that the evident paediatric institutional failure to at least acknowledge such reasonable scientific debate is rooted in broader value-laden concerns that relate to safeguarding of children.

These and the other issues I have highlighted raise questions about whether criminal prosecutions related to alleged SBS should be pursued in Scotland if they are primarily based on incriminating expert opinion that continues to endorse the conventional view that the triad *alone* is a strong marker of an inflicted injury.

In the preceding chapters of this thesis, I have revealed significant potential frailties in paediatric expert evidence that raise questions about how such evidence should be managed by the experts themselves, and dealt with by legal actors and the courts. Further I have shown in this chapter that many of these concerning issues are relevant to criminal cases involving alleged SBS.

In the light of those concerns, I suggest that it is no longer appropriate for paediatric experts to opine that the presence of the triad, in isolation, is strongly indicative of a shaking injury. A much more cautious approach is required.

---

Note 712 above.
Given the scientific uncertainties and potential for subjective expert biases, and based on the evolving approaches of the prosecuting authorities and the English courts to such a scenario, I propose that criminal prosecutions of alleged SBS based on the presence of triad alone should not be pursued in the Scottish courts.

**Conclusion**

In this final substantive chapter I have moved from the perspective of the expert witness to that of the courts and examined one of the key challenges at that interface – the management of incriminating paediatric expert evidence in relation to alleged shaken baby cases, and the challenges for the courts in assessing the reliability of such evidence to ensure that, as far as possible, only sufficiently reliable expert evidence is admitted and appropriately presented.

I have shown that the evidence base underlying the interpretation of the clinical features in alleged SBS is in a state of relative flux, where on the one hand many experts continue to strongly support the institutional consensus view that the triad alone remains an extremely strong indicator that a baby has been shaken, while a respectable minority are now beginning to challenge that established paradigm, and where a variety of alternative hypotheses of varying and uncertain validity are also sometimes presented as exculpatory explanations for the clinical findings.

What the SBS controversy also raises is a more general issue - the generally differential application of reliability and admissibility standards applied by the courts depending on whether an expert’s opinion reflects conventional rather than novel views of the state of particular disputed scientific issue, even when, as in the case of latent fingerprint evidence, in reality there is much emerging scientific uncertainty.

---

See, for example, R v P [2010] EWCA Crim 2895, involving a judicial decision not to proceed with a prosecution of a “triad case” in the face of uncertainties about the ability to exclude innocent or unknown explanations.
Here I do not argue that this institutional tendency to be slow to recognise scientific reversals necessarily justifies a more explicitly restrictive approach by the courts in applying differential admissibility thresholds for opinions based on “established” as compared to novel science that the SC in *Kennedy* continues to support. However, I argue that that the courts need to ensure that experts explicitly justify the current evidential foundations of even well-established science in order to clarify its on-going status.

In this chapter I have demonstrated that many, if not all, the potential frailties in relation to paediatric expert evidence that I highlighted in the preceding chapters come together in alleged shaken baby cases. In both the criminal and family court settings, they represent arguably the most challenging cases that come before these courts. I have not sought to provide solutions to those challenges, but tried to expose some of the issues that must be considered in addressing them.

As the English (and perhaps Scottish) criminal courts increasingly seek to manage expert testimony, and demand that experts explicitly comply with their duties to the court, one of the increasingly clear key expectations is that the expert must clearly explain the basis for their opinion, the alternatives they have considered, and ensure that the certainty with which they express their opinion matches the reliability of the science underlying it.

As the contentious debate about expert evidence in cases of alleged SBS shows, expert paediatricians face particular challenges in meeting these expectations, given the evident uncertainties about the foundations of evidence base on which they must base their opinions.

In my concluding chapter I summarise the potential frailties intrinsic to paediatric expert evidence that I have exposed in the preceding chapters, and briefly consider the practical implications of the findings.
Chapter 8 – Conclusion: some inconvenient truths, and thoughts about future directions

Introduction

As Professor Sir Roy Meadow and Professor (soon to be Dame) Sue Black entered their respective witness boxes some 12 years apart and swore/ affirmed that they would “tell the truth, the whole truth, and nothing but the truth”, I suggest that we can be sure of one thing – neither consciously intended to provide flawed expert evidence that would result in a wrongful conviction. But as Clark and Hainey illustrate, as each stood in that intimidating liminal space, straddling the worlds of paediatrics or anthropology and the law, offering their opinions to juries tasked with deciding if a mother had committed what many would regard as the ultimate crime, even the most distinguished and experienced expert witness can get things wrong.

Both were judged to have made the same error – each had provided opinion evidence beyond the boundaries of their relevant areas of expertise. And of course, other errors were evident, for both of these renowned experts were unaware or chose to ignore published evidence that should at least have modified the evident certainty in their conclusions. They had been overly dogmatic, and the science they espoused seemed to have moved on.

In noting the above, I do not intend an ad hominem attack on those individuals, but to reflect back on the origins of this thesis: if such renowned experts could make such errors as they stepped into the law’s domain, what of the “ordinary” child abuse expert? And this led me to a broader question: could it be that such “child cases” pose particular problems?

Thus I set out to address a gap in our understanding of the reasons why, based on an apparent disproportionate prominence of wrongful convictions involving cases of alleged serious child abuse or murder, paediatric expert evidence seemed to be particularly problematic.

This thesis is the result of that inquiry.
It represents the first comprehensive and focused examination of potential weaknesses that may underlie core elements of paediatric expert forensic practice and associated opinion evidence in relation to alleged serious child abuse. While some of the issues that I consider here have been explored in relation to other areas of forensic practice, this thesis represents the first in-depth and integrated account of this topic from a paediatric forensic perspective.

What was the research problem and how did I address it?

The key question I sought to address was whether there are specific potential frailties related to the professional ecology and process of paediatric expert assessments in cases of suspicious injury or child deaths that may undermine the reliability of an expert paediatrician’s decision-making and judgment, and that need to be recognised and accounted for in order to ensure, or at least improve, the provision of reliable opinion evidence.

The work contained within the core of this thesis sought to reveal these issues. Given the dearth of relevant paediatric scholarship available in relation to most of the topics I identified, it was necessary to explore and apply knowledge available from other areas of forensic science or other relevant learning. My arguments were further informed by my personal perspective and insights as an experienced paediatric expert witness and clinician.

And, while this was essentially an inward-looking exercise, particularly focused around the workings of clinical paediatric forensic practice, where appropriate I also considered the wider child protection environment, including the potential impact of other agencies, and the influence their professional actors may have on the paediatric expert’s judgments.

The common thread that runs through the core of most of the preceding chapters is that they track the sequential process that the paediatric expert follows in assessing a case of alleged child abuse or murder.

In systematically examining these key elements, I sought potential threats to a paediatric expert’s objective opinion, and identified a number of issues that individually and
collectively may adversely influence the forensic interpretation and expert opinions of suspicious childhood injuries. Some of these, such as the emotional impact of possible violent crime, may be considered generic and relevant to all forms of forensic interpretation, although I argue that in the setting of alleged child abuse assessments such effects are likely to be enhanced. Other influential factors that I identify are more particular to the setting of paediatric expert forensic practice.

Thus, the complex professional landscape of safeguarding and child protection and the multiple simultaneous responsibilities that expert paediatricians must fulfil provide a challenging environment in which to disentangle and appreciate the expectations of the various agencies that all intersect around a child with suspicious injuries. Other notable features peculiar to such cases include the domestic environment in which most cases originate, and that case outcomes may depend almost exclusively on the expert paediatrician’s forensic interpretation.

One of the consequences of a study focused on “frailties” or problems of a professional system is that it might engender a negative view of the relevant practitioners and their competencies. This would be an unfortunate and unjustified effect. This inquiry and its conclusions do not set out to criticise or undervalue the commitment and professionalism of the many expert paediatricians and associated specialists who daily must respond to rising demands to assess suspicious injuries in children against a background of an increasingly risk-averse public and multiagency environment, and a blame-orientated professional regulatory attitude. Rather, it exposes issues that might otherwise be unrecognised, and in doing supports them in their work by hopefully reducing the possibility of personal misjudgment and error.
Key themes specific to the paediatric forensic practitioner

A number of key issues that might influence the paediatric expert’s judgment emerge from my analysis. While I have separated these out in the listing below, reflecting the sequence of an assessment process, it is important to acknowledge that such a distinction is somewhat artificial. In the practical reality of dealing with such cases they are interrelated, and their varying combined impacts coalesce and separate dynamically during different phases of the assessment.

The risk of unconscious expert bias

First, I highlighted the largely intuitive nature of most human reasoning, and examined how that relates to the heuristics of the medical diagnostic process, contrasting that with the different expectations of a forensic expert assessment. Drawing on empirical work outwith the paediatric sphere, I suggested that a number of separate influences – the culture and values intrinsic to paediatric professional life, societal and policy expectations, and particular case-specific contextual features - might individually and collectively undermine the objectivity of paediatric expert opinion evidence. In particular, I highlighted the ethical tensions at play here – to “do the right thing by the child”, and yet to be open with the legal actors and the courts about the uncertainties underlying such evaluations. I also argued that the particular emotional milieu in which such forensic evaluations are made might separately bias such judgments.

The relative weakness of the clinical evidence base on which to found expert opinions

Second, I critically examined the available clinical evidence base on which objective expert opinion evidence must be based. I questioned the core foundation on which the forensic interpretation of suspicious injury is set, and raised concerns about the scientific and epistemological basis of some important areas of paediatric expert testimony, given that
features putatively indicative of abuse are primarily validated by confessions or admissions of uncertain provenance, or by the outcomes of criminal trials, civil family court judgments and child protection hearings, rather than empirical science.

*The absence of an agreed reporting lexicon*

Third, I discussed the particular semantic difficulties faced by the forensic paediatrician in expressing appropriate degrees of certainty in the conclusions they reach. I argued that the absence of agreed discipline-specific terminology standards in relation to the reporting of expert conclusions matched to the certainty with which they are held is a significant deficiency that may result in miscommunication both between expert disciplines, and between experts and the courts.

*The challenge for the courts in resolving conflicting expert opinions*

Finally, I highlighted the challenges faced by the courts in reconciling evolving, and dissolving, scientific understanding with the reliability standards that should determine whether opinions based on such science are safe to be admitted in evidence to the courts. Here, illustrated by the polarised professional disputes around the causation of “shaken baby” cases, I argued that while the courts may have a relatively well-established and largely cautious approach in dealing with admissibility questions arising from expert testimony in relation to novel or emerging science, there is an evident tendency to continue to admit “established” scientific testimony without challenge, even when the science on which such opinions are based becomes less certain. Delay in at least questioning the status of such knowledge is often compounded, if not led, by conservative attitudes of the medical or other forensic institutions themselves, which are often slow to acknowledge slippage in the reliability status of hitherto established science.

860 Black and Nic Daeid, n. 69 above.
And of course we must recognise that the law can only react to the evidence put before it, and it is up to the various scientific communities to reach a consensus on any particular scientific controversy. In the face of scientific disputes, the law must await their resolution, and do the best it can in the meantime.

**Relevance and implications of this study**

One of the key problems that seems to underlie some of the issues that I have raised here in relation to paediatric expert evidence is that, with the exception of paediatric forensic pathology, even for paediatricians dealing almost exclusively with alleged child abuse, a true “forensic culture” remains relatively underdeveloped.

An understanding of the relevant law and its requirements has, until recently, been addressed in an unstructured fashion, and even now the focus is still primarily on paediatricians who wish to provide independent expert forensic assessments, rather than those involved in the routine but equally challenging “coalface” as expert clinicians in child protection services. This lack of a distinct “forensic paediatrics” specialty, which may otherwise interact and learn from other forensic disciplines, may partly explain the difficulties that have been seen.

The solution may be to develop such forensic skills and knowledge within the routine safeguarding/child protection NHS services, but with a truly cross-cultural perspective. This would provide a formal bridge between clinical child protection and the law, where a common understanding and rules of engagement could be developed and refined.

**Implications for practice**

Wrongful criminal convictions are the extreme example of some of the particular complexities and challenges presented by paediatric expert opinion evidence that I have exposed here. The issues are equally generalisable to the civil courts where the outcomes of

---

861 The RCPCH has recently launched a series of training events for paediatricians who currently or wish to act as expert witnesses in child protection: “Expert Witness in Child Protection: developing excellence”.
flawed fact-finding are arguably just as serious. The unjustified removal of a child from their natural family into care, often nowadays followed by irreversible adoption, and where the burden of proof is on the balance of probability, also demands that the relevant expert opinion evidence, and the clinical and scientific basis on which it is founded, must exhibit the highest standards of forensic judgment.

The issues I have identified in this thesis demand practical responses, and indeed, during the course of this study, as some of the issues have begun to be recognised, at least partial solutions have been proposed. A detailed overview is not appropriate here, but some brief examples illustrate some of these proposed responses, and the on-going challenges.

Minimising Bias

In order to minimise the potential impact of the cognitive biases and the other human factors on paediatric forensic interpretation that this thesis highlights, raising awareness and understanding of the potential adverse influences of cognitive contamination and associated threats to objectivity would seem a basic and desirable response. However, given that many of these effects are implicit, the potential benefits of simple raising awareness about such risks are likely to be modest at best. Thus organisational changes designed to passively reduce such effects are required. However, these are not straightforward in a paediatric setting.

Unlike its applicability in some other forensic environments, blinding to contextual information is unrealistic in a paediatric clinical environment, where by definition the work involves multi-agency working and information sharing. Similarly, while separating those

---

862 David, n. 5, and Skellern, n. 222 above.
with treating responsibilities from those tasked with providing a forensic role to minimise role conflict may be theoretically attractive, such a strategy is generally impractical in all except the largest UK Children’s Hospitals.

Skellern has suggested a number of particular strategies that may be usefully considered. For example, she suggests the paediatric forensic expert should explicitly discount social and other contextual factors in their assessment of the origin of an injury, instead restricting the expert’s role to interpreting the biomechanical aspects of any injury in the context of the child’s developmental stage, and comparing that with the forces implicit in any explanation provided.

In a similar vein, in considering the potential causes of an injury, by actively constructing a differential diagnostic list of all possible explanations, and systematically considering all of these, she suggests that the potential risks of tunnel vision may be reduced. Such a “cognitive forcing” approach, designed to activate “type 2” (analytical) reasoning, and balance unconscious heuristic diagnostic responses has been applied in other challenging clinical settings, and seems a useful concept.

Whether such strategies might work in practice is unknown.

**Formulating forensic opinions**

As I highlighted earlier, the variability with which experts express their opinions is also problematic. Returning to contextual bias, Skellern suggests that in formulating their reports, experts should openly acknowledge any contextual information that may have been absorbed

---

866 Skellern, n. 222 above.
into their forensic judgment. This may, at least, expose the risk of double counting of
evidence that might otherwise occur.

More broadly, as I noted in Chapter 6, there is a clearly a need for paediatric expert
witnesses to approach the forensic examination and its recording of findings in a common
way (already largely achieved), but to develop a common language to describe findings and
conclusions to triers of fact. Thus the paediatric forensic community would be wise to set up
clear and unambiguous standards for reporting the probability that their forensic findings
represent an abusive injury. A rating scale with exemplar case vignettes, together with a
standard reporting language might improve the expression of the likelihood of abuse in any
particular case.

In addition, the format of expert reports now mandated by the English criminal and civil
courts not only provides a useful structure but makes explicit the general duties of experts in
providing opinion evidence, and includes specific expectations that the expert provides a
detailed rationale for the final conclusions to which they come.

That would seem a useful approach that could be adopted in Scotland, where experts’
responses remain largely unstructured.

These brief responses to some of the problems I identified earlier simply serve to illustrate
the types of practical solutions that might be developed in the light of the issues I identified.

---

868 See Skellern and Donald, n. 525 above.
869 Lindberg et al., n. 54 above.
870 Although I acknowledge some modest recent progress on that front Crown Office & Procurator
Fiscal Service for Scotland (COPFS), ‘Guidance Booklet for Expert Witnesses. The Role of the Expert
Witness and Disclosure’, available from
uidance%20booklet%20for%20expert%20witnesses%20%20June%202015.pdf (accessed 18 May
2018).
Strengths and limitations of the research

In many ways, the thesis represents a scoping exercise; indeed, the focus of each substantive chapter would arguably merit an individual thesis. However, given the dearth of scholarship and associated literature focused on the particular issues I have identified, this research begins to fill a major gap in existing scholarship in relation to some key aspects of paediatric expert evidence, and a broad overview of the issues explored here is the appropriate first step.

Given the lack of existing scholarship focused specifically on paediatric forensic evidence, it was appropriate to apply relevant knowledge from published literature set in other forensic domains. However, while I have been careful to select what seemed appropriate relevant literature, my selection is arguably selective and subjective. Thus it is important that the issues I have raised here are subjected to wider professional scrutiny and challenge.

However, having had the opportunity to present and discuss this research at a number of national paediatric professional meetings, the interest and positive feedback I have received indicates that for many paediatricians providing expert child protection opinions, the issues that I raise here are both novel and relevant to their practice. Notably, the issues of heuristics and contextual bias in the context of forensic assessment provoked great interest and there was strong encouragement that systematic learning on these topics should be widely disseminated in the child protection medical community.

Recommendations for further research

Specialist child protection paediatrics is a relatively new sub-specialty compared to many other clinical and forensic specialist areas, and the training of the experts who work within it is still evolving. Much of the related research agenda to date has been concentrated around
the recognition of abuse, generally directed at the interpretation of specific suspicious
clinical findings. This largely continues to be the prominent research focus.

System issues have generally concentrated on the expert paediatrician’s role within the
safeguarding / child protection framework and multiagency working, although more recently
the RCPCH has been developing an expanded professional development curriculum that
recognises that child protection paediatricians need a better understanding of the civil and
criminal justice systems in the relevant UK jurisdictions, and how they should interact with
them.

From a more general perspective however, there is almost no empirical research from the
UK directly addressing any of the topics I consider in the bulk of this thesis, and such
research is also sparse and disjointed on the international scale.

Many of the questions that this thesis raises are currently unanswered. Thus, for example, we
need a better understanding of contextual bias in the context of paediatric expert forensic
assessment, and whether restricting task-irrelevant information could be an effective
cognitive debiasing tool in the setting of alleged abusive injury evaluation. Similarly, we
need to understand the operational impact of emotion on paediatricians’ judgments.

And of course, as I began to explore in Chapter 7, the potential frailties that underlie
paediatric expert evidence may extend far beyond challenges for the experts themselves. As
Clark and Hainey demonstrate, such wrongful convictions also reflect failings on the legal
side of the forensic medicine/law interface.

Thus, an examination of the particular issues that paediatric expert evidence poses for the
courts is likely to be of itself a fruitful endeavour. There is significant potential for cross-
cultural studies here, exploring, for example, whether the proxy of confessions as a claimed

Daniel M. Lindberg and others 'Research Priorities for a Multi-Center Child Abuse Pediatrics
Lesley Munshi 'Forensic science in the dock' (2005) 155 NLJ 1162.
validation of the “scientific” justification of the diagnosis of shaking injury in reality meets the reliability standard of the criminal courts.

**Closing comments**

This thesis has identified apparent inherent frailties in the provision of paediatric expert evidence in alleged child abuse, a highly important area of interaction between medicine and the law that has, at times, been the focus of much public and professional controversy. It is an area that has received very little focussed research.

This thesis begins to reduce that deficiency.

Adam C Forensic Evidence in Court: Evaluation and Scientific Opinion in (John Wiley & Sons, Ltd, Chichester 2016a) 315.

—'Scientific Opinion and the Law in Practice' in Forensic Evidence in Court (John Wiley & Sons 2016b).


Adamsbaum C and others 'Abusive Head Trauma: Judicial Admissions Highlight Violent and Repetitive Shaking' (2010) 126(3) Pediatrics 546.


Alfonso MP, Thompson JL and Standen VG 'Re-evaluating Harris lines--a comparison between Harris lines and enamel hypoplasia' (2005) 29(2) Coll Antropol 393.


Anon., 'NSPCC reports 58% rise in calls referred to social services' The Scotsman (http://rss.feedsportal.com/c/33846/f/610141/p/1/s/e39f51/0L0Scotsman0N0Cworried0Escots0Ecalls0Eto0Enspcc0Eup0E30A0Ein0E20A130E10E3285124/story01.htm, 13 Oct 2016) Health Feed, accessed 18 May 2018.


Barnes P 'Child Abuse—Nonaccidental Injury (NAI) and Abusive Head Trauma (AHT)—Medical Imaging: Issues and Controversies In the Era of Evidence-Based Medicine' (2017) 50(3) U Mich J L Reform 679.


Baroness Helena Kennedy 'Sudden Unexpected Death in Infancy: a multi-agency protocol for care and investigation' (The Royal College of Pathologists and The Royal College of Paediatrics and Child Health 2004).


Bearing Good Witness: Proposals for reforming the delivery of medical expert evidence in family law cases - A report by the Chief Medical Officer (Department of Health, 2006)

Beattie JO, 'Deciding about life-sustaining treatment for incompetent children: doctors, parents and the courts' (MML, School of Law, University of Glasgow 2010).


Berger CEH and others 'Evidence evaluation: A response to the court of appeal judgment in R v T' (2011) 51(2) Sci Jus 43.


Betts S, 'A critical analysis of medical opinion evidence in child homicide cases' (PhD, Faculty of Law, UNSW 2013).


Bilo RAC and others 'Using the table in the Swedish review on shaken baby syndrome will not help courts deliver justice' (2017) 106(7) Acta Paediatr 1043.


Bonati M, Jacqz-Aigrain E and Choonara I 'Licensed medicines, off-label use or evidence-based. Which is most important?' (2016) 102(1) Arch Dis Child 53.


Bragge P 'Asking good clinical research questions and choosing the right study design' (2010) 41, Supplement 1(0) Injury S3.


Brookman F and Nolan J 'The Dark Figure of Infanticide in England and Wales: Complexities of Diagnosis' (2006) 21(7) J Interpers Violence 869.


Caffey J 'Multiple fractures in the long bones of infants suffering from chronic subdural hematoma' (1946) 56 AJR 163.


*Care Crisis Review: options for change* (Family Rights Group, 2018)


Clarke A, 'Secondary traumatic stress in child welfare work' (Masters in Education, University of Lethbridge 2010).


Cowley LE and others 'Letter to Editor' (2018) Law, Prob & Ris mgy002.


Crowe S and others 'Grief and Burnout in the PICU' (2017) 139(5) Pediatrics e20164041.


Cunliffe E 'Weeping on cue: the socio-legal construction of motherhood in the Chamberlain case' (LMM edn University of British Columbia 2003).


David TJ 'Avoidable pitfalls when writing medical reports for court proceedings in cases of suspected child abuse' (2004) 89 Arch Dis Child 799.


Debelle GD and others 'Abusive head trauma and the triad: a critique on behalf of RCPCH of "Traumatic shaking: the role of the triad in medical investigations of suspected traumatic shaking” (2018) 103 Arch Dis Child 606.


Decety J and Fotopoulou A 'Why empathy has a beneficial impact on others in medicine: unifying theories' (2014) 8 Front Behav Neurosci 457.


Deng Y and others 'Gender Differences in Emotional Response: Inconsistency between Experience and Expressivity' (2016) 11(6) PLoS ONE e0158666.

Department of Health 'Guidelines for Responding to Child Abuse, Neglect and the Impact of Family and Domestic Violence' (Government of Western Australia 2004).

— Delivering high quality, effective, compassionate care: Developing the right people with the right skills and the right values (HMSO, 2015).


Dias MS and others 'Defining 'reasonable medical certainty' in court: What does it mean to medical experts in child abuse cases?'' (2015) 50 Child Abuse Negl 218.


Dror IE and Hampikian G 'Subjectivity and bias in forensic DNA mixture interpretation' (2011) 51(4) Sci Jus 204.


Dyer C 'Pathologist in Sally Clark case suspended from court work' (2005) 530 BMJ 1347.

—'Great Ormond Street and Baby P: was there a cover up?' (2011) 343 BMJ 286.


Dwyer O 'Doctor charged with misconduct over murder claim' (2004a) 328(7453) BMJ 1393.


— 'The 'science' of miscarriages of justice' (2014) 37(1) UNSWLJ 376.


Edmond G and Roberts A 'Procedural Fairness, the Criminal Trial and Forensic Science and Medicine' (2011a) 33 Syd L Rev 359.


Edmond G 'Advice for the courts? Sufficiently reliable assistance with forensic science and medicine (Part 2)' (2012a) 16(3) IJEP 263.

— 'Is reliability sufficient? The Law Commission and expert evidence in international and interdisciplinary perspective (Part 1)' (2012b) 16(1) IJEP 30.


— 'Legal versus non-legal approaches to forensic science evidence' (2016) (20) IJEP 3.


Edmond G and others 'Contextual bias and cross-contamination in the forensic sciences: the corrosive implications for investigations, plea bargains, trials and appeals' (2015) 14(1) Law, Prob & Ris 1.


Edmond G and others 'Thinking forensics: Cognitive science for forensic practitioners' (2017) 57(2) Sci Jus 144.


Elaad E, Ginton A and Ben-Shakhar G 'The effects of prior expectations and outcome knowledge on polygraph examiners' decisions' (1994) 7(4) J Behav Decis Making 279.


Eraut M and du Boulay B 'Developing the Attributes of Medical Professional Judgement and Competence' (University of Sussex 2000).


Evidence-Based Medicine Working Group 'Evidence-based medicine. A new approach to teaching the practice of medicine' (1992) 268(17) JAMA 2420.


Findley KA and others 'Shaken Baby Syndrome, Abusive Head Trauma, and Actual Innocence: Getting it Right' (2012) 12(2) Hous J Health L & Pol'y 209.


Foreman DM and Williams J 'Medical law and protection of children' (2008) 337 BMJ.


Frankish K and Evans JSBT, 'The duality of mind: An historical perspective' in Evans, Jonathan St B. T. and Keith Frankish (eds), In two minds: Dual processes and beyond (Oxford University Press 2009).
Frasier LD and others 'International issues in abusive head trauma' (2014) 44(Suppl 4) Pediatr Radiol S647.


General Medical Council 'Protecting children and young people: the responsibility of all doctors' (2012).


Gigerenzer G and Goldstein DG, 'Betting on one good reason: The Take The Best heuristic' in Gigerenzer, Gerd, Peter M. Todd and & the ABC Group (eds), Simple heuristics that make us smart (Oxford University Press 1999).


Gigerenzer G and Todd PM, 'Fast and frugal heuristics: The adaptive toolbox' in Gigerenzer, Gerd, Peter M. Todd and & the ABC Group (eds), Simple heuristics that make us smart (Oxford University Press 1999).


Gilbert R and others 'Burden and consequences of child maltreatment in high-income countries' (2009a) 373(9657) Lancet 68.


Godsey M *Blind Injustice: A Former Prosecutor Exposes the Psychology and Politics of Wrongful Conviction* in (University of California Press, Oakland 2017) 244.


Goldstein I and Shabat LW *Medical Trial Technique* in (Callaghan and Company, Chicago 1942).


Greenhalgh T 'How to read a paper: getting your bearings (deciding what the paper is about)' (1997) 315(7102) BMJ 243.

Greenhalgh T and Heath I 'Measuring quality in the therapeutic relationship' (The King’s Fund 2010).


Günther S, 'The recent history of child protection in Scotland and the impact of intra-familiar child abuse inquiries such as Orkney on today's child protection work' (Masters, Glasgow Caledonian University 2000).


Haack S 'Of Truth, in Science and in Law' (2008a) 73(3) Brook Law Rev 985.


Haack S Defending science - within reason: between scientism and cynicism in (Prometheus Books, Amherst 2007).

Haber RN and Haber L 'Experimental results of fingerprint comparison validity and reliability: A review and critical analysis' (2014) 54(5) Sci Jus 375.


Harris HA 'The growth of the long bones in childhood: With special reference to certain bony striations of the metaphysis and to the role of the vitamins' (1926) 38(6) Arch Intern Med 785.


Hatala R and Guyatt GH 'Evaluating the teaching of evidence-based medicine' (2002) 288(9) JAMA 1110.


Emily Henderson and Fred Seymour, *Expert Witnesses Under Examination in the New Zealand Criminal and Family Courts* (School of Psychology, University of Auckland, 2013).


Hettler J and Greenes DS 'Can the initial history predict whether a child with a head injury has been abused?' (2003) 111(3) Pediatrics 602.


Hojat M, 'Empathy and Gender: Are Men and Women Complementary or Opposite Sexes?' in Hojat, Mohammadreza (ed), Empathy in Health Professions Education and Patient Care (Springer International Publishing 2016).


Huber PW and Kay B 'Galileo's Revenge: Junk Science in the Courtroom' (British medical Association 2011).

Hughes J and others 'Biomechanical characteristics of head injuries from falls in children younger than 48 months' (2016) 101(4) Arch Dis Child 310.


Jackson G, Aitken CGG and Roberts P '4: Case assessment and interpretation of expert evidence: guidance for judges, lawyers, forensic scientists and expert witnesses',

Jackson G and Black S 'Use of data to inform expert evaluative opinion in the comparison of hand images: the importance of scars' (2014) 128(3) Int J Legal Med 555.


Jeffrey D and Downie R 'Empathy - can it be taught?' (2016) 46 J R Coll Physicians Edinb 107.

Jeffrey RC The Logic of Decision in (University of Chicago Press, Chicago, IL 1965).


Jenny C and others 'Analysis of missed cases of abusive head trauma' (1999) 281(7) JAMA 621.


Johnson SLJ, 'Legal and Forensic Medicine in Paediatrics' in Beran, Roy G. (ed), Legal and Forensic Medicine (Springer-Verlag 2013).


Kahneman D and Tversky A 'Prospect Theory: An Analysis of Decision under Risk' (1979) 47(2) Econometrica 263.
Kerasidou A and Horn R 'Making space for empathy: supporting doctors in the emotional labour of clinical care' (2016) 17(1) BMC Medical Ethics 8.


Khomami N, 'Parents cleared of abuse told they may not be reunited with their child' The Guardian (9 Oct 2015).


Kmietowicz Z 'Complaints against doctors in child protection work have increased fivefold' (2004) 328(7440) BMJ 601.


Lamine A and others 'Solving constrained optimization problems by solution-based decomposition search' (2016) 32(3) J Comb Optim 672.

Lane WG and others 'Racial differences in the evaluation of pediatric fractures for physical abuse' (2002) 288(13) JAMA 1603.


Laville S, 'Paediatricians shy away from dealing with abuse as backlash from parents grows' The Guardian (7 June 2004) Children.


Leventhal JM and Edwards GA 'Flawed theories to explain child physical abuse: What are the medical-legal consequences?' (2017) 318(14) JAMA 1317.


Levin AP and others 'Secondary traumatic stress in attorneys and their administrative support staff working with trauma-exposed clients' (2011) 199(12) J Nerv Ment Dis 946.


Loewenstein G and others 'Risk as feelings' (2001) 127(2) Psychol Bull 267.


Lynch M 'God's signature: DNA profiling, the new gold standard in forensic science' (2003) 27(2) Endeavour 93.


Lynøe N and others 'Is accepting circular reasoning in shaken baby studies bad science or misconduct?' (2017b) 106(9) Acta Paediatr 1445.

—'Easier to see the speck in your critical peers' eyes than the log in your own? Response to Debelle et al' (2018) 103(7) Arch Dis Child 714.


Maguire S and others 'What does the recent literature add to the identification and investigation of fractures in child abuse: an overview of review updates 2005-2013' (2013) 8(5) Evid Based Child Health 2044.


Manson M, 'Living in the Age of Outrage' *Observer* (20 June 2016).

Marion SB and others 'Lost proof of innocence: The impact of confessions on alibi witnesses' (2016) 40(1) Law Hum Behav 65.


Marquis R and others 'Discussion on how to implement a verbal scale in a forensic laboratory: Benefits, pitfalls and suggestions to avoid misunderstandings' (2016) 56(5) Sci Jus 364.


Martire KA and others 'On the likelihood of “encapsulating all uncertainty”' (2017) 57(1) Sci Jus 76.


McCartney M 'Statins for all?' (2012) 345 BMJ.

McIntosh N and others 'Guidelines for the ethical conduct of medical research involving children. Royal College of Paediatrics, Child Health: Ethics Advisory Committee' (2000) 82(2) Arch Dis Child 177.


Meehl PE Clinical versus statistical prediction: A theoretical analysis and a review of the evidence in (University of Minnesota Press, Minneapolis, MN 1954) 149.


Montgomery K *How Doctors Think: Clinical Judgment and the Practice of Medicine* in (Oxford University Press, 2006).


— 'The Supreme Court Screws Up the Science: There is No Abusive Head Trauma/Shaken Baby Syndrome 'Scientific' Controversy' (2016) 2014(5) Utah L Rev 1357.


Morrison Z, Australian Institute of Family Studies and Australian Centre for the Study of Sexual Assault 'Feeling heavy': vicarious trauma and other issues facing those who work in the sexual assault field in (Melbourne: Australian Institute of Family Studies, 2007).


Mullen C and others 'Perception problems of the verbal scale' (2014) 54(2) Sci Jus 154.


Munshi L 'Forensic science in the dock' (2005) 155 NLJ 1162.

Myers JEB 'What is the meaning of “reasonable medical certainty”? Your guess is as good as mine' (2015) 50 Child Abuse Negl 228.


Narang SK and others 'Acceptance of Shaken Baby Syndrome and Abusive Head Trauma as Medical Diagnoses' (2016) 177 J Pediatr 273.


Nardelli A, 'Turn of phrase: EU referendum result may come down to wording of question' The Guardian (3 June 2015).


Naughton AM and others 'Emotional, behavioral, and developmental features indicative of neglect or emotional abuse in preschool children: a systematic review' (2013) 167(8) JAMA Pediatr 769.

Naughton AM and others “Mind the Gap” - Observable features of parent-child interactions amongst neglected children aged 0 to 13 years' (2014) 99(Suppl 1) Arch Dis Child A72.


Newton BW and others 'Differences in medical students' empathy' (2000) 75(12) Acad Med 1215.


Nicholl R and others 'Can we have a permanent end to 'temporary brittle bone disease'?' (2012) 97(8) Arch Dis Child 762.


Nørdgaard A and others 'Scale of conclusions for the value of evidence' (2012) 11(1) Law, Prob & Ris 1.


Ommaya AK, Faas F and Yarnell P 'Whiplash injury and brain damage' (1968) 204 JAMA 75.


Paediatrics & Child Health Division *Genital examinations in girls and young women - a clinical practice guideline* in Royal Australasian College of Physicians, August 2009 (ed) (Royal Australasian College of Physicians, Sydney 2009).


Papageorgopoulou C and others 'Harris lines revisited: prevalence, comorbidities, and possible etiologies' (2011) 23(3) Am J Hum Biol 381.


Piteau SJ and others 'Clinical and radiographic characteristics associated with abusive and nonabusive head trauma: a systematic review' (2012) 130(2) Pediatrics 315.


Probst B, 'Why This Book Is Needed' in Probst, Barbara (ed), Critical Thinking in Clinical Assessment and Diagnosis (Springer International Publishing 2015).


The radiological investigation of suspected physical abuse in children (The Royal College of Radiologists, 2017).


Raoof N and others 'Retinal haemorrhage in infants with pertussis' (2017) 102(12) Arch Dis Child 1158.


RCPCH/ Family Justice Council, Paediatricians as expert witnesses in the family courts in England and Wales: Standards, competencies and expectations ()


Rivard JR, 'Confirmation bias in witness interviewing: Can interviewers ignore their preconceptions?' (PhD, Florida International University 2014).


Robertson BW, Vignaux GA and Berger CEH 'Extending the Confusion About Bayes' (2011) 74(3) Mod Law Rev 444.

Robins J 'Rough justice' (2016) 166(7701) NLJ 22.


Royal College of Paediatrics and Child Health 'Intercollegiate document - Safeguarding Children and Young people: roles and competences for health care staff' (2010).


Rudolf MC and others 'A search for the evidence supporting community paediatric practice' (1999) 80(3) Arch Dis Child 257.

Ruperto N and others 'Extrapolation or controlled trials in paediatrics: the current dilemma' (2017) 102(10) Arch Dis Child 949.

Russell J and Hollander N 'The Disappearing Trial' (2017) 8(3) NJECL 309.


Safeguarding Children Toolkit for General Practice (RCGP/NSPCC, 2014).


Saks MJ and Neufeld SL 'Convergent evolution in law and science: the structure of decision-making under uncertainty' (2011) 10(2) Law, Prob & Ris 133.


Saunders D and others 'Throwing the baby out with the bath water — response to the Swedish Agency for Health Technology Assessment and Assessment of Social Services (SBU) report on traumatic shaking' (2017) 47(11) Pediatr Radiol 1386.


Schiff GD and others 'Diagnostic error in medicine: analysis of 583 physician-reported errors' (2009) 169(20) Arch Intern Med 1881.


Schott E and others, 'Relationship of the Personality Traits of Early Career Pediatricians and Career Satisfaction and Burnout' (Pediatric Academic Societies Annual Meeting Baltimore Apr 30, 2016).


Shapiro J, 'The Paradox of Teaching Empathy in Medical Education' in Decety, Jean (ed), Empathy: From bench to bedside (Social Neuroscience Series, MIT Press 2011).


— Reason in Human Affairs in (Stanford University Press, Stanford, CA 1983) 128.


Singer T and Bolz M (eds), Compassion. Bridging Practice and Science (Max Planck Institute 2013).


Singer T and others 'Empathy for Pain Involves the Affective but not Sensory Components of Pain' (2004) 303(5661) Science 1157.


Smith R, 'Doctors should report child abuse suspicions "on a hunch"' The Telegraph (10 July 2012).

Smith SM, Hanson R and Noble S 'Parents of battered babies: a controlled study' (1973) 4(5889) BMJ 388.


Spence D 'Evidence based medicine is broken' (2014) 348 BMJ g22.


Tardieu A 'Etude médico-légale sur les services et mauvais traitements exercés sur les enfants' (1860) 13 Ann Hyg Pub Med Leg 361.

Taylor F and others 'Statins for the primary prevention of cardiovascular disease' (2013) 1 Cochrane Database Syst Rev CD004816.


Tone EB and Tully EC 'Empathy as a "risky strength": a multilevel examination of empathy and risk for internalizing disorders' (2014) 26(4 Pt 2) Dev Psychopathol 1547.

Traumatic shaking – The role of the triad in medical investigations of suspected traumatic shaking: A systematic review (Swedish Agency for Health Technology Assessment and Assessment of Social Services (SBU), 2016255E).


*Vicarious trauma: the consequences of working with abuse* (NSPCC, 2013).


Ward MG and Bennett S 'Studying child abuse and neglect in Canada: we are just at the beginning' (2003) 169(9) CMAJ 919.


Wardrop M, 'Baby P 'failed by incompetent staff from every agency" *The Telegraph* (26 Oct 2010).


Zylke JW, Rivara FP and Bauchner H 'Challenges to excellence in child health research: Call for papers' (2012) 308(10) JAMA 1040.