THE UNIVERSITY *of* EDINBURGH

Towards a Data-Driven Personalised Approach to Gestational Diabetes Care

THE UNIVERSITY *of* EDINBURGH

Jasmine R. Kirkwood, BSc (Hons)

Thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy,
The University of Edinburgh
2025

Declaration

I, Jasmine R. Kirkwood, declare that this thesis has been composed by myself and has not been submitted for any other degree or professional qualification. I can confirm that all work is my own, except where co-authors have contributed to publications that have been included in this thesis. I have contributed considerably to the publications as indicated by my first authorship and detailed throughout the thesis.

Acknowledgments

There are not many moments in life when we take the time to reflect and express gratitude to those around us, and it is not only a privilege to have had the opportunity to undertake a PhD, but also to thank those who supported me throughout.

I would like to thank my supervisors, Professor Rebecca Reynolds, Dr Debbie Wake, Dr Robert Lindsay and Dr Areti Manataki, for your continuous support and guidance. I would also like to thank Dr Jane Dixon, Dr Marryat Stevens, Dr Natasha Galloway, West of Scotland Safehaven, and the study participants, without whom none of this research would have been possible. Additionally, I am would like to express my gratitude to my funders Medical Research Scotland and MyWay Digital Health.

I would also like to thank my thesis panel chair, Dr Roland Stimson, for his constructive feedback and support throughout my research, and Dr Sara White and Professor Kathrin Cresswell for agreeing to conduct my final examination.

I am most grateful to my friends both within and outside academia, whom I have made while in Edinburgh. A special thank you, to Bohee, Druga and Thulani, thank you for all the fun, support, procrastination and writing retreats. You were my supporting pillars throughout my PhD. To my sister, Annie, who has always and will always be there for me and me for you. And to Matt, thank you for being by my side.

Most importantly, I would like to thank my mum, if it were not for her love, encouragement, and persistence in taking me to all my extra tutoring and extracurricular activities as a child, I do not believe I would be here writing acknowledgements in a thesis today. She is simply an extraordinary, strong, resilient person, and I do not know how you managed everything. Thank you is not enough.

Lay Summary

Gestational diabetes mellitus (GDM) is a common pregnancy complication characterised by high blood sugar (glucose) that develops during pregnancy and usually resolves after childbirth. However, individuals who experience GDM are at a greater risk of facing long-term health issues later in life, such as type 2 diabetes and cardiovascular disease. This makes them an important population to monitor for future health problems. GDM can also lead to complications during pregnancy, including delivering a larger baby; however, these risks can be mitigated by effectively managing blood sugar levels. Blood sugar is monitored during pregnancy through finger-prick tests using a glucometer, and the results are reviewed during hospital appointments that are in addition to standard pregnancy care. To maintain blood sugar levels, dietary adjustments and light exercise are advised, and in some cases, medication may be necessary to achieve the target blood glucose range.

Digital tools like mobile apps and web-based dashboards for GDM have allowed for remote monitoring and decreased the need for in-person visits. Research has shown that using digital tools for GDM care produces outcomes for both mothers and babies that are comparable to traditional GDM care. This indicates that using digital tools is safe and effective in alleviating the need for in-person appointments and minimising the challenges associated with managing GDM. Additionally, they have not been leveraged as a means to predict medication risks, which could enhance care efficiency and optimise the allocation of staff and hospital resources.

This thesis proposes a more personalised approach to managing GDM. To achieve this, I first focus on developing a medication risk prediction model, and then the design of a digital tool, 'MyGDM'.

Firstly, I aimed to explore existing machine learning models for predicting medication in GDM published by other researchers. To do this, I conducted a comprehensive 'scoping review' of the published literature. This review

revealed that the majority of models used an algorithm known as logistic regression and incorporated common clinical variables, including age and body mass index (BMI).

From a dataset of pregnancies and births in Greater Glasgow and Clyde that received care between 1st April 2022 to 31st December 2023, I identified pregnancies that had been complicated by GDM. I then analysed these pregnancies and found that those pregnancies that were complicated by GDM, particularly those that needed medication, had higher age, BMI. Pregnancies not complicated by GDM experienced lower rates of Caesarean births, had fewer large babies and babies admitted to the neonatal intensive care unit than those that did have GDM.

Using these findings from the results and the pregnancy data in Glasgow, I created a predictive model to assess the risk of insulin for a GDM-complicated pregnancy. The best model was logistic regression using health details and blood test results, including age, BMI, ethnicity, GDM diagnostic test value and the gestational week of the GDM diagnostic test.

Finally, I designed 'MyGDM', an app and dashboard to assist in managing GDM. To ensure the design reflected the needs of end-users, I consulted with healthcare professionals, women with GDM, and researchers. The final design received positive feedback from users.

In summary, GDM care could be moved to a more personalised approach that utilises risk-prediction models and digital tools. There is an opportunity to further develop the concepts discussed in this thesis into a clinically viable model of care for GDM.

Table of Contents

List of Tables

List of Appendices

List of Figures

Abbreviations

ACCORD	Academic and Clinical Central Office for Research and Development
ADA	American Diabetes Association
AI	Artificial Intelligence
ANOVA	One-Way Analysis of Variance
ASIPD	Australasian Diabetes in Pregnancy Society
AUROC	Area Under the Receiver Operating Characteristic Curve
BMI	Body Mass Index
CART	Classification and Regression Tree
CDSS	Clinical Decision Support System
CGM	Continues Glucose Monitors
CI	Confidence Interval
CRediT	Contributor Roles Taxonomy
EHR	Electronic Health Records
GCT	Glucose Challenge Test
GDM	Gestational Diabetes Mellitus
HbA1c	glycosylated haemoglobin
HCPs	Healthcare Professionals
IASPSG	International Association of Diabetes and Pregnancy Study Groups
ICD-10	International Classification of Diseases version 10
IQR	Interquartile Range
JBI	Joanna Briggs Institute
LASSO	Least Absolute Shrinkage and Selection Operator
LGA	Large for Gestational Age
MCAR	Missing Completely at Random
NHS	National Health Service
NICE	National Institutes of Health and Care Excellence
NICU	Neonatal Intensive Care Unit

NIH	National Institutes of Health
NPV	Negative Predicative Value
OGTT	Oral Glucose Tolerance Test
PCC	Population, Concept and Context
PHS	Public Health Scotland
PPV	Positive Predictive Value
PRIMISA-ScR	Preferred Reporting Items for Systematic reviews and Meta-Analyses – Scoping Review
PROBAST	Prediction model Risk Of Bias Assessment Tool
Q-Q	Quantile-Quantile
RCT	Randomised Control Trial
ROC	Receiver Operator Curve
RR	Relative Risk
SD	Standard Deviation
SIGN	Scottish Intercollegiate Guidelines Network
SIMD	Scottish Index of Multiple Deprivation
SMBG	Self-Monitoring Blood Glucose
SMOTE	Synthetic Minority Oversampling Technique
SMR02	Scottish Morbidity Record - Maternity interaction/stay
TRIPOD + AI	Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis Or Diagnosis with Artificial Intelligence
WHO	World Health Organisation

Abstract

Introduction

Gestational diabetes mellitus (GDM) is high blood glucose that is first recognised during pregnancy and is one of the most common pregnancy complications. GDM is managed through lifestyle modification and may require pharmacological therapy to normalise blood glucose levels. Self-monitoring blood glucose readings are reviewed at clinical appointments, where treatment may be escalated. Current care in Scotland is time-consuming and does not utilise the potential of risk-stratification and digital tools. This thesis proposes an alternative method of care that does so by risk-stratifying women with GDM based on their need for insulin and remotely monitoring them through a digital tool.

Methods & Results

This is an interdisciplinary mixed-methods thesis, the quantitative methods focusing on the development of a need for an insulin treatment prediction model, and the qualitative methods focusing on the design of a digital tool to support both women and clinicians in the management of GDM. To aid with readability, the results chapters' methods and results are reported together.

In the first result chapter I report the findings of a scoping review of 17 studies describing 44 machine learning models to predict the need for pharmacological therapy in GDM from four electronic databases between 1st July 2007 and 31st August 2024.

All published literature were binary classifiers with 61.4% (27/44) of models predicting need for any pharmacological therapy and 38.6% (17/44) predicting need for insulin. Models had a median area under the receiver operator curve (AUROC) of 0.75. Common clinical variables were found to be predictors, such as history of GDM, gestational week at GDM diagnosis, pregestational body mass index (BMI), maternal age, HbA1c, fasting – and 1-hr-glucose from

a 75g oral glucose tolerance test (OGTT), with logistic regression being a popular algorithm. There was a lack of external validation and clinical implementation.

In the following two results chapters, I present my quantitative analyses. From a database of 30,666 pregnancy episodes in Greater Glasgow and Clyde, I selected 10,694 pregnancy episodes that had a booking date and received care between 1st April 2022 to 31st December 2023. The selected data were cleaned, and a cohort of singleton pregnancies that were complicated by GDM (10.4%, 1,109) was identified alongside a complementary non-GDM cohort (89.6%, 9,585). The maternal characteristics and pregnancy outcomes were described.

In the descriptive analysis of 10,694 singleton pregnancies in Glasgow, the rate of GDM was high (10.4%, 1,109) despite only screening women with a BMI ≥ 35kg/m^2. Women with pregnancies complicated by GDM were older (GDM: mean[SD] 32 [5.3], non-GDM: 31 [5.4] years, p<0.001), had a higher BMI (32 [7.4], 27 [5.8] kg/m^2, p<0.001), and more likely to live in the most deprived areas (Scottish Index of Multiple Deprivation quantile 1, 45.4%(504), 39.5% (3,789), compared to those without GDM. Among women with GDM, those managing it through diet were younger than those who required pharmacological treatment (Diet: 32 [5.4], Metformin: 33 [4.8], Insulin: 33 [5.3] years, p = 0.001). Women requiring insulin had a higher BMI (32 [7.2], 32 [7.6], 35 [7.5]kg/m^2, p<0.001). HbA1c at booking was higher in the insulin-treated GDM (35 [3.3] mmol/L, 35 [3.1] mmol/L, 36 [3.8] mmol/L, p=0.038) as were fasting OGTT results (5.2 [0.5] , 5.3 [0.46] , 5.6[0.57] mmol/L, p<0.001) which was taken earlier (26 [6], 24 [5.6] ,23 [5.6] gestational weeks, p<0.001) . Caesarean birth rates were higher in GDM pregnancies (53.5%(593) vs. 40.0%(3,834), p < 0.001), with insulin-treated GDM showing increased elective Caesarean rates (29.8%(239), 35.9%(79), 41.4%(36), p=0.033). Large for gestational age (LGA) was more common in GDM-complicated pregnancies (15.0%(166), 9.0%(841), p<0.001). Neonatal unit admissions were higher in GDM pregnancies compared to non-GDM (41.2%(158), 9.4%(899), p<0.001).

Although admissions were more frequent in GDM managed with insulin, the difference was not statistically significant (13.8%(111), 14.1%(31), 18.4%(16), p=0.513).

The penultimate results chapter uses the results of the scoping review and statistical analysis to direct the development of machine learning models for predicting insulin for GDM. Two algorithms (logistic regression and Classification and Regression Tree (CART)) using univariate feature selection and least absolute shrinkage and selection operation (LASSO) were compared. Synthetic Minority Oversampling Technique (SMOTE) was used to address the unbalanced data. Using a complete case analysis, 996 GDM complicated pregnancy episodes were included in the dataset which split into 70% for training and 30% for testing. The training performance was assessed using 10-fold cross-validation, and the final model performance was validated on the unseen test data.

Both models generalised well, and overall were sensitive but not specific. The logistic regression had a mean 10-fold cross validation AUROC in train data of 0.79 [0.07] and AUROC of 0.71 on the unseen data. The CART model had mean 10-fold cross validation AUROC in train data of 0.78 [0.02] and AUROC of 0.71 on the unseen data.

In the final results Chapter, a user-centred digital tool, 'MyGDM', was designed from ethnographic observation and 11 semi-structured interviews with end-users (6 healthcare professionals and 5 women with GDM). The initial design was evaluated in feedback sessions with 31 participants (17 healthcare professionals, 14 researchers) and 13 questionnaires with women with GDM.

MyGDM has a clinical dashboard linked to a patient-facing app, aiming to enhance clinical workflows, identify off-target blood glucose levels and provide GDM-specific information. The initial design drew upon existing literature and insights from 11 semi-structured interviews conducted with HCPs and women with GDM. In a survey of 13 women with GDM, every

participant reported that the tool would fit well into their lifestyles and aid in managing their GDM. Educational resources, along with the 'request a call' feature, were particularly well received, with 61.5% (8 out of 13) and 69.2% (9 out of 13) indicating they were very likely or likely to utilise these options, respectively. End-user evaluations of the interactive design were favourable, confirming that it effectively met their needs.

Conclusion

GDM care could be personalised through risk-stratification and digital tools. There is an opportunity to translate the key concepts explored in this thesis into a clinically implementable model of care for GDM.

CHAPTER 1

Introduction and Background

1.1 INTRODUCTION

This PhD thesis presents a data-driven approach to gestational diabetes mellitus (GDM) through the development of a user-centred patient-facing app and clinical dashboard, which is supported by a pharmacological therapy prediction model that provides clinical decision support.

In a mixed-methods approach, the PhD has two parts: part one focuses on developing a machine learning predictive model, and part two focuses on designing a digital tool.

I use the term woman or women throughout the thesis to refer to women and birthing people who are pregnant or who have recently given birth. It refers to people who share the protected characteristic of pregnancy and maternity.

1.2 BACKGROUND

1.2.1 Gestational diabetes mellitus (GDM)

One of the most common pregnancy complications (Hivert et al., 2024), GDM, is hyperglycaemia that is first recognised during pregnancy (World Health Organisation (WHO) & International Diabetes Federation (IDF), 2021, 13 April). Depending on population, location and diagnostic criteria, it is currently estimated to affect approximately 14% of pregnancies globally (Sweeting et

al., 2024; Wang et al., 2022) and up to 30% in some populations (White et al., 2023). The prevalence of GDM is expected to increase as we see a general trend in the rise of obesity among women of reproductive age (Ferrara, 2007; Johns et al., 2018) and the increased age of those becoming pregnant (Johns et al., 2018), both of which are risk factors for GDM.

1.2.1.1 Complications associated with GDM

GDM can cause short-term complications during pregnancy, including adverse maternal and neonatal outcomes, and long-term health complications for both the mother and offspring.

During pregnancy, there is an increased risk of pre-eclampsia (Bryson, 2003), infant born large for gestational age (LGA), macrosomia, pre-term birth, Caesarean birth, respiratory distress syndrome, neonatal jaundice and admission to the neonatal intensive care unit (NICU) (Ye et al., 2022).

Subsequent pregnancies are also more likely to be affected by GDM (Buchanan et al., 2012). Most noteworthy, those affected with GDM are ten times more likely to go on to develop type 2 diabetes compared to those with a normoglycaemic pregnancy (Vounzoulaki et al., 2020) and have a two-fold increase in developing early cardiovascular disease (Daly et al., 2018). In addition, the offspring have an increased risk of childhood and adolescent obesity (Buchanan et al., 2012).

1.2.1.2 Diagnosis

GDM is more pronounced in lower socio-economic groups and is positively associated with body mass index (BMI), age, parity, and macrosomia (Collier et al., 2017; Vounzoulaki et al., 2024). Despite well-known risk factors and established long-term complications associated with GDM, there is no consensus on screening and diagnosis. Historically, some guidelines have tried to reduce the number of GDM diagnoses to manage healthcare professionals' (HCPs) workload (Simmons et al., 2024).

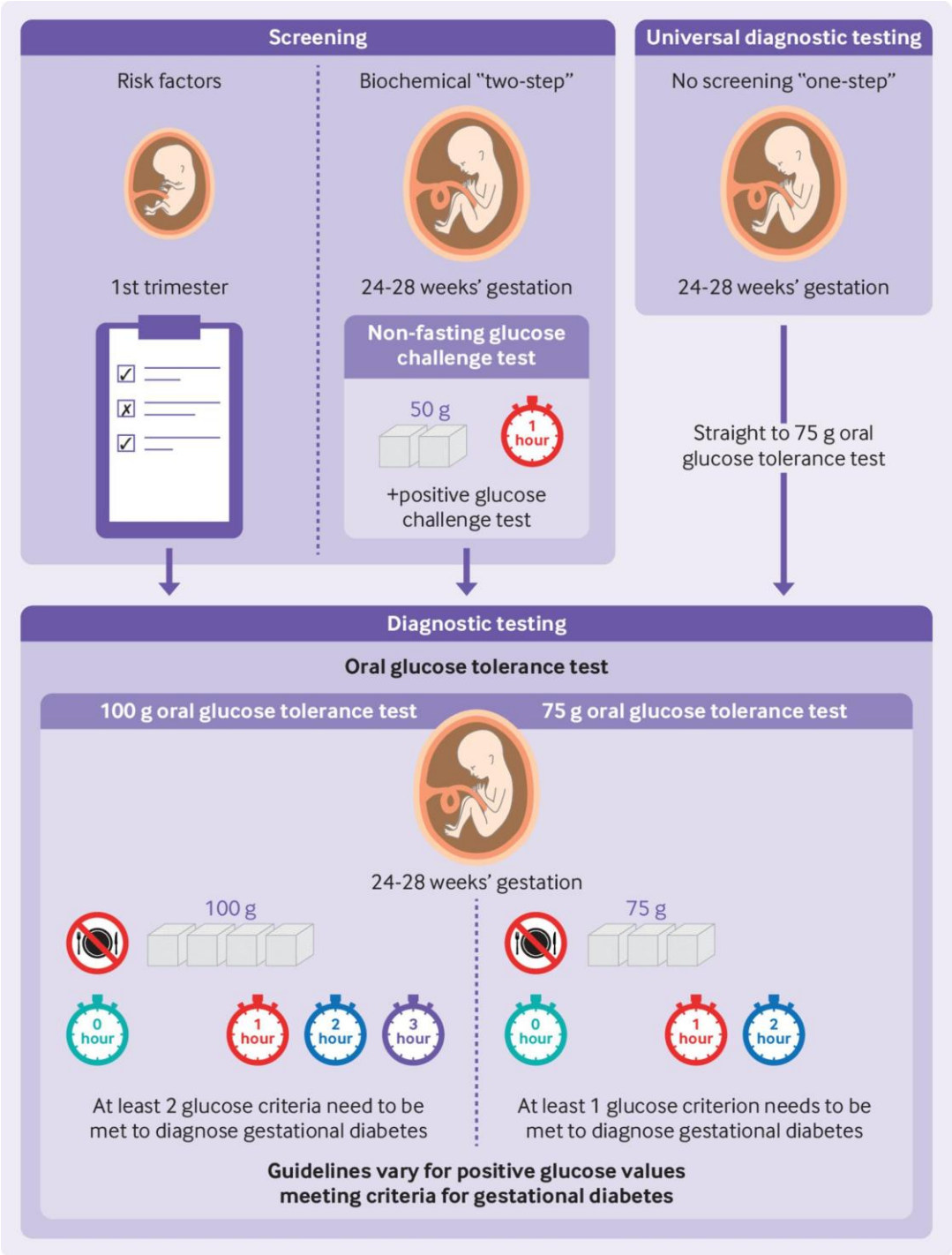


Figure 1.1 *GDM diagnostic methods. Reproduced from 'Screening and diagnosis of gestational diabetes', White S L, Ayman G, Bakhai C, Hillier T A, Magee L A., 381, e071920, 2025, with permission from BMJ Publishing Group Ltd.*

There are three main approaches to GDM screening and diagnosis, shown in **Figure 1.1** , 1) risk factor-based screening in the first trimester 2) biochemical "two-step" screening with a 50g 1-hour glucose challenge test (GCT) at 24-28

weeks. Both 1) and 2) are then followed by an oral glucose tolerance test (OGTT). Or 3) universal screening, which goes straight to the OGTT diagnostic test at 24-28 weeks gestation (White et al., 2023). GDM is diagnosed using an OGTT, of which there are two common types: 75g-OGTT and 100g-OGTT. A blood sample is taken after an overnight fast and then 1-hour and 2-hours after a glucose load has been taken and additionally 3-hours for the 100g-OGTT. GDM is then diagnosed if one blood glucose reading is over a threshold in a 75g-OGTT or if two are over a threshold in a 100g-OGTT. The thresholds vary depending on the guidelines used.

The UK National Screening Committee does not recommend universal screening for GDM (GOV.UK, 2021). In Scotland, GDM is diagnosed through risk-based screening with a 75g-OGTT at 24-28 weeks gestation and reviewing postload glucose levels against the International Association of Diabetes and Pregnancy Study Group's (IADPSG) thresholds (Scottish Intercollegiate Guidelines Network, 2024).

Risk factors for GDM screening in Scotland includes: a BMI ≥ 30 kg/m^2 (some locations this is ≥ 35 kg/m^2), previous macrosomic infant, previous GDM, a family history of diabetes in a first-degree relative, or family origin of South Asia, Middle Eastern or Black African/Caribbean (Scottish Intercollegiate Guidelines Network, 2024). It is also recommended that those aged 35-40 years with a history of polycystic ovary syndrome and those over 40 years should also be screened.

Those with a risk factor are then invited to take a 75g-OGTT and are diagnosed with GDM if they exceed any of the following thresholds: fasting plasma glucose ≥ 5.3 mmol/L, one-hour blood glucose ≥ 10.6 mmol/L or a two-hours blood glucose ≥ 9.0 mmol/L (Scottish Intercollegiate Guidelines Network, 2024). These thresholds have recently increased in 2024, from fasting glucose ≥ 5.1 mmol/L or a two-hours glucose ≥ 8.5 mmol/L (Scottish Intercollegiate Guidelines Network, 2010). Furthermore, to detect overt diabetes in those with GDM risk factors, a glycosylated haemoglobin (HbA1c) test is

recommended to be taken in the first trimester, and overt diabetes is diagnosed if HbA1c ≥ 48 mmol/L.

1.2.1.3 Care and treatment during and after GDM

Following diagnosis, blood glucose needs to be monitored for the remainder of the pregnancy. For GDM, self-monitoring blood glucose (SMBG) is done through finger-prick testing and a glucometer. The SMBG readings are recorded either manually in a paper logbook or electronically with a Bluetooth-enabled meter and portal, such as Glooko (Glooko Inc, United States) (Mayne et al., 2025). SMBG readings need to be taken four times a day, once fasting, and then at each meal, either pre- or post-prandial, depending on the hospital Trust. The Scottish Intercollegiate Guidelines Network (SIGN) recommends SMBGs fasting < 5.5 mmol/L, one-hour postprandial < 8 mmol/L and two-hours postprandial < 7 mmol/L (Scottish Intercollegiate Guidelines Network, 2024), however timings of reading and targets may differ between Trusts, (Mayne et al., 2025). These targets are set to reduce the maternal and neonatal adverse risks of LGA, Caesarean birth, neonatal hypoglycaemia and pre-eclampsia (Scottish Intercollegiate Guidelines Network, 2024). The SMBGs are typically reviewed fortnightly by an HCP at an antenatal-diabetes clinic, which may be in-person or via a telephone, depending on the clinic's capabilities. This is to ensure that SMBG values are within target and therefore current treatment is correct. In addition, fetal growth is monitored through ultrasounds (Venkatesh et al., 2024). These additional appointments can place considerable stress on a woman during their pregnancy (Fraser et al., 2023; Draffin et al., 2016; Nazarpour et al., 2024); meanwhile, an appointment typically takes over two hours, excluding travel time, making it time-consuming for a woman with GDM (Alqudah et al., 2019).

The first line of treatment for GDM is lifestyle modification through diet and exercise. If this is not sufficient to meet SMBG targets, then pharmacological intervention is needed. Education on diet and lifestyle change is provided through group education sessions upon diagnosis of GDM; these include dietary strategy, carbohydrate awareness, and recommended exercise for 150

minutes per week (Scottish Intercollegiate Guidelines Network, 2024). Consequently, if SMBG targets are not met with lifestyle modification, then medication is prescribed. In the UK, pharmacological therapy can include metformin, an oral medication, and insulin, which is injectable(Mayne et al., 2025). Other countries have used glyburide as an oral medication as well; however, this is no longer provided. Metformin and insulin are commonly used globally, with glyburide less commonly used due to the association with neonatal morbidities (Sweeting et al., 2024). However, metformin may cause some people gastrointestinal issues and therefore may not be suitable. Furthermore, metformin does cross the placenta, which can have effects on the offspring's growth (Tarry-Adkins et al., 2019). While insulin does not cross the placenta, there may be an aversion to self-injecting (Sweeting et al., 2024). Hence, it is necessary to find the treatment plan that works best for each woman with GDM.

For women with GDM, it is recommended to give birth no later than 40+6 weeks of gestation. To ensure this, elective birth, Caesarean birth and induction of labour are offered (Scottish Intercollegiate Guidelines Network, 2024). After birth, GDM resolves, SMBG do not need to be taken, any medicine for GDM is stopped, and a referral to a weight management program is given (Scottish Intercollegiate Guidelines Network, 2024).

To confirm that GDM has resolved and there is no underlying diabetes, a postpartum test should be taken; at 6-13 weeks postpartum, a fasting glucose test or, if this was missed, after 13 weeks postpartum, an HbA1c test is recommended (Scottish Intercollegiate Guidelines Network, 2024; Simmons et al., 2024). Due to the risk of type 2 diabetes, annual HbA1c tests should also be taken (Scottish Intercollegiate Guidelines Network, 2024). However, postpartum and annual testing are poorly attended; a 48% attendance rate has been reported within 0 to 1 year postpartum, which drops to 33% after 2 to 3 years postpartum (Daly et al., 2018).

1.2.1.4 Impact of GDM

Being diagnosed with GDM can be psychologically challenging; thus, diagnosis is associated with depression, anxiety, and stress (Sweeting et al., 2024). Stigma, both internalised and external, brings feelings of shame, guilt, and failure, which can be driven by societal pressure and expectations of motherhood, alongside friends and family's lack of understanding of GDM (Davidsen et al., 2022). Consequently, this has led some women to severely restrict their diets out of fear of spiking their blood glucose, avoid screenings, and misreporting their blood glucose readings to avoid medication (Davidsen et al., 2022; Davidsen et al., 2023; Draffin et al., 2016). Despite these burdens, their child's health is a strong motivator for change (Skar et al., 2018).

GDM is prevalent, but the cost remains unclear. The cost-effectiveness and treatment of GDM are influenced by various factors that differ depending on the population and screening methods. It is estimated that GDM costs USD 5.5 billion per year in China and USD 1.6 billion annually in America (Hivert et al., 2024). Factors influencing the cost-effectiveness of GDM are detailed by Sweeting et al. (Sweeting et al., 2024). These include the following: early screening for GDM can reduce adverse outcomes but may also be costly; universal testing may only improve outcomes under circumstances where screening and treatment are not well established and regardless of the screening method, low treatment adherence will result in poor outcomes. Therefore, in populations with a high prevalence of GDM, screening and treatment are likely to be most cost-effective.

1.2.2 Machine learning

1.2.2.1 Overview

Machine learning is the use of algorithms to make predictions within a set range of data by identifying patterns and trends from the input data (Handelman et al., 2018). It is often deployed in solving real-world problems in daily lives (Rebala et al., 2019), such as the spam filter on an email inbox. It is a subfield of artificial intelligence (AI) (Rebala et al., 2019), which has become

a commonplace “buzzword” partly thanks to the increased accessibility of generative AI.

In machine learning, the algorithm learns from the available data; how it learns depends on the input data’s structure, and how it learns to produce an output depends on the problem. In this section, I will give an overview of machine learning from “An Introduction to Machine Learning” by Rebala et al. (Rebala et al., 2019) and “Machine Learning Algorithms” by Bonaccorso, (Bonaccorso, 2018) . These textbooks have been chosen as they give a clear, comprehensive overview of machine learning. Where applicable, I have illustrated each model with an application relevant to GDM.

1.2.2.1.1 Learning models

The primary learning models include supervised learning, unsupervised learning, semi-supervised learning, and reinforcement learning. The way the data are labelled and presented depends on the learning approach of the algorithm. A summary of machine learning appears in **Figure 1.2**. Reinforcement learning is beyond the scope of this thesis and therefore will not be discussed further.

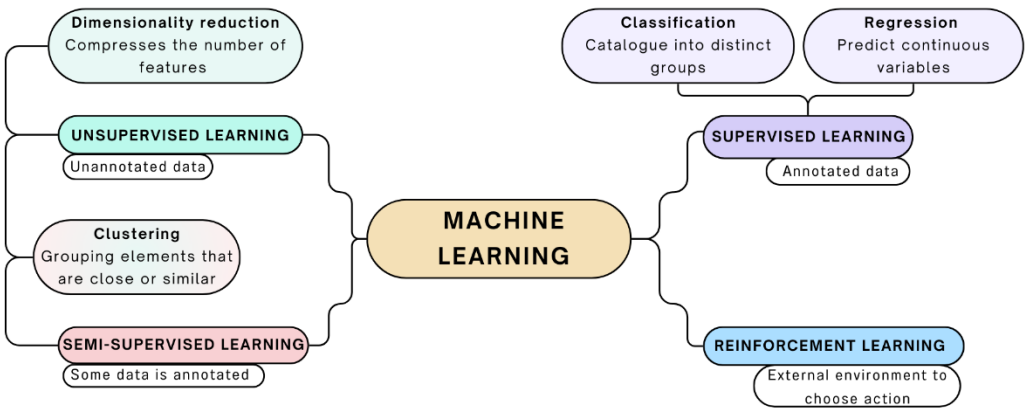


***Figure 1.2** Summary of machine learning’s learning models and problems*

1.2.2.1.1.1 Supervised machine learning

Supervised machine learning requires an annotated dataset in which the algorithm can learn key characteristics of the labelled predictor and use that knowledge on unseen data. For instance, a supervised machine learning algorithm could predict which women with GDM would require pharmacological intervention based on a dataset of GDM-affected pregnancies that indicate whether pharmacological intervention was needed. This illustrates a classification problem. Supervised machine learning encompasses two main types of problems: a classification problem, which predicts the distinct class or category to which an item belongs, and a regression problem, which predicts a continuous variable.

1.2.2.1.1.2 Unsupervised machine learning

Unsupervised machine learning does not require an annotated dataset. Given a large enough dataset, an algorithm can identify trends and patterns that recognise clusters or groups. These include dimensionality reduction, where the number of features is compressed or reduced without loss of information, and clustering, where elements that are 'close' to each other are grouped together. For example, in a dataset with SMBG data consisting of blood glucose values and timestamps, an unsupervised clustering algorithm could group these readings into associated meals with which they were taken.

1.2.2.1.1.3 Semi-supervised learning

If the data is partially annotated, then it is in between supervised and unsupervised learning. These use clustering algorithms, as described above (1.2.2.1.1.2), to make predictions. Semi-supervised learning can be useful to quicken the machine's learning process, as some groups and clusters will already be labelled; similarly, if labelling the data is manual, it reduces the time taken to label data.

1.2.2.1.2 Data

Data plays a huge role in machine learning. The quality and reliability of predictions and outcomes from a machine learning model depend

considerably on the quality of the data (Sandfeld, 2024). If the data are of poor quality or lacking in appropriate quantity, then model development will similarly lack meaningful results.

Within health data science, one of the main sources of data is electronic health records (EHR). Utilising this data source can allow HCPs to learn more about their patient groups and discover trends within patient groups (Goyal and Malviya, 2023). However, this can only be achieved if the data are well-structured, clean and balanced (Fregoso-Aparicio et al., 2021). Approximately 20% of health data is structured, with the remainder being unstructured, non-standardised and multimedia, leading to issues around the quality of data, including incompleteness, inconsistencies, accuracy, heterogeneity and data fragmentation (Goyal and Malviya, 2023).

1.2.2.1.3 Model evaluation and validation

In this section, I will discuss the common predictive performance methods for machine learning models.

During the development of a machine learning model, the data are often split into a training and a testing dataset, which allows assessment of how well the model is performing (Brink et al., 2017) and is often referred to as internal validation (Ramspek et al., 2021). The most common methods are cross-validation and holdout. Cross-validation splits the data into training and test, usually 5 or 10, which are called folds. The model is then trained and tested on each fold, and the results are accumulated (Brink et al., 2017). Holdout is a form of cross-validation, where a proportion of the data, usually 20-40%, is used for testing and the rest of the data is used for training (Brink et al., 2017). Another method is bootstrapping, which is a resampling method. A random selection of the data is sampled with replacement; used for training and testing, the data can be sampled multiple times. This is repeated numerous times, often 100 to 200 times and the results are pooled (Ramspek et al., 2021).

To understand how well a model is performing, metrics such as accuracy, confusion matrix and area under the receiver operating characteristic (AUROC) curve are frequently used. The most basic metric is accuracy, which is the proportion of predictions that were correct (Brink et al., 2017). However, more information on how the model is actually performing for each class can be obtained from the confusion matrix (Brink et al., 2017). This is a matrix that shows the proportion of predictions that were true or false for each class; in other words, it shows the number of predictions that were correct or not. For each class, the matrix reports the true-positive rate (sensitivity), the false-negative rate (miss rate), the false-positive rate (fall-out) and the true-negative rate (specificity). The most common performance metric is the AUROC, which plots the false-positive rate against the true-positive rate; the larger the area under the curve is the better the model is performing (Brink et al., 2017).

To fully test the model and assess the generalisability of the model's prediction, external validation is performed. External validation tests the model on unseen structurally different data from the training and testing data (or internal validation data) (Ramspek et al., 2021). These are often temporal, where data are from a different time period or geographical, where data are from a different region or location (Ramspek et al., 2021).

1.2.2.2 Machine learning in diabetes and maternity care

The benefits of AI or machine learning in healthcare settings are considered ambivalent by some. In a review of data-driven AI in clinical decision support systems (CDSS) in healthcare, results were 'limited and contradictory' (Cresswell et al., 2020). Yet, machine learning has been shown to provide statistically significant improvements in health outcomes when incorporated into digital health interventions within real-life studies (Triantafyllidis and Tsanas, 2019), as described within this section.

The application of machine learning within healthcare includes medical imaging, disease diagnosis and forecasting, personalised medicine, drug discovery and development, virtual assistants and chatbots, healthcare

administration and resourcing, and healthcare robots (Haldorai et al., 2024; Hasan et al., 2023) using EHR, genomic data and data from wearable devices (Goyal and Malviya, 2023). It is a wide-reaching area with diverse applications across many different health conditions. Within diabetes, it has been established for blood glucose management (section 1.2.2.2.1). In maternity care, including GDM, predictive models have been developed, yet are limited due to the lack of robust external validation (sections 1.2.3.2.2 and 1.2.2.3).

1.2.2.2.1 Use of machine learning in diabetes care

The integration and analysis of data sources give remit for personalised medicine through tailored recommendations, particularly on optimal glucose levels, nutrition, and overall health, which can empower people living with diabetes to have an active role in their condition and help improve self-care and outcomes, while assisting HCPs to effectively manage diabetes (Sarma and Devi, 2025; Rhee and Rhee, 2024). This can be achieved through machine learning. Machine learning and AI have a wide application within diabetes care, including patient education and self-management, CDSS, including treatment and screening support and complication prediction and using multiple data sources and types (Mackenzie et al., 2024). Furthermore, risk prediction models can facilitate earlier intervention for complications such as retinopathy (Sarma and Devi, 2025) and diabetic foot (Mackenzie et al., 2024). Diabetes care is already utilising machine learning through insulin pumps and continuous glucose monitors (CGM) to monitor and alert blood glucose, insulin dosage and medication management (Dankwa-Mullan et al., 2019).

1.2.2.2.2 Use of machine learning in maternity care

Recent studies have demonstrated the potential that AI and machine learning have within maternity care to improve accuracy, clinical decision making and healthcare outcomes (Giaxi et al., 2025). Current applications include embryo selection, prediction of pregnancy complications, fetal monitoring and assessment of maternal and neonatal outcomes (Sarno et al., 2023; Giaxi et al., 2025).

Several systematic reviews on AI-CDSS in obstetric and pregnancy care have highlighted the potential for predicting and forecasting complications, as well as aiding in clinical decision making. This enables early intervention, which is critical for improving outcomes in prenatal care (Abdalrahman Mohammad Ali et al., 2025; Lin et al., 2024). However, there was a lack of externally validated models, raising concerns over the models' reliability and generalisability (Abdalrahman Mohammad Ali et al., 2025; Du et al., 2023). Few models and CDSSs have been implemented and user-tested; nevertheless, where they have, they have been received positively by HCPs and shown to have a positive effect on pregnancy care (Du et al., 2023).

Despite a considerable volume of research in this area, this area is still in its infancy, with most of the models being developed in silo and lacking validation, halting the movement towards implementation.

1.2.2.3 Machine learning in GDM

Machine learning has been applied to different aspects of GDM care and management, for example, to predict adverse neonatal outcomes (Pintaudi et al., 2018), predict blood glucose levels (Pustozerov et al., 2018) and to identify factors associated with not attending post-partum follow-up testing (Periyathambi et al., 2022). Here, I will illustrate examples from the available literature of machine learning used to predict GDM, pregnancy outcomes or complications associated with GDM, pharmacological therapy in GDM and type 2 diabetes after GDM.

1.2.2.3.1 Predicting GDM

A cost-effective machine learning screening method can predict GDM (Zhang et al., 2022) which could be used clinically in countries where gestational care is limited due to geography, facilities and cost, or to reduce the number of unnecessary or missed OGTT by identifying women who need to be screened.

In a systematic review and meta-analysis, Zhang et al. (Zhang et al., 2022) identified 25 studies from 2004 to 2020 that used machine learning to predict women who would develop GDM. The authors reported that studies using

machine learning had high accuracy in the early detection of GDM, with a pooled area under the receiver operating characteristic curve (AUROC) of 0.85, sensitivity 0.69 (95% confidence interval (CI) 0.68-0.69; p<0.001; I^2=99.6%), and specificity of 0.75 (95% CI 0.75-0.75; p<0.001; I^2=100%). Logistic regression was the most popular algorithm used (n=17, 68%), but it was outperformed by other algorithms, with an AUROC of 0.82 compared to an AUROC of 0.89. The models used pre-pregnancy factors to predict GDM, the most common being maternal age, family history of diabetes, pre-pregnancy BMI and fasting blood glucose. Despite the high performance of these models, just over half were internally validated (n=13, 52%), and only four (16%) were externally validated. None were reported to be used in a clinical setting, and it was reported that screening for GDM using machine learning is not widely recommended by guidelines.

In addition, Du et al. (Du et al., 2022) reviewed nine studies between 2015 and 2021 that predicted GDM through machine learning; none were designed for a clinical setting. Studies often did not address class imbalances and did not address the issues associated with different cultural or ethnic backgrounds in GDM prediction, which would result in a more biased model being developed. In general, this results in a less clinically useful or accurate overall model.

1.2.2.3.2 Predicting pregnancy outcomes and complications associated with GDM

Women who have GDM have a high-risk pregnancy; having GDM puts a woman at a greater risk of having other pregnancy complications, such as pre-eclampsia and a LGA baby (World Health Organisation (WHO) & International Diabetes Federation (IDF), 2021, 13 April). Women with high-risk pregnancies could benefit from targeted care; equally, health boards could benefit by targeting resources. Machine learning has been used to identify those women with GDM who are at a higher risk of developing adverse pregnancy outcomes.

Cooray et al. (Cooray et al., 2020) identified five studies that contributed ten models for predicting pregnancy complications associated with GDM using machine learning. The most common complication predicted was LGA; other

complications that were predicted included adverse neonatal outcomes, Caesarean birth and pre-eclampsia. Characteristics that were used most frequently in the models as indicators of pregnancy complications were glycaemic measures, BMI, and maternal age. Models that predicted a single outcome had a higher accuracy than models that predicted a composite outcome, and there was a large variation in the performance of the models, with the AUROC ranging from 0.52 to 0.91. Cooray et al. found that the models had high bias due to methodological limitations and statistical analysis and there was a lack of externally validated. Alongside the lack of external validation, there was also a lack of measures of clinical usefulness. Prediction of pregnancy complications is complex, as many factors could contribute. There has not been much research into this area; the models that have been developed lack clinical relevance and are likely to exhibit a bias towards the training data due to a lack of external validation.

1.2.2.3.3 Prediction of pharmacological therapy for GDM

Achieving good glycaemic control prevents adverse pregnancy outcomes for both mother and child (Hashimoto and Koga, 2015). It is thought that identifying women who are more likely to require pharmacological therapy to achieve optimum blood glucose levels reduces adverse pregnancy outcomes (Meshel et al., 2016), potentially due to closer monitoring and earlier interventions with pharmacological therapies. In addition, early identification could allow for better therapeutic strategies and allocation of resources.

Alvarez-Silvares et al. (Alvarez-Silvares et al., 2022) presented a systematic review and meta-analysis to identify risk factors that may indicate the need for insulin therapy upon the diagnosis of GDM. The authors reviewed 18 observational studies that included 14,951 women with GDM, of which 5,731 were on insulin, from 1975 to 2021; these studies did not exclusively use machine learning. It was found that there was a strong causal relationship and statistically significant association between having a BMI ≥ 30 kg/m^2 (relative risk(RR):2.2; 95% CI: 1.44–3.41), a family history of type 2 diabetes (RR:1.74; 95%CI: 1.56–1.93), prior history of GDM (RR:2.10; 95%CI: 1.56–2.82), and

HbA1c at diagnosis (RR:2.12;95%CI: 1.77–2.54), and basal glycaemia (RR: 1.2; 95%CI: 1.12–1.28), with the need for insulin treatment in GDM.

Further, Benham et al. (Benham et al., 2023b), investigated the predictors for the need for pharmacological therapy in addition to lifestyle intervention. Thirty-four studies were identified, 23 of which were included in the meta-analysis and 11 in a narrative synthesis. They reported that nulliparity, lower BMI, no previous history of GDM, lower HbA1c, fasting, 1-hour and 2-hour glucose, later gestation of diagnosis of GDM and later gestation at initiation of the oral agent were markers for pharmacological therapy. While lower maternal age, nulliparity, lower BMI, no previous history of GDM, lower HbA1c, fasting, 1-hour, 2-, and 3-hour glucose, no family history of diabetes, later gestation of diagnosis of GDM, and no macrosomia in previous pregnancies were markers for the prediction of success of lifestyle intervention.

The use of machine learning to predict pharmacological therapy is further explored in a scoping review in Chapter 3.

1.2.2.3.4 Prediction of type 2 diabetes after GDM

Women with GDM are at an elevated risk of developing type 2 diabetes (Vounzoulaki et al., 2020). Identifying the subgroup that is at the highest risk of developing type 2 diabetes would allow for earlier and more targeted intervention and could reduce anxiety in women who had GDM.

A systematic review and meta-analysis of 13 studies, between 2011 to 2024, with 11,320 women with GDM and 22 machine learning models, examined the progression of GDM to type 2 diabetes (Zhao et al., 2025). Zhao et al. reported a pooled C-statistic of 0.82 (95% CI: 0.79 ~ 0.86), a sensitivity of 0.76 (0.72 ~ 0.80), and a specificity of 0.57 (0.50 ~ 0.65). Machine learning had a good diagnostic accuracy, and there was no significant difference in the C-statistic between the different algorithms employed, although logistic regression did have a lower accuracy in comparison to other algorithms.

Furthermore, Belsti et al. (Belsti et al., 2023), conducted a systematic review of 15 papers on predicting postpartum glucose intolerance in women with a

history of GDM. The predicted power of the included studies ranged from an AUROC of 0.66-0.92, with less than half (n=7, 47%) internally validated and none externally validated. Common variables for prediction were BMI, fasting glucose during pregnancy, maternal age, family history of diabetes, biochemical variables, OGTT, use of insulin in pregnancy, postnatal fasting glucose level, genetic risk factors, HbA1c and weight. Traditional statistical methods were more common than machine learning.

1.2.3 Digital health

Digital health is broadly the application of AI, machine learning and data science through technology and is often referred to as telemedicine, telehealth (tele referring to telephone), eHealth (electronic health) and mHealth (mobile health) (Istepanian, 2022). Digital health is provided through voice and video calls, messages, and apps connected through the phone networks, the internet, and Bluetooth. Telehealth systems primarily focus on remote communication, whereas digital health interventions focus more on self-management through online education and data-driven approaches; however, some interventions may encompass all elements. Digital health technology has made advances within healthcare by increasing the efficiency of communication between HCPs and patients, increasing patient engagement, providing real-time analytics, and streamlining care decisions and work (Mohamed et al., 2025).

Digital health technologies can be a useful addition to GDM care as they can reduce the lag between data inputs (i.e., SMBG) and treatment adjustments by HCPs (Xie et al., 2020). They can also help in standardising and reducing the human error when recording SMBG (Garnweidner-Holme et al., 2015), such that women with GDM are not reliant on waiting for their next appointment for any treatment changes or advice, and similarly, HCPs are not reliant on the patient's accuracy and record keeping and can effectively implement treatment earlier. Digital health technologies can reduce the number of clinic appointments (Dalfrà et al., 2009), making care more accessible to women living rurally (Given et al., 2015) and less of a burden to those who would have

to arrange care or time off work (Harrison et al., 2017). Furthermore, it can reduce the anxiety associated with a high-risk pregnancy and make the understanding of risk and how to manage it less overwhelming (Yang et al., 2018; Conway et al., 2019). In all, it can give the freedom for women with GDM to have a more normal pregnancy (Dalfrà et al., 2009).

Though there are many benefits to digital health technologies, they have been found to increase workloads of HCPs. In a retrospective study on the introduction of GDm-Health (further discussed in 1.2.3.3.2.1.2) within a London hospital (Hyams et al., 2024), it was found to increase the number of interactions per women between GDM diagnosis to delivery from a median of 5 (interquartile range (IQR) 3-6) to 7 (IQR 5-10), $p<0.001$. This also included an increase in telephone interaction from 13.7% to 88.4%. Similarly, increase in HCPs workloads was found in a systematic review of 15 papers investigating the unintended effects of mHealth interventions (Cao et al., 2025). Cao et al. identified 26 distinct effects of mHealth on the community and institutions, interpersonal and individuals. The author highlights issues around the changes in HCPs roles as mHealth technologies are introduced which may move to a more administrative task or duplications of tasks on multiple systems, as well as less physical interactions with patients. There are also concerns over the data that mHealth collects, including its safety and accuracy, despite there being more available data it may cause more noise rather than more information. In conduction, if the digital health technology also includes as CDSS, then this raises issues around the ambiguity of responsibility for any clinical decision (Rowland et al., 2022), particularly if information could be distorted or missing within the technology.

1.2.3.1 Evolution of digital health

Digital health has progressed in line with technological advancements. In the early 2000s, voice calls and text messages were commonly used (Ameyaw and Baatiema, 2025). For example, Pérez-Ferre et al. (2009) used text messages to send SMBG from women with GDM to HCPs, and HCPs were able to send texts back. Women in the text message group had a 62.0%

reduction in the number of unscheduled face-to-face visits, which was 87.2% lower in the insulin-treated women, improving patient satisfaction while maintaining similar pregnancy and neonatal outcomes. Within the 2010s, smartphones became widespread, allowing for tracking and personalised health information, and now in the 2020s, there is a drive for integration into the healthcare system (Ameyaw and Baatiema, 2025). The World Health Organisation (WHO) recognised this need for the integration of digital health interventions into the healthcare system and, in 2019, published guidelines (World Health Organisation (WHO), 2019). The catalyst force behind the integration into healthcare systems came from the pandemic, where there was a pressing need to safely provide care remotely and minimise in-person interaction (Mohamed et al., 2025; Ameyaw and Baatiema, 2025). During the pandemic, GDm-Health (Huma, United Kingdom), an mHealth app for GDM (further discussed in 1.2.3.3.2.1.2), was made free to use for the National Health Service (NHS) in England and is now used in 47% of Trusts in England (Lu et al., 2023).

1.2.3.2 Digital health in diabetes and maternity care

There is mounting evidence that digital health systems can have a beneficial impact on the self-management of chronic illnesses such as diabetes (El-Gayar et al., 2013). However, in many systems, there is a lack of CDSS (Klonoff and True, 2009), meaning that the method simply facilitates HCPs to provide care remotely rather than in person. Moreover, a lack of personalised feedback for behavioural changes reduces the effectiveness of the intervention (Cotter et al., 2014).

1.2.3.2.1 Digital health in diabetes

Digital health interventions are well-established and common within diabetes care. Predominantly, the use of Bluetooth-enabled glucometers with mobile apps and clinical dashboards is used to monitor and track blood glucose levels and allows for remote access to data (Rhee and Rhee, 2024; Bauer, 2024). This allows for faster responses to blood glucose fluctuations; however, the use of mobile apps in diabetes has not been robustly shown to

improve long-term clinical outcomes (Rhee and Rhee, 2024). Nevertheless, enhancement in digital health technologies for type 1 diabetes now includes CGM and insulin pumps, improving the self-management and empowerment of people living with diabetes (Bauer, 2024). With the large volume of data that is generated from diabetes care, from EHR, regular blood glucose readings, wearables and apps (Mackenzie et al., 2024), machine learning can be applied to extract information (Kavakiotis et al., 2017). Digital health systems can then provide a person living with diabetes and their healthcare team a more complete overview of their condition, especially when linked to a personal health record, including test results, and educational materials (Conway et al., 2019). This can streamline the care and self-management of a person living with diabetes by having all their information in one place.

In Scotland, MyWay Digital Health (MyWay Digital Health Ltd, United Kingdom), (Wake et al., 2016; Conway et al., 2019; Cunningham et al., 2013), a University of Dundee spinout which launched in 2008, is an online platform for people living with diabetes and their carers. It provides a variety of multimedia resources about diabetes and self-management. In addition, users can access their personal health records, including diabetes related test results. The personal health record is linked to SCI-Diabetes, NHS Scotland's diabetes record (Scottish Care Information Diabetes Collaboration, 2015), with data from primary and secondary care, specialist screening and lab results. A survey conducted by MyWay Digital Health on user experiences of using such a system found that it is a helpful and useful addition to standard care for diabetes self-management (Conway et al., 2019).

1.2.3.2.2 Digital health in maternity care

Among women of reproductive age, >90% are actively engaged in mHealth apps (Venkatesh et al., 2024), and such apps can be helpful in pregnancy as a source of information both antenatally and postnatally for themselves and their child (Kusyanti et al., 2022). Uses of common apps include prenatal tracking, fetal monitoring, labour and delivery support, breastfeeding support and postpartum care platforms (Ameyaw and Baatiema, 2025). Digital health

gives an accessible platform for data-driven innovation, providing information synchronously or asynchronously (Rhee and Rhee, 2024).

In a scoping review of 126 studies on the progression and application of digital health in antenatal care (Mohamed et al., 2025), the focus of interventions was on patient-HCPs consultations, remote monitoring, and health education, which were either complementary to care (n=90, 71.4%) or a substitution (n=36, 28.9%). These were mostly targeting within general maternal care (n=36, 28.7%), GDM (n=19, 15.1%), and mental health (n=17, 13.5%); other areas included gestational weight gain, high-risk pregnancy, maternal education and hypertension. Digital health was provided through real-time teleconsultation (n=32, 25.4%), mHealth (n=40, 31.7%), mHealth with target feedback (n=41, 32.5%), and real-time teleconsultation with mHealth (n=13, 10.3%). Mohamed et al. reported that findings showed high satisfaction across the range of digital health, alongside positive outcomes in managing the condition, increasing health knowledge and birth preparedness and improvement in access to care, while also being cost-effective.

Mohamed et al. performed a comprehensive scoping review including a diverse range of study designs, allowing for an understanding of the current literature on the application of and outcomes from digital health technologies in antenatal care. However, as publications with positive results are more likely to be published, a publication bias may have been introduced, potentially overstating the benefits of digital health. In addition, 61.1% (n = 77) of the included studies were from high-income countries, which may limit the applicability of the result to low- and middle-income countries.

1.2.3.3 Digital health in GDM

Digital health tools can help manage GDM by providing real-time glucose monitoring, education, personalised recommendations and remote consultations. Different aspects of GDM care and management have been targeted, such as GDM diagnosis, GDM management, education and advice, and blood glucose monitoring and control (Daley et al., 2021). Primarily, the

use of telephones and smartphones has been utilised in this area. In this section, I will highlight current research on digital health in GDM by first reporting and updating previous literature reviews and then providing a synopsis of key mHealth interventions in GDM found in the literature.

1.2.3.3.1 Reviews of digital health in GDM

The most common feature of GDM mhealth apps is blood glucose management; the majority provide some form of automatic upload from the glucometer to a platform, which allows for monitoring and analysis. In a scoping review of 17 studies, with 1,871 women with GDM, on smartphone apps that provided self-management and remote monitoring through direct blood glucose uploads from glucometers, Smyth et al. (Smyth et al., 2022) highlighted the feasibility, reliability and acceptability of mHealth apps in GDM. All apps in the review had directed upload from the glucometer to the patient-facing app, feedback and advice were provided bidirectional (n=8, 47%), in-app communication (n=6, 35%) or automatically (n=3, 18%). mHealth-assisted care was noninferior to standard care, in terms of maternal and neonatal outcomes, alongside glycaemic indices of the mHealth users improved in nine (53%) of the studies. In studies that reported satisfaction (n=7, 42%), it was high, including HCPs.

Furthermore, regarding mHealth apps for GDM that include AI or CDSS, a scoping review of 18 papers describing 11 apps between 2014 to 2019 was conducted by Daley et al. (Daley et al., 2021). Three of the apps (GDm-Health, MobiGuide (European Commission 7th Framework Program, Spain) and SineDie (Universidad Politécnica de Madrid, Spain, detailed in sections 1.2.3.3.2 and 1.2.3.3.2.1) having gone through clinical trials, with only one (GDm-Health, detailed in section 1.2.3.3.2.1.2) being clinically implemented. Contextual user feedback for self-management was provided in 67% (n=7) of the apps. Twelve (67%) of the studies were to be used by women with GDM and HCPs, with five (28%) to be used for research or clinical use only. It was reported that there were four clinical focuses: diagnosis (n=5, 28%),

management (n=13,72%), ongoing support (n=12, 67%) and data collection in between clinical visits (n=8, 44%). None had a holistic approach to GDM.

Daley et al. scoping review covered literature from 2014 to 2019, applying the same search strategy and extending the search to 2025 (further details in **Appendix A**) two additional papers were found that would have met their inclusion criteria. The additional papers described the experience of using the SineDie app (described in 1.2.3.3.2.1.7) during the pandemic (Albert et al., 2020). The other paper presented a development study on an AI-based app for GDM diagnosis (Shen et al., 2020). From updating the Daley et al. review, there remain three apps that are well developed, but now two have been clinically implemented (GDm-Health and SineDie). Contextual feedback was provided by the two additional papers, and both were intended to be used by women with GDM and HCPs. The clinical focus was ongoing support, data collection between clinical visits (Shen et al., 2020; Albert et al., 2020), management (Albert et al., 2020), and diagnosis (Shen et al., 2020). Other results of Daley et al. review remain unchanged.

In another review, Lu et al. gave an overview of digital health and machine learning technologies for blood glucose monitoring and management of GDM (Lu et al., 2023). Clinical trials of GDM devices and applications before 8th Aug 2021 were reviewed, using the keyword "gestational diabetes" on the National Institutes of Health (NIH) ClinicalTrials.gov database, including only completed trials. Lu et al. identified 19 included complete clinical trials on digital health in GDM. The main focus of the mHealth tools was on blood glucose monitoring, diet and exercise, and medication support. Few had been commercialised and successfully integrated into the clinics.

Applying the same search strategy, and extending the search from 08/08/2021-31/03/2025, returned seven study on digital health technologies for GDM patient monitoring (further details in **Appendix A**). It is noteworthy that there were four clinical studies on using CGM for GDM. Interestingly, these looked at different time points of GDM care where CGM could be applied, two

investigated the monitoring of blood glucose during pregnancy, one during labour and delivery, and one during the post-partum period. The other three clinical trials compared an mHealth app against standard care, the use of text messages for education on GDM, and investigated the use of a wearable ring as a holistic approach in preventing type 2 diabetes.

1.2.3.3.2 mHealth apps for GDM

1.2.3.3.2.1 mHealth in GDM care and management

As seen in the previous two sections, there is a lot of research within digital health for GDM. Different mHealth tools have different objectives, for instance, the Pear study (Ainscough et al., 2020) focused on the prevention of GDM in the overweight and obese pregnancy population, while most focus on the self-management of GDM. In this section, I will summarise the key well-developed mHealth tools for GDM self-management from the literature. **Table 1.1** gives an overview of the key tools discussed.

Table 1.1 *Summary of key well-developed mHealth tools for GDM self-management from the literature*

Name, reference	Development status	Main features	Evaluation highlights	Evaluation highlights details
eMOM (Kytö et al., 2022b)	Prototype testing (10 participants) RCT[a] (148 participants)	Visualisation of BG[b], physical activity and nutrition CGM[c] and physical activity sensors Self-management and learning	Fewer off-target fasting	Control 18.7% (17.3-20.1 95% CI[d]) vs intervention: 15.1%(13.8-16.4), p=0.003)
			Increase in vegetable uptake	Percentage difference between control and intervention 11.8% (0.4 - 23.3), p=0.043
			Increase in light physical activity	Percentage difference between control and intervention 22.8% (5.7 to 39.9) p=0.009
			Fewer macrosomia babies	Percentage difference between control and intervention 19.7% (10.5-20.0) vs 6.6% (1.0-12.2), p=0.36
GDm-Health (Mackillop et al., 2014)	Feasibility and acceptance study (49 participants) RCT[a] (203 participants)	Remote BG[b] monitoring Bidirectional communication	High user satisfaction	Control median 44.5,(IQR[e] 41-46); intervention: median 43 (39-46), p=0.049
			High BG[b] compliance	Mean readings per day control 2.63[1.71 SD[f]] vs intervention 3.80[1.80] p<0.001
			Fewer Caesarean births	Control: 47/102, 46.1%, intervention: 27/101, 26.7%, p=0.005

Name, reference	Development status	Main features	Evaluation highlights	Evaluation highlights details
	Implemented clinically, England UK	Feedback through BG[b] visualisation and colour coding	Cost-effective	Intervention cost compared to standard care –£1044 (95% CI[d] £2186 to £99).
GDMapp (Smyth et al., 2025)	Observational pilot study (168 participants)	Remote BG[b] monitoring Feedback through BG[b] visualisation and colour coding Education	Fewer off-target fasting BG[b]	Mean fasting blood glucose historical 5.0 (0.92 SD[f]) (mmol/l) vs intervention 4.8(0.4), p=0.022 Mean postprandial blood glucose 6.6(1.2) vs 6.1(0.4), p <0.001
MobiGuide (García-Sáez et al., 2014)	Feasibility study (19 participants)	Remote monitoring of BG[b], ketonuria, steps and blood pressure Personalised advice	High user satisfaction	-
			High monitoring compliance	Mean historical blood glucose compliance 0.87 (0.28 SD[f]) vs intervention 1.01 (0.10)
			Improved clinical overview	-
MoTHer			User satisfaction	87.1% of app users were satisfied with the app for managing GDM[g]

Name, reference	Development status	Main features	Evaluation highlights	Evaluation highlights details
(Butten et al., 2024)	Feasibility and acceptance study (40 participants) Implemented clinically, Brisbane Australia	Remote BG[b] , weight, step and blood pressure monitoring Feedback through BG[b] visualisation and colour coding Education Language options	High BG[b] compliance	Average of 3 readings a day
			Similar understanding of GDM[g] and dietary changes	-
Pregnant+ (Garnweidner-Holme et al., 2015)	Prototype testing (21 participants) Acceptability study (9 participants) RCT[a] (238 participants)	Remote BG[b] monitoring Feedback through BG[b] visualisation and colour coding Education	High user satisfaction	-
			Reduction in emergency Caesarean sections	Control: 27(22.1%), intervention: 10 (8.8%), mode of delivery p=0.33
			No difference in postpartum OGTT[h]	Mean control: 6.0(95% CI[d] 5.6 to 6.3), intervention: 6.7 mmol/L (6.2 to 7.1), p=0.22

Name, reference	Development status	Main features	Evaluation highlights	Evaluation highlights details
		Language options	No difference in dietary behaviour	-
SineDie (Caballero-Ruiz et al., 2017)	Implemented clinically, Spain Service evaluation (90 participants)	Remote monitoring of BG[b] and ketonuria Machine learning tagged BG[b] with meal Treatment recommendations	High user satisfaction	-
			High monitoring compliance	Average 3.9 reading per day Data sent an average of every 3.5 days
			Reduces time to evaluate the patient	Median time reduced by 27.4%
			Reduction of in-person appointments	Reduced by 88.6%

[a]RCT Randomised control trail, [b]BG Blood glucose, [c]CGM Continuous glucose monitoring, [d]CI Confidence interval, [e]IQR interquartile range [f]SD Standard deviation, [g]GDM Gestational diabetes mellites, [h]OGTT Oral glucose tolerance test

1.2.3.3.2.1.1 eMOM

eMOM (University of Helsinki, Finland) (Kyto et al., 2024; Kytö et al., 2024; Kytö et al., 2022a; Kytö et al., 2023; Määttänen et al., 2025; Kytö et al., 2022b) was developed in 2022 in Finland. The mobile app visualises CGM data alongside physical activity, sleep and stress data from sensors and nutrition inputs and aims to improve self-management and learning of diet-controlled women with GDM.

In a mixed-methods study, the prototype was tested with 10 women with GDM (Kytö et al., 2023; Kytö et al., 2022a), where the desirable features were clarified and the acceptability of sensors was established. A Randomised Control Trial (RCT) followed this (Kytö et al., 2024; Määttänen et al., 2025) of 148 women, 72 in the control and 76 in the intervention arm. Adherence to eMOM was high and those using the eMOM intervention had a fewer off-target fasting blood glucose (control (% (95% CI): 18.7% (17.3-20.1) vs intervention: 15.1%(13.8-16.4), p=0.003), an increase in vegetable intake from baseline (24-28 weeks gestation) to 35-37 weeks gestation (percentage difference between control and intervention (95% CI) 11.8% (0.4 - 23.3), p=0.043), increase in light physical activity (22.8% (5.7 to 39.9) p=0.009) and fewer instances of macrosomia (19.7% (10.5-29.0) vs 6.6% (1.0-12.3), p=0.36).

1.2.3.3.2.1.2 GDm-Health

GDm-Health (Huma, United Kingdom) (Mackillop et al., 2018; Mackillop et al., 2014; Mackillop et al., 2016; Hirst et al., 2015a) was developed from 2014 to 2016 in England and licensed by Sensyne Health plc in 2017, which was acquired by HUMA Health in 2022 (Lu et al., 2023). To aid with the management of women with GDM during the pandemic, it was made free for NHS England, and as of 2021 is paid for and available in 47% of Trusts in England (Lu et al., 2023).

The mHealth tool is a digital blood glucose management system that facilitates remote SMBG and bidirectional communication with HCPs and women with GDM. It consists of a patient-facing app and a clinical web-based

portal. Glucometers are linked via Bluetooth to the women's app and transferred to the HCPs portal via an internet connection. Visualisation of SMBG through graphs, summaries, and colour-coded readings is used to prompt self-management of GDM.

During the development of GDm-Health, the acceptability and satisfaction were studied (Hirst et al., 2015b). A total of 49 women completed a questionnaire postpartum about using GDm-Health from GDM diagnosis to birth. This questionnaire reviewed the general satisfaction, equipment issues, and relationship with the diabetes care team. It was rated highly: 92% (n=45) found care was satisfactory, 96% (n=47), believed the equipment was convenient and 88% (n=43) believed it was reliable respectively, 86% (n=42) agreed GDm-Health fitted into their lifestyle, and 94% (n=46) agreed they had a good relationship with their care team.

Following the user acceptance study, a RCT was conducted, with 203 women (n=101 intervention, n=102 control) (Mackillop et al., 2018) to determine if the use of GDm-Health remote blood glucose management was as effective as standard care (e.g., paper log-diary and regular in-person clinical visits). They reported that GDm-Health was comparable to standard care, and those using the app had significantly higher number of SMBG readings, ($p < 0.001$, mean 3.80 [Standard Deviation (SD) 1.80] and mean 2.63 [SD 1.71] readings per day in the intervention and control groups, respectively), higher satisfaction ($p=0.049$, intervention: median 43, IQR 39-46; control: median 44.5, IQR 41-46) and fewer Caesarean births ($p=0.005$, intervention: 27/101, 26.7% vs control: 47/102, 46.1%), with other glycaemic, maternal and neonatal outcomes being similar in both groups. There was no significant difference between direct health groups of the two groups; the mean cost difference of the GDm-Health compared to the standard care group was estimated as –£1044 (95% CI –£2186 to £99).

A pharmacological therapy forecasting model was developed using the SMBG and maternal data of the app users (Velardo et al., 2021). This predicts the

likelihood of needing pharmacological therapy in the upcoming weeks during GDM using the SMBG recorded in the app. There has not been any literature indicating that this would be integrated into GDm-Health, moreover, this paper was published with Sensyne Health, which has since dissolved.

1.2.3.3.2.1.3 GDMapp

GDMapp (Royal College of Surgeons Ireland, Dublin) (Smyth et al., 2025), was developed in Dublin Ireland, through focus groups with HCPs and women with GDM. The app is linked to a Bluetooth glucometer and provides instant feedback through colour coded icons and educational materials.

In an observational pilot study between January to December 2021, 168 women with GDM used GDMapp and their outcomes were compared with 162 women with GDM from a historic cohort identified directly before pilot study recruitment, which can be assumed to be during the pandemic. Their results report that both the fasting and postprandial SMBG reading were significantly lower in the intervention group ($p=0.022$, and $p<0.001$ respectively), and fewer instances of out of target SMBG. The maternal and neonatal outcomes were not significantly different between the intervention and historic group.

1.2.3.3.2.1.4 MobiGuide

MobiGuide (European Commission 7th Framework Program, Spain) (García-Sáez et al., 2014; Peleg et al., 2017a; Peleg et al., 2017b; Shalom et al., 2015), developed between 2014-2017 in Spain, is an interactive guideline for GDM. It provides personalised decision support based on data extracted from EHR and biosensors (blood glucose, ketonuria, and accelerometer). It gives women with GDM advice on monitoring and therapy through messages with acknowledgement to reinforce compliance and reminders to follow therapy prescription and to monitor specific parameters, alongside clinical assessment advice suggesting changes in therapy and when to contact the hospital. This is delivered through a smartphone app.

The evaluation of MobiGuide (Peleg et al., 2017a) was through a feasibility study with 19 women with GDM in Spain. It was received positively by women

with GDM and HCPs. Women with GDM were compliant to the monitoring targets. Ten of the women started insulin during the study, MobiGuide notified the HCP of two women with GDM to start insulin before their weekly appointment; therefore, these two women started insulin earlier due to MobiGuide detections. HCPs reported that they were able to check their patients' SMBG readings and ketonuria more frequently.

1.2.3.3.2.1.5 MoTHer

The MoTHer (Australian eHealth Research Centre, Australia) app and clinical dashboard (Varnfield et al., 2021; Butten et al., 2024), developed in Australia between 2019-2023, provides remote blood glucose monitoring through automatic transfer from a Bluetooth-enabled glucometer alongside weight, diet and step monitoring. It has graphs and trends of SMBG readings and colour-coded feedback. It also has educational materials through multimedia and is available in 12 languages.

Using the MoTHer app alongside usual care, in a feasibility and acceptability study (Varnfield et al., 2021), 40 first-time diagnosed women with GDM were recruited between August 2017 to April 2018. HCPs and women with GDM indicated satisfaction with the system, and there was an increase in the number of SMBG recorded per day in comparison to historical matched GDM cases.

The MoTHer system was implemented in the Brisbane tertiary centre, Australia, with a major change to the care model. The effects of this change were analysed through pre- and post- implementation surveys. The app-experience survey showed high levels of satisfaction with care that was received (91.5%, 162/177) and with the app for managing GDM (87.1%, 154/177).

1.2.3.3.2.1.6 Pregnant+

Pregnant+ (Oslo Metropolitan University, Norway) (Garnweidner-Holme et al., 2015; Borgen et al., 2019; Garnweidner-Holme et al., 2020; Garnweidner-Holme et al., 2018; Skar et al., 2018) in Norway, allows automatic transfer of

SMBG from the glucometer to app with instant feedback through visualisations and colour-coded smileys. Its focus is on behavioural diet changes, and the app provides information on healthy eating, physical activity, and GDM, some of which is sent automatically if blood glucose levels are too high. Information is provided in multiple languages, and the app links to recipes from the Norwegian Diabetes foundation, including culturally appropriate options.

The Pregnant+ intervention has been assessed at the prototype level, in a RCT and a qualitative study. The user-involvement study (Garnweidner-Holme et al., 2015) discussed the prototype with 21 women with GDM, firstly in interviews (n=10) considering the use of the smartphone app and presentation, and secondly in a talk-aloud (n=11) session performing tasks with the prototype. Nine HCPs interviewed perceived mHealth, particularly the Pregnant+ app, as an appropriate tool for the care of women with GDM, and the findings suggest it was seen as a useful tool to enhance the care provided to women with GDM (Garnweidner-Holme et al., 2018).

After the design and development, a multicentred RCT was conducted with 238 women (115 intervention, 123 control) with GDM to assess if the behavioural diet changes implemented with the use of the Pregnancy + app had an effect on the postpartum 2-hour OGTT (Borgen et al., 2019). They found that there was no difference between the Pregnant+ group and the control group's 2-hour postpartum OGTT (intervention: 6.7 mmol/L (95% CI 6.2 to 7.1), control: 6.0 mmol/L (95% CI 5.6 to 6.3)). There was a reduction in the intervention group's number of emergency Caesarean births, after adjusting for parity (intervention: 10 (8.8%), control: 27(22.1%), mode of delivery p=0.33). There were no significant differences in other variables, including birth weight, breastfeeding practice, obstetric complications, or NICU admission, and no other adverse effects were reported. As the focus of Pregnant+ was behavioural changes, they also reported that the intervention did not have any significant effects on the dietary behaviour of participants during pregnancy (Garnweidner-Holme et al., 2020).

Of the 115 women with GDM who used the Pregnant+ app during the clinical trial, 17 of them were interviewed about their experience using it (Skar et al., 2018). Women with GDM found that Pregnant+ increased their confidence in their self-management of GDM and their motivation for behavioural change. However, for some women, this contributed to feelings of frustration or obsession. In addition, some technological problems and a lack of support from HCPs limited several women from using the Pregnant+ app.

1.2.3.3.2.1.7 SineDie

SineDie (Universidad Politécnica de Madrid, Spain) (Caballero-Ruiz et al., 2017; Caballero-Ruiz et al., 2016; Albert et al., 2020), developed in 2017 in Spain, is a CDSS system for the management of GDM, through a web-based platform that allows remote monitoring of SMBG, ketonuria, and dietary compliance. Through machine learning, the system automatically labels meals associated with SMBG, and through rule-based logic algorithms recommends treatment changes such as diet modification and proposes to HCPs if insulin needs to be started.

It was clinically evaluated over 17 months with 90 women with GDM (60 in intervention and 30 in control) (Caballero-Ruiz et al., 2017). They reported that the median time for HCPs to evaluate their patient was reduced by 27.4%, alongside the number of face-to-face appointments reduced by 88.6%. All therapy recommendations were safe. Women with GDM satisfaction were high, and they measured their blood glucose 3.9 times per day and sent their monitoring data every 3.5 days. Following the clinical evaluation, it was clinically implemented and used during the pandemic (Albert et al., 2020). No other outcomes were reported.

1.2.3.3.2.1.8 Apps outside the literature

In this area, there is an intersection between academia and industry. Reporting solely on the literature does not give the full picture of what is available or being developed. To demonstrate the other apps available to women with GDM, I searched “gestational diabetes” on the Google Play Store

and Apple App Store (01/04/2025). Popular apps found from the search are listed below. This shows what is currently commercialised, however, it does not give a sense of what is in production.

Carbs and Cals, Diet and Diabetes (Chello Publishing Limited, United Kingdom), not specifically for GDM, provides diet tracking with a focus on glucose levels and is recommended by NHS Dietitians.

DiabetesM (Sirma medical, Bulgaria), for all diabetes types, tracks blood glucose levels, food and nutrition. The app provides reports with visualisations, statistics and charts to understand trends.

Gestational Diabetes (GD) tracker (Rotem Maoz, Unknown) was developed by a woman with GDM during her maternity leave. The app logs blood glucose readings and meals and gives reminders to take blood glucose readings.

Hlgedi (Wildcard OY/ Julija Udodova, Estonia) is a community app developed by a woman who had GDM. The app's focus is on sharing advice, stories and recommendations with women with GDM. It also provides daily recipes and opportunities for meet-ups.

Malama Health (Malama Health, United States), specifically for GDM and available in Spanish and English, is an app providing analysis and tracking of SMBG and educational materials. AI is used to identify food triggers for high blood glucose.

mySugr (mySugr GmbH, Austria), for all types of diabetes, allows tracking of blood glucose, food, and activity, and provides reports.

1.2.3.3.2.1.9 Cost of mHealth in GDM care and management

There is limited evidence of the cost-effectiveness of digital tools in general, particularly in GDM. However, there seems to be favourable effects on the cost and health outcomes, with a positive impact in studies using apps or web-based portals (Gentili et al., 2022). Costs when reported are generally about the cost to the clinic, not for the individual woman. Comparing 98 women with GDM using a smartphone-based remote blood glucose

monitoring platform with 94 historical GDM cases, there was an AUD $23 per patient reduction in cost, including the HCPs' time and cost of the app (Poulter et al., 2022). Similarly, in the RCT of GDm-Health (Mackillop et al., 2018), there was a mean cost difference in the intervention group of – GBP £1044 (95% CI –£2186 to £99), with no significant difference in direct healthcare costs. This was also found in a survey of 73 NHS Trusts on glucose management (Mayne et al., 2025). Theoretically, by risk-stratifying care some of the time pressure on HCPs could be alleviate and generate more time available for higher-risk and more complex patients. In addition, this could reduce the cost to women with GDM who need to come into the hospital to attend appointments.

1.2.3.3.2.1.10 Summary of mHealth functions for GDM

In summary, there is a growing range of mHealth interventions being applied to every aspect of GDM care and management in the literature. **Table 1.2** compares the features of the mHealth apps discussed in this section (1.2.3.3.2.1). It is difficult to draw digital health interventions into a review, given the range of functions and differences in implementation. While there were three RCTs (Garnweidner-Holme et al., 2015; Mackillop et al., 2014; Kytö et al., 2022a), the remaining studies were feasibility studies, historical comparisons and had small sample sizes, which introduced the risks of selection bias and limited generalisability.

Overall, within the literature, there is a shared focus on enhancing remote monitoring and self-management to support women with GDM and HCPs. The use of smartphone apps with a Bluetooth glucometer linked to the clinical portal for remote monitoring is well established and has high user satisfaction. These allow tracking of blood glucose and help with self-management through data visualisations, giving real-time or near-real-time feedback and communication, and personalised care. Though some mHealth tools included alternative language or culturally appropriate dietary advice, most often this is not included. Similarly, there is a lack of education or information about GDM and the future health risks associated with it.

The clinical effectiveness is varied, but from RCTs, mHealth tools have shown that remote monitoring is not inferior to standard care for GDM, it reduces the number of in-person appointments and increases compliance and engagement, with some evidence that it may reduce the rate of caesareans and be cost-neutral. Outcomes often focused on acceptability and satisfaction with the mHealth tool, alongside some maternal and neonatal outcomes. There was a lack of long-term follow-up on mother and child outcomes, reporting on digital literacy, how the tool integrates into clinical workflow and multicentred replication, alongside cost analysis.

Commercially available apps typically tend to be limited to tracking key parameters. Collectively, the evidence supports that mHealth is feasible and safe for GDM care and management, with some already being clinically implemented. However, there is insufficient evidence to conclude that it is greater than standard care, as further studies on the cost and long-term outcomes are required.

Table 1.2 *Summary of features in GDM mHealth apps that facilitate self-management, from the literature and commercially.*

				Literature^					Commercial only					
	Device name	**eMom**	**GDm-Health**	**GDMapp**	**MobiGuide**	**MoTHer**	**Pregnant+**	**SineDie**	**Carbs and Cals: Diet & Diabetes ****	**DiabeteM****	**Gestational Diabetes (GD) Tracker**	**Hlgedi: Gestational Diabetes**	**Malama Health**	**mySugr****
	Author/ Reference	(Kytö et al., 2022b)	(Mackillop et al., 2014)	(Smyth et al., 2025)	(García-Sáez et al., 2014)	(Varnfield et al., 2021)	(Garnweidner-Holme et al., 2015)	(Caballero-Ruiz et al., 2017)	-	-	-	-	-	-
Blood glucose recording	**Automatic***	x	x	x	x	x	x			x		x	x	x
	Manual		x		x		x			x	x	x	x	x
Logging and tracking	**Meals**	x				x		x	x	x	x	x	x	x
	Medication		x							x				x
	Exercise	x	x			x				x				
	Ketonuria				x			x						
	Weight	x				x			x					
Educational and information	**Diet**	x			x	x	x							
	Exercise	x			x	x	x							
	Diabetes & GDM[a]	x			x	x								
	Unspecified			x									x	

				Literature^					Commercial only					
	Device name	**eMom**	**GDm-Health**	**GDMapp**	**MobiGuide**	**MoTHer**	**Pregnant+**	**SineDie**	**Carbs and Cals: Diet & Diabetes ******	**DiabeteM****	**Gestational Diabetes (GD) Tracker**	**Hlgedi: Gestational Diabetes**	**Malama Health**	**mySugr****
Reminders	**Blood glucose testing**									x	x			
	Appointments													
	Data visualisation	x	x	x	x	x	x			x			x	x
Data sharing	**Downloading a report**						x			x	x		x	x
	Automatic transfer		x			x								
	Contact with the clinical team		x											
	Decision support				x			x						
	ML[b] or AI [c]	Future development	Unclear		Unclear			x	Unclear				x	
	Multiple languages					x	x						x	
	Integration with other devices	x								x				x
	Forum											x		
	Recipes						x					x		

*^May also be commercially available. *Connection with a Bluetooth-enabled blood glucose meter or continuous glucose monitor. **Not gestational diabetes specific. [a]GDM: Gestational diabetes mellitus [b]ML: Machine learning, [c]AI: Artificial intelligence*

1.3 RESEARCH GAP

From large RCTs on GDM-specific mHealth apps (Borgen et al., 2019; Mackillop et al., 2018), there was little to no significant difference in maternal and neonatal outcomes, indicating that digital tools in GDM do not compromise care. Therefore, the true benefits of digital health tools for GDM lie in a more qualitative sphere. The use of digital tools allows for greater freedom and autonomy over one's health, empowering women with GDM to access care when it suits them. It also provides opportunities to personalise care, ranging from simple tailored messages to more complex predictive models, while simultaneously offering clinical decision support to HCPs.

The use of digital tools in GDM care and management is established, but these tend to only provide information about blood glucose monitoring. There is a gap to provide a holistic approach to GDM, by improving care experience through streamlining care, and integrating into SCI-Diabetes and MyWay Digital Health to provide life-long diabetes care.

We are seeing an increasing rate of obesity in women of reproductive age, contributing to higher rates of GDM. Additionally, this population possesses a high level of digital literacy and access to smartphones. As GDM becomes more prevalent, it continues to add strain to the healthcare system. Current care is time-consuming, with women being reviewed fortnightly, regardless of their medication risk. Thus, there is an opportunity to streamline care through risk stratification of medication needs and remote monitoring (**Figure 1.3**).

A risk prediction model could identify how likely a woman with GDM is to need insulin. If they are at high risk, they would enter a care stream with more intensive monitoring and appointments, similar to current care; if they were at low risk, they would transition to a care stream that primarily involves remote monitoring with fewer in-person appointments. This would then alleviate some of the pressure on the healthcare system, whilst allowing more time for patients that have more complex needs and adjustments. Monitoring would be conducted through a digital tool, consisting of a patient-facing app and a

clinical dashboard. This digital tool would provide education to enhance understanding of GDM and reminders or notifications for postpartum OGTT. The digital tool would be integrated into MyWay Digital Health and SCI-Diabetes, improving GDM records and providing lifelong diabetes care for those who develop type 2 diabetes.

It is hypothesised that a digital tool for GDM would be positively received by end-users in Edinburgh and Glasgow, and that it has the potential to enhance clinical workflow and self-management of GDM, while also serving as a platform for an insulin risk-prediction model that could stratify care. The schematic of this alternative approach to GDM care and the overall ambition of this thesis is shown in **Figure 1.3**, and the aims and objectives are detailed in Chapter 2.

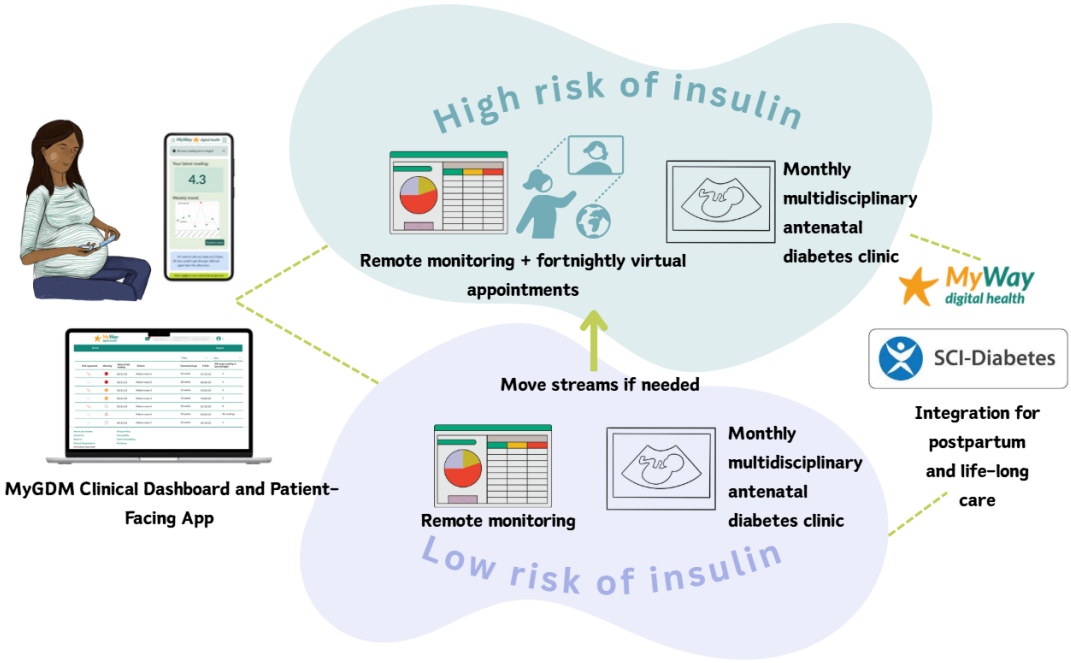


Figure 1.3 *Thesis schematic of gestational diabetes care streamlined through insulin risk and remote monitoring through "MyGDM", a clinical dashboard and patient-facing app.*

CHAPTER 2

AIMS, OBJECTIVES AND THESIS STRUCTURE

2.1 AIMS AND OBJECTIVES

The overall aim of this PhD is to develop a novel data-driven personalised digital tool for women with GDM in Scotland, which incorporates clinical decision support on treatment risk prediction.

The specific objectives are to:

I. Assess current literature around the application of machine learning to predict pharmacological therapy for women with GDM.
II. Identify a cohort of GDM-affected pregnancies in Glasgow. And to analyse and compare the outcomes of: (a) the cohort against treatments for GDM and (b) between GDM and non-GDM pregnancies.
III. Develop and validate a machine learning model to predict insulin therapy in women with GDM in Glasgow.
IV. Assess the needs, acceptability and barriers around digital tools in GDM with women who have GDM and HCPs who care for them.
V. Design and evaluate a user-centred digital tool for GDM care and management.

2.2 THESIS STRUCTURE

This is a mixed-methods PhD, so the specific methods for each component will be described within the corresponding chapter.

Chapter 3 presents a scoping review of the use of machine learning to predict pharmacological therapy for women with GDM (objective I). This has been has been published in the Journal of Diabetic Medicine (Kirkwood et al., 2025b). The results of this chapter will give direction the selection of clinical data variables described in Chapter 4 and set the groundwork for the modelling described in Chapter 6.

Chapter 4 described the cleaning and identification of a cohort of singleton pregnancies from three study dataset of 30,666 pregnancies in Glasgow between 2021-2023. It then identifies the GDM-complicated pregnancies within the cohort (objective II), in preparation for statistical analysis, Chapter 5 and machine learning modelling Chapter 6.

Chapter 5 analyses the cohort described in Chapter 4 (objective II). Comparing maternal characteristics and outcomes between different GDM treatment methods and between GDM and non-GDM affected pregnancies.

Chapter 6 described the development and internal validation of an insulin risk prediction model (objective III).

Chapter 7 presents the user-centred design of a digital tool for GDM, including the user-needs assessment, design and evaluation (objectives IV and V). The qualitative user-needs assessment has been published at the Medical Informatics Europe Conference 2025 (Kirkwood et al., 2025a), and the design and evaluation have been published in the Journal of Diabetes Science and Technology (Kirkwood et al., 2024).

Finally, **Chapter 8** will summarise and discuss the content of this PhD and present future work.

CHAPTER 3

Scoping Review of Machine Learning Application for Predicting Pharmacological Therapy in GDM

3.1 INTRODUCTION

GDM is treated through lifestyle changes, including diet and exercise, or pharmacological therapy, such as oral agents (metformin or glyburide) or insulin. If SMBG readings are not within target, an HCP will adjust the treatment plan for a woman with GDM. Early identification of women needing pharmacological therapy could enhance treatment approaches and optimise therapeutic outcomes. Previous reviews (described in1.2.2.3.3) identified risk factors for pharmacological therapy but have not focused on the application of machine learning to predict pharmacological therapy.

Therefore, a scoping review was conducted to investigate the current literature on the application of machine learning in predicting

pharmacological therapy in GDM. The scoping review results informed the variable selection for the data analysis and modelling described in Chapters 4, 5 and 6.

3.2 AIMS AND ENDPOINTS

This Chapter will focus on objective: I Assess current literature around the application of machine learning to predict pharmacological therapy for women with GDM. The specific aims and endpoints of this Chapter are:

3.2.1 Aims

1) Map the available evidence regarding the use of machine learning to predict pharmacological therapy in GDM within the literature, and in doing so, identify the knowledge gaps.
2) Comprehensively review the methods, variables and quality in machine learning used to predict the need for prescribing or the escalation of pharmacological therapy in GDM, both at the start of or throughout pregnancy.

3.2.2 Endpoints

1) Understand key variables (features) that have been used in the literature to predict pharmacological therapy in GDM and use this to create a list of potential variables to make my own pharmacological therapy predictive model (Chapters 4 and 6).
2) Inform the development of my models (Chapter 6) and highlight areas of improvement by identifying prior work and current research gaps.

3.3 CHAPTER STRUCTURE

This chapter presents the scoping review entitled 'The use of machine learning to predict pharmacological therapy in gestational diabetes: a scoping review' (Kirkwood et al., 2025b), published in the Journal of Diabetic Medicine in November 2025. Below, in **Table 3.1**, the Contributor Roles Taxonomy (CRediT) (CRediT, 2022) has been used to provide an understanding of each author's contribution to the article.

To aid with readability, the full text of the published manuscript has been formatted into the style of the thesis (section 3.4), all supplementary material can be found in **Appendix B**; where appropriate, some supplementary material has been included with the chapter. Otherwise, the structure and content of the published manuscript has not been altered.

Table 3.1 *Authors' contributor role, following CRediT, for the publication of The Use of Machine Learning to Predict Pharmacological Therapy in Gestational Diabetes: a Scoping Review, Journal of Diabetic Medicine.*

Contributor Roles	**Author's initials**
Conceptualisation	JRK
Data curation	JRK, NG
Formal analysis	JRK
Funding acquisition	RMR, DJW, RSL, AM
Investigation	JRK
Methodology	JRK
Project administration	JRK
Resources	Not applicable
Software	Not applicable
Supervision	RMR, DJW, RSL, AM
Validation	NG
Visualisation	Not applicable
Writing – original draft	JRK
Writing – review & editing	RMR, DJW, RSL, AM, NG

JRK Jasmine R Kirkwood, NG Natasha Galloway, RSL Robert S Lindsay, AM Areti Manataki, DJW Deborah J Wake, RMR Rebecca M Reynolds

3.4 The Use of Machine Learning to Predict Pharmacological Therapy in Gestational Diabetes: A Scoping Review [Published in the Journal of Diabetic Medicine]

The use of machine learning to predict pharmacological therapy in gestational diabetes: a scoping review

Short title: Predicting pharmacological therapy in GDM: a review

Jasmine R Kirkwood[1], Natasha Galloway[2], Robert S Lindsay[3], Areti Manataki[4], Deborah J Wake[5], Rebecca M Reynolds[1]

1. Centre for Cardiovascular Science, Queen's Medical Research Institute, The University of Edinburgh, EH16 4TJ, UK
2. Edinburgh Centre for Endocrinology and Diabetes, NHS Lothian Hospitals Trust
3. School of Cardiovascular and Metabolic Health, The University of Glasgow, G12 8TA, UK
4. School of Computer Science, University of St Andrews, KY16 9SX, UK
5. Usher Institute, The University of Edinburgh, EH16 4T, UK

Corresponding author, Jasmine R Kirkwood, j.r.kirkwood@sms.ed.ac.uk.

Conflicts of interest: There are no conflicts of interest.

Novelty statement

- Machine learning has been applied to many aspects of pregnancy care and diabetes.
- Logistic regression is a popular algorithm to predict pharmacological therapy for gestational diabetes. Most current models lack external

validation, and as far as can be ascertained by a review of the published literature, none have been implemented clinically.

- All models are binary classifiers predicting either insulin or pharmacological therapy (grouping insulin and oral agents). Alongside external validation and clinical implementation, there is a gap in developing a multi-class classifier to identify the risk for an oral agent or insulin for a person with gestational diabetes.

Abstract

Aims: Early identification of pharmacological therapy for gestational diabetes mellitus (GDM), a common pregnancy complication, through machine learning could allow for better therapeutic strategies and improved treatment efficiency. This scoping review aims to comprehensively review the machine learning models used to predict the need for pharmacological therapy in GDM.

Methods: Four electronic databases: Embase, Medline, IEEE Xplore, and Web of Science, were searched for publications between 01/07/07 and 31/08/24. Studies predicting pharmacological therapy for GDM using machine learning were included. The Joanna Briggs Institute and PRISMA-ScR checklist was followed, and the Prediction model Risk Of Bias ASsessment Tool (PROBAST) was used to assess quality.

Results: Included were 17 studies presenting 44 models, 61.4% (27/44) predicted any pharmacological therapy use, and 38.6% (17/44) predicted insulin use alone. All were binary classifiers, and logistic regression was typically used. The overall area under the receiver operator curve (AUROC) had a median of 0.75. Common clinical variables were found to be predictors, such as history of GDM, gestational week at GDM diagnosis, pregestational body mass index, maternal age, HbA1c, fasting – and 1hr- glucose from 75g oral glucose tolerance test. Though 65.9% of models were validated, there was a lack of external validation. There was no evidence of clinical application of the models.

Conclusion: Logistic regression with common clinical variables was often used to predict pharmacological therapy for GDM. Few models were externally validated or clinically applicable.

Keywords: Gestational diabetes mellitus (GDM), Prediction algorithms, Machine learning, Pharmacological therapy, Insulin, Oral agents

3.4.1 Introduction

Gestational diabetes mellitus (GDM) is one of the most common pregnancy complications, affecting 13.4% of live births worldwide in 2019 (Magliano et al., 2021) and is rising in prevalence (Ferrara, 2007). Women with GDM typically attend multi-disciplinary specialised clinics fortnightly, where their self-monitoring blood glucose values (SMBG) and diet and exercise modifications are reviewed (Mackillop et al., 2014; National Institute for Health and Care Excellence (NICE), 2020). If blood glucose targets are unmet, some women will need medication, such as oral agents and/or insulin injections (National Institute for Health and Care Excellence (NICE), 2020).

GDM increases the risk of adverse pregnancy and neonatal outcomes if glucose levels are not well-controlled (Kc et al., 2015; Ye et al., 2022; Hashimoto and Koga, 2015). There is a 10-fold increased risk of women with GDM developing subsequent type 2 diabetes, (Vounzoulaki et al., 2020) and a two-fold increased risk of developing premature cardiovascular disease (Kramer et al., 2019; Carr et al., 2006; Daly et al., 2018). In addition, children of women who had GDM are at a higher risk of obesity during their childhood and adolescence (Buchanan et al., 2012). Early identification of women with GDM who need pharmacological therapy could allow for more efficient therapeutic strategies for each woman and better allocation of resources.

Machine learning is a subset of artificial intelligence that uses algorithms to 'learn' from the data to optimise the performance metric (Jordan and Mitchell, 2015), allowing insight into data that may elude human analysis. Machine learning has been used within healthcare alongside electronic health records for many different applications, including predicting and identifying diseases

and treatments (Islam et al., 2022). In combination with the clinician's expertise, machine learning could give an opportunity to personalise the care of women with GDM and allow for prompt effective therapeutic recommendations, particularly within the remit of predicting pharmacological therapy.

Previous systematic reviews have identified risk factors for the need for pharmacological therapy to achieve good glycaemic control, but have not looked at the modelling algorithms or techniques used in detail (Alvarez-Silvares et al., 2022; Benham et al., 2023a). Here, we investigated the algorithms and their performance rather than just identifying risk factors.

This scoping review aimed to comprehensively review the methods, variables and quality in machine learning used to predict the need for prescribing or the escalation of pharmacological therapy in GDM, both at the start of or throughout pregnancy.

3.4.2 Methods

The scoping review was conducted using the Joanna Briggs Institute (JBI) checklist (Peters et al., 2020), and the Preferred Reporting Items for Systematic Reviews and Meta-Analyses – Scoping Review (PRISMA-ScR) checklist (Tricco et al., 2018) and quality was assessed using the Prediction model Risk Of Bias ASsessment Tool (PROBAST) (Wolff et al., 2019). The protocol for the scoping review was registered on the Open Science Framework website, using the JBI protocol template (osf.io/24dhk, **Appendix C**).

3.4.2.1 Search strategy

Four electronic databases: Embase, Medline, IEEE Xplore, and Web of Science, were searched for published literature (**Appendix B**). Electronic databases were chosen, and search terms were developed and refined in consultation with the study team and research librarian.

3.4.2.2 Inclusion and exclusion criteria

Studies published between 1st July 2007 and 31st August 2024 were included. The start date was chosen to be after the reporting of the Metformin in Gestational Diabetes trial (Rowan, 2007), after which Metformin was more widely used for GDM. Inclusion and exclusion criteria are shown in **Table 3.2**, using the JBI (Peters et al., 2020) population, concept, and context (PCC) framework. We included papers that had a population of women with GDM defined by the author and predicted pharmacological therapy for GDM using machine learning. As logistic regression could be classified as either a statistical or machine learning model, studies that reported predominantly statistical metrics were excluded.

Table 3.2 *Inclusion and exclusion criteria using the Population Concept Context framework*

	Inclusion criteria	**Exclusion criteria**
Population	Pregnant women who have GDM[a]	Non-pregnant participants Pregnant women without GDM[a] Women with T1DM[b] or T2DM[c]
Concept	Predicting pharmacological therapy for GDM[a] using machine learning	Predicting pharmacological therapy for GDM[a] without using machine learning, or not predicting pharmacological therapy
Context	All patient groups and settings	
Types of studies	Any	Conference papers Review papers Editorial Commentary Letters Essays Books and book chapters

[a]*GDM: Gestational diabetes mellitus*
[b]*T1DM: Type 1 diabetes mellitus*
[c]*T2DM: Type 2 diabetes mellitus*

3.4.2.3 Screening and selection of studies

Using the search terms, 6,917 studies were identified from the four databases and imported into Covidence. Two additional papers were added from the references of included papers, giving a total of 6,919 studies identified. 803 were duplicates, leaving 6,116 to be screened by their title and abstract. Screening and selection of the studies are shown in the PRISMA diagram in **Figure 3.1**. Title and abstract screening were completed by the first reviewer (JRK), and 9% (n=531) were completed by the second reviewer (NG). 6,080 studies were excluded, and 36 studies had a full-text review. The full-text review was completed by JRK, and NG completed 11% (n=4). Following the full-text review, nineteen studies were excluded, either for not using machine learning (n=14), not predicting pharmacological therapy (n=4) or being a conference paper (n=1). This resulted in 17 studies being included in this scoping review. Disagreements between reviewers were resolved through discussion.

3.4.2.4 Charting the data

Data were extracted from the 17 included studies in Covidence and then exported and analysed using Microsoft Excel and R. The extracted data included general information on the study characteristics (title, author, country, research aims, study design, length of study, and population), treatment predicted, GDM diagnosis criteria, variables, algorithms, performance and validation of the models. Studies' quality was assessed using the PROBAST checklist, focusing on risk of bias and applicability within four domains: participants, predictors, outcomes and analysis.

For analysis, the models were grouped into two treatment prediction categories. These were (1) predicting pharmacological therapy, all of which were a change from diet, and (2) predicting insulin, which was a change from either diet or oral agents.

Due to a large proportion of the models (45.5%, 20/44) coming from Liao et al. (2022) a sensitivity analysis was performed, excluding this study.

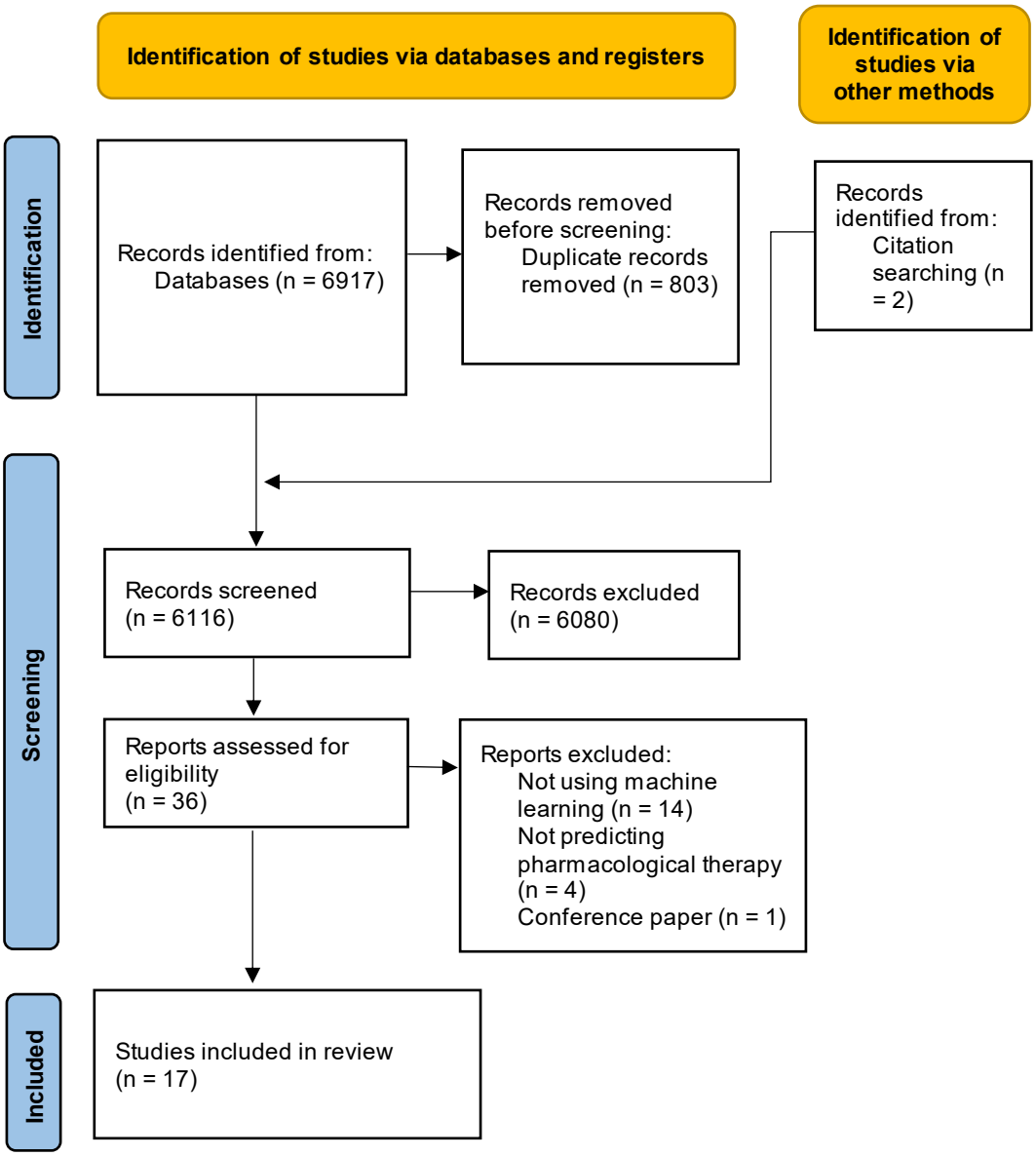


Figure 3.1 *PRISMA-ScR flowchart*

3.4.3 Results

3.4.3.1 Overview

A total of 17 studies were included, as shown in the PRISMA flowchart (**Figure 3.1**), which were published between 2016 and 2023. An overview of the characteristics of the included studies is shown in **Table 3.3** (a detailed summary of the included studies' characteristics is provided in **Appendix B**). Published studies included participants located in Europe (25.3%, 6/17) (Yerlikaya et al., 2018; Ducarme et al., 2019; Weschenfelder et al., 2021; Eleftheriades et al., 2021; Khin et al., 2018; Velardo et al., 2021); Asia (23.4%, 4/17) (Tang et al., 2019; Watanabe et al., 2016; Nishikawa et al., 2018; Tamagawa et al., 2021); North America (17.6%, 3/17) (Harper et al., 2016; Feghali et al., 2019; Liao et al., 2022); Australia (11.8%, 2/17) (Barnes et al., 2016; Ford et al., 2022) and South America (Souza et al., 2019; Zaccara et al., 2023). One study was a secondary analysis of a prospective observational study (5.9%, 1/17) (Ducarme et al., 2019) and the rest were cohort studies, either retrospective (76.5%, 13/17) (Ford et al., 2022; Feghali et al., 2019; Harper et al., 2016; Khin et al., 2018; Nishikawa et al., 2018; Souza et al., 2019; Tamagawa et al., 2021; Tang et al., 2019; Velardo et al., 2021; Watanabe et al., 2016; Weschenfelder et al., 2021; Yerlikaya et al., 2018; Zaccara et al., 2023); prospective (2/17, 11.8%) (Barnes et al., 2016; Eleftheriades et al., 2021) and population-based (5.9%, 1/17) (Liao et al., 2022).

Table 3.3 *Summary of study characteristics*

	Author	Country	Type of prediction study	GDM[a] diagnostic criteria	Data collection period (years)	Number of participants	Algorithm (s) described in the paper	Number of models described in the paper
Predicting pharmacological therapy	Feghali et al. 2019	United States	Development and validation	Carpenter and Coustan's criteria	3	1174	Logistic regression	2
	Liao et al. 2022	United States	Development and validation	Carpenter and Coustan's criteria	10	30474	CART[e] , LASSO[f], Simple super learner[g], Complex super learner[h] , Logistic regression	20
	Velardo et al. 2021	United Kingdom	Development and validation	IADSPG[c] and national guidelines	5	1789	Logistic regression	1
	Yerlikaya et al. 2018	Austria	Development	IADSPG[c]	2	203	Logistic regression, Random Forest	4
Predicting insulin	Barnes et al. 2016	Australia	Development and validation	ADIPS[b]	23	3317	Logistic regression	1
	Ducarme et al. 2019	France	Development and validation	IADSPG[c] and national guidelines	1	200	Logistic regression	1
	Eleftheriades et al. 2021	Greece	Development and validation	IADSPG[c]	8	775	CART[e]	1
	Ford et al. 2022	Australia	Development and validation	ADIPS[b]	1	2048	Logistic regression	1
	Harper et al. 2016	United States	Development and validation	Carpenter and Coustan criteria	6	360	Logistic regression	2
	Khin et al. 2018	United Kingdom	Development	IADSPG[c]	3	228	Logistic regression	1
	Nishikawa et al. 2018	Japan	Development	IADSPG[c]	1	529	Logistic regression	1

	Author	Country	Type of prediction study	GDM[a] diagnostic criteria	Data collection period (years)	Number of participants	Algorithm (s) described in the paper	Number of models described in the paper
	Souza et al. 2019	Brazil	Development and validation	IADSPG[c]	3	408	Logistic regression	1
	Tamagawa et al. 2021	Japan	Development	IADSPG[c] and national guidelines	9	388	Logistic regression	1
	Tang et al. 2019	China	Development	IADSPG[c] and national guidelines	3	534	Logistic regression	1
	Watanabe et al. 2016	Japan	Development	IADSPG[c]	6	37	Logistic regression	1
	Weschenfelder et al. 2021	Germany	Development	IADSPG[c] and national guidelines	5	454	Logistic regression	4
	Zaccara et al. 2023	Brazil	Development	ADA[d]	8	869	Logistic regression	1

[a]*GDM Gestational diabetes mellitus*
[b]*ADIPS Australasian Diabetes in Pregnancy Society*
[c]*IADPSG International Association of the Diabetes and Pregnancy Study Group*
[d]*ADA American Diabetes Association*
[e]*CART Classification and regression tree*
[f]*LASSO Least absolute shrinkage and selection operator*
[g]*Simple super learner could have been included response-mean, least absolute shrinkage and selection operator regression, and classification and regression tree*
[h]*Complex super learner could have been response-mean, least absolute shrinkage and selection operator regression, classification and regression tree, random forest, and extreme gradient boosting*

GDM was diagnosed with different diagnostic criteria, shown in **Figure 3.2** alongside the country of the included studies. Mostly GDM was diagnosed mostly with the International Association of the Diabetes and Pregnancy Study Groups (IADPSG) or IADPSG in combination with national guidelines, for example, the National Institute for Health and Care Excellence (NICE), (64.7%, 11/17) (Eleftheriades et al., 2021; Khin et al., 2018; Nishikawa et al., 2018; Souza et al., 2019; Watanabe et al., 2016; Yerlikaya et al., 2018; Ducarme et al., 2019; Tamagawa et al., 2021; Tang et al., 2019; Weschenfelder et al., 2021; Velardo et al., 2021), followed by the Carpenter and Coustan criteria (17.6%, 3/17) (Feghali et al., 2019; Harper et al., 2016; Liao et al., 2022), Australasian Diabetes in Pregnancy Society (ADIPS), (11.8%, 2/17) (Barnes et al., 2016; Ford et al., 2022), American Diabetes Association (ADA) (5.9%, 1/17) (Zaccara et al., 2023).

From the 17 included studies, 44 models were described. Pharmacological therapy, either oral agents or insulin, was predicted by 61.4% (27/44) of the models (Feghali et al., 2019; Liao et al., 2022; Velardo et al., 2021; Yerlikaya et al., 2018), of which 74.1% (20/27) were from Liao et al. (Liao et al., 2022). The remaining 38.6% (17/44) predicted insulin, as a change from diet (Barnes et al., 2016; Ducarme et al., 2019; Eleftheriades et al., 2021; Ford et al., 2022; Nishikawa et al., 2018; Souza et al., 2019; Tamagawa et al., 2021; Tang et al., 2019; Weschenfelder et al., 2021; Watanabe et al., 2016; Zaccara et al., 2023), glyburide (Harper et al., 2016) or metformin (Khin et al., 2018).

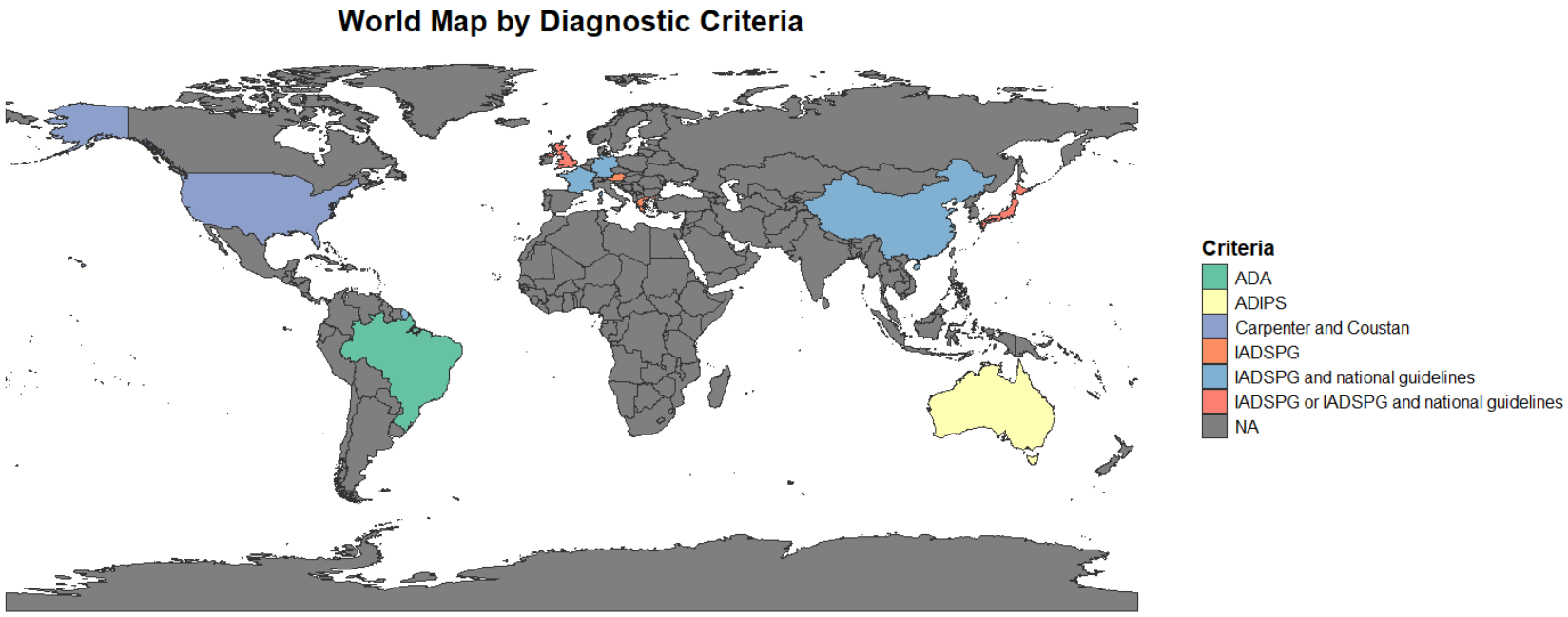


Figure 3.2 *GDM diagnostic criteria and country of the included studies.*

3.4.3.2 Study characteristics and population

The median duration for data collection was 5 years (range 1-23 years). There was a median per model of 1919 participants (range 37-30,474), with a median of 61.2% (range 30.2-89.2%) of study participants per model in the control groups and 38.8% (range 10.8-69.8%) in the prediction group. Only one study explicitly reported a sufficient sample size (Weschenfelder et al., 2021). However, using the commonly used calculation of one variable for ten adverse outcomes (Riley et al., 2020), it was found that all but one model (Watanabe et al., 2016) had a sufficient sample size.

Ethnicity was reported in 41.2% (7/17) (Liao et al., 2022; Velardo et al., 2021; Khin et al., 2018; Harper et al., 2016; Ford et al., 2022; Feghali et al., 2019; Barnes et al., 2016) studies. Common exclusion criteria included multi-fetal pregnancies (58.8%, 10/17) (Barnes et al., 2016; Ducarme et al., 2019; Eleftheriades et al., 2021; Ford et al., 2022; Khin et al., 2018; Souza et al., 2019; Tamagawa et al., 2021; Weschenfelder et al., 2021; Zaccara et al., 2023); having other types of diabetes (58.8%, 10/17) (Ducarme et al., 2019; Eleftheriades et al., 2021; Ford et al., 2022; Liao et al., 2022; Tamagawa et al., 2021; Tang et al., 2019; Velardo et al., 2021; Watanabe et al., 2016; Zaccara et al., 2023); and missing data (47.1%, 8/17) (Barnes et al., 2016; Zaccara et al., 2023; Yerlikaya et al., 2018; Weschenfelder et al., 2021; Velardo et al., 2021; Souza et al., 2019; Khin et al., 2018; Ducarme et al., 2019).

3.4.3.3 Clinical implementation

None of the identified studies had implementation in a clinical setting. One study developed a nomogram for clinical interpretation (Souza et al., 2019), and another (Velardo et al., 2021), indicated that it could potentially be integrated into their existing mHealth app.

3.4.3.4 Algorithm

Overall, the most used algorithm was logistic regression (59.1%, 26/44) (Liao et al., 2022; Barnes et al., 2016; Harper et al., 2016; Watanabe et al., 2016; Khin et al., 2018; Nishikawa et al., 2018; Yerlikaya et al., 2018; Ducarme et al.,

2019; Feghali et al., 2019; Souza et al., 2019; Tang et al., 2019; Tamagawa et al., 2021; Velardo et al., 2021; Weschenfelder et al., 2021; Ford et al., 2022; Zaccara et al., 2023), and this was also true for both prediction categories. Following was the classification and regression tree (CART), (11.4%, 5/44) (Eleftheriades et al., 2021; Liao et al., 2022), least absolute shrinkage and selection operator (LASSO) (9.1%, 4/44) (Liao et al., 2022), simple super learner (either response-mean, LASSO, and CART) (9.1%, 4/44) (Liao et al., 2022), complex super learner (either response-mean, LASSO, CART, random forest or extreme gradient boosting) (9.1%, 4/44) (Liao et al., 2022) and random forest (9.1%, 4/44) (Yerlikaya et al., 2018).

3.4.3.5 Variables

The most frequently used variables in the models overall were history of GDM (47.7%, 21/44), gestational week at GDM diagnosis (45.5%, 20/44), pregestational BMI (40.9%, 18/44), and maternal age (38.6%, 17/44). This was the same for predicting pharmacological therapy, history of GDM (51.9%, 14/27), gestational week at GDM diagnosis (51.9%, 14/27), pregestational BMI (48.1%, 13/27), and maternal age (41.8%, 13/27). However, for predicting insulin this differed, with 1hr glucose in the 75g OGTT (64.7%, 11/17), fasting glucose in the 75g OGTT (58.8%, 10/17), history of GDM (41.2%, 7/17) and HbA1c at GDM diagnosis (41.2%, 7/17), being the most frequently used model variables.

3.4.3.6 Performance metrics

The models were evaluated through several different methods, including the area under the receiver operating curve (AUROC) (95.5%, 42/44) (Eleftheriades et al., 2021; Ford et al., 2022; Liao et al., 2022; Velardo et al., 2021; Yerlikaya et al., 2018; Zaccara et al., 2023; Harper et al., 2016; Nishikawa et al., 2018; Tang et al., 2019; Watanabe et al., 2016; Barnes et al., 2016; Ducarme et al., 2019; Feghali et al., 2019; Tamagawa et al., 2021; Weschenfelder et al., 2021), sensitivity and specificity (36.4%, 16/44) (Weschenfelder et al., 2021; Souza et al., 2019; Khin et al., 2018; Tamagawa et al., 2021; Feghali et al., 2019; Ducarme et al., 2019; Barnes et al., 2016; Watanabe et al., 2016; Tang et al.,

2019; Nishikawa et al., 2018; Harper et al., 2016) and positive predictive values (PPV) and negative predictive value (NPV) (25.0%, 11/44) (Barnes et al., 2016; Ducarme et al., 2019; Feghali et al., 2019; Tamagawa et al., 2021; Weschenfelder et al., 2021; Khin et al., 2018; Souza et al., 2019). AUROC had a median of 0.75 (range 0.61-0.93), which did not exhibit considerable changes between algorithms and prediction pathways or GDM diagnostic criteria, **Table 3.4**. It also did not seem affected by class imbalances (originally published as supplementary material, **Appendix B**).

Table 3.4 *Performance of the whole model within class imbalance. Original published as supplementary material (Appendix B).*

Percentage of participants in the predictive group, %	**AUROC[a] Median (range)**
10-20	0.76 (0.72-0.87)
21-30	0.78 (0.71-0.86)
31-40	0.76 (0.61-0.93)
41-50	0.72 (0.71-0.78)
>51	0.77 (0.70-0.83)

[a] *AUROC Area under the receiver operating curve*

3.4.3.7 Model validation

Validation of the models was conducted on 65.9% (29/44) of the models. Validation methods included temporal (Barnes et al., 2016; Liao et al., 2022), geographical (Barnes et al., 2016; Eleftheriades et al., 2021), bootstrapping (Harper et al., 2016) and cross-validation, either 10-fold (Liao et al., 2022; Eleftheriades et al., 2021; Feghali et al., 2019), 5-fold (Velardo et al., 2021; Ford et al., 2022) or leave-one-out (Souza et al., 2019). The validation performance AUROC had a median of 0.72 (range 0.59-0.82); all models performed slightly worse on the validation set, but no more than 0.12.

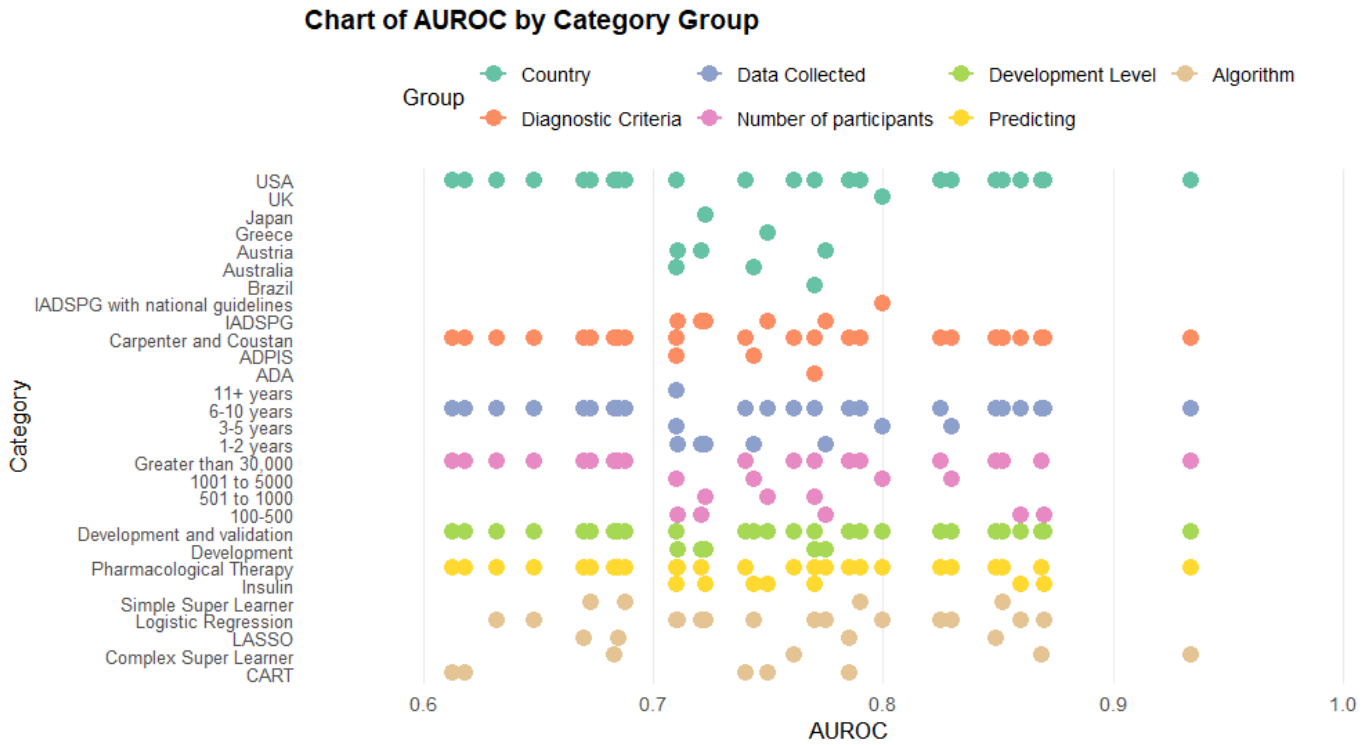


Figure 3.3 *AUROC whole model performance for categories grouped by country, GDM diagnosis criteria, length of data collection and number of participants within the study, development level of the model, the prediction of the model and the algorithm use d.*
ADA, American Diabetes Association; ADIPS, Australasian Diabetes in Pregnancy Society; AUROC, area under the receiver operating characteristics; CART, classification and regression tree; IADSPG, International Association of the Diabetes and Pregnancy Study Group; LASSO, least absolute shrinkage and selection operator; UK, United Kingdon; USA, United States of America. Simple super learner could have been included response-mean, least absolute shrinkage and selection operator regression, and classification and regression tree, Complex super learner could have been response- mean, least absolute shrinkage and selection operator regression, classification and regression tree, random forest, and extreme gradient boosting.

3.4.4 Quality assessment

The PROBAST tool was used to assess the models' risk of bias and applicability within four domains: participants, predictors, outcomes and analysis, as summary of the results in shown in **Figure 3.4** (detailed summary is provided in **Appendix B**). It was found overall that there was a predominantly high risk of bias (88.6%, 39/44, unclear: 9.1%, 4/44; low: 2.3% 1/44), and concern over applicability was varied (high: 59.1%, 26/44; unclear: 27.3% 12/44; low: 13.6% 6/44). The risk of bias between the models predicting insulin or those predicting pharmacological therapy was similar (high: 88.2% 15/17 vs 88.9%, 24/27, respectively). There was a higher concern over applicability in the pharmacological predicting group (77.8%, 21/27) compared with the insulin predicting group (29.4%, 5/17).

The high risk of bias introduced by the selection of participants, in 61.4% (27/44) of the models, was often because of the inclusion of women with pre-diabetes or the use of SMBG data. There was a low risk of bias (97.7%, 43/44) introduced by the predictors, which were defined, assessed and available appropriately for the models, which led to a low concern (95.5%, 42/44) as well. All had a low risk of bias regarding the outcomes of the models; however, there were concerns about the outcomes (low: 65.9%, 292/44; unclear: 29.5%, 13/44), mostly due to the data used not matching the outcome described. Finally, half of the models' analysis had a low risk of bias (high: 38.6%, 17/22; unclear 11.4%, 5/44; low: 50.0%, 22/44), high and unclear risk of bias was due to unbalanced or small datasets, unclear approaches to missing data or not appropriately assessing the models.

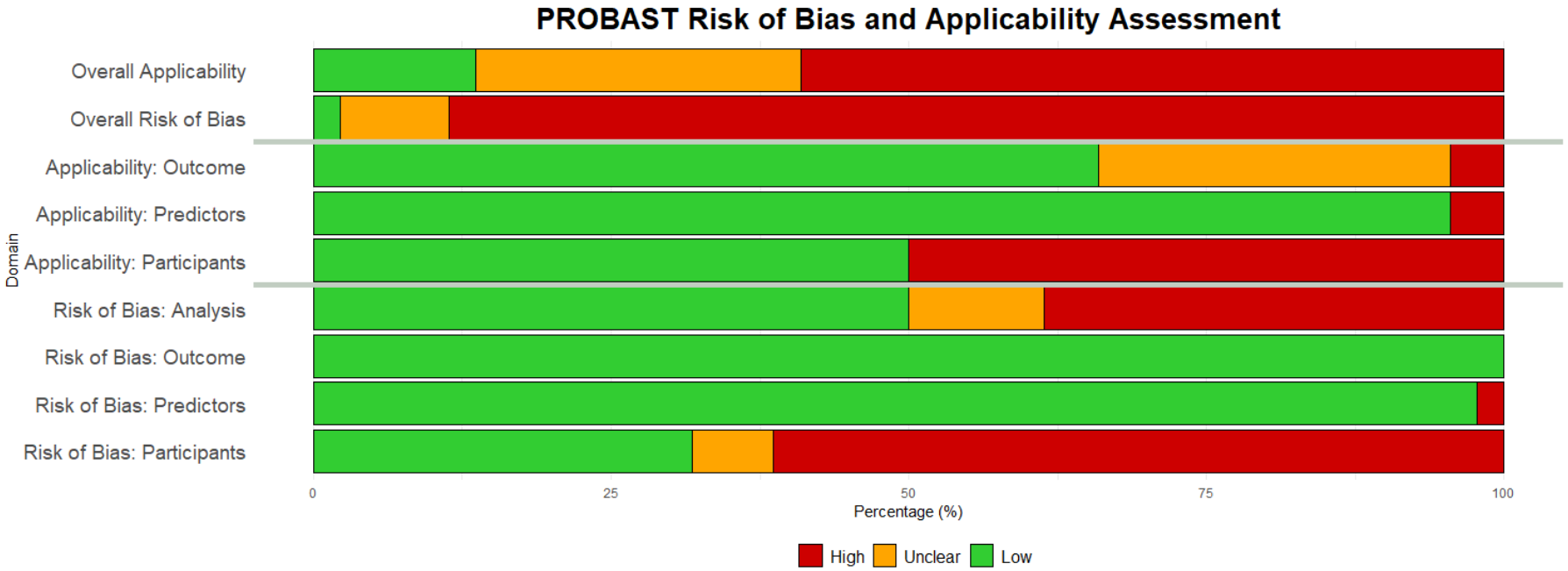


Figure 3.4 *Risk of bias summary using PROBAST. The review authors' judgements on each risk of bias and applicability domains for all included studies in the analysis*

3.4.5 Sensitivity analysis

Liao et al. (2022), presented 20 models, using five different algorithms, (logistic regression, CART, LASSO, simple super learner, complex super learner), four levels of data, (1) 1-year preconception to last menstrual period; (2) last menstrual period to before diagnosis of GDM; (3) at the time of diagnosis of GDM; (4) 1 week after diagnosis of GDM (including SMBG data), 30,474 participants and 25 variables. As this skewed the results, a sensitivity analysis was conducted where Liao et al. (2022) was removed from the results (**Appendix B**).

After the removal of Liao et al. (2022), the percentage of models predicting insulin increased (70.8%, 17/24) and the prediction of pharmacological therapy decreased (29.2%, 7/24). The median study period of 5 years (range 1-23) did not change. There was a reduction in the number of participants per model from 1919 participants to 304 (range 37-2217), and an increase in the percentage of participants in the control group from 61.2% to 65.1% (range 30.2-89.2) and a slight decrease in the prediction group from 38.8% to 35.0% (range 10.8-69.8%).

Liao et al. (2022) was the only study to use LASSO, simple super learner and complex super learner; as a result, logistic regression became the most widely used (97.7%, 22/24), with CART and random forest just used once each (4.2%, 1/24).

Common predictive variables in the models were now 1hr glucose in the 75g OGTT (58.3% 14/24), fasting glucose in the 75g OGTT (54.2%, 13/24), maternal age (41.7%, 10/24) and gestational week at GDM diagnosis (41.7%, 10/24). There were also changes regarding the predictive variables for pharmacological therapy, which are now maternal age (85.7%, 6/7), gestational week at GDM diagnosis (57.1%, 4/7), pregestational BMI (57.1%, 4/7) and parity (57.1%, 4/7).

AUROC was used to assess the performance of 91.7% (22/24) of the models, which had a minor decrease in the median from 0.75 to 0.74 (range 0.70-0.87). There was a large reduction in validated studies from 65.9% to 37.5% (9/24).

The overall risk of bias followed a similar pattern (high: 79.2%, 19/24; unclear: 16.7%, 4/24, low: 4.2%, 1/24). However, the concerns over applicability became more unclear (high: 25.0%, 6/24; unclear: 50.0%, 12/24; low: 25.0%, 6/24).

Overall, through the sensitivity analysis, it can be seen that Liao et al. (2022) was most influential in increasing the participant numbers, adding additional algorithms and altering predictive variables.

3.4.6 Discussion

This review has comprehensively investigated the use of machine learning algorithms, variables, and performance in predicting pharmacological therapy in GDM, and assessed their quality. There was a total of 17 studies that described 44 models included. All the models were binary classifiers; none were multi-class. We grouped them into two groups, predicting pharmacological therapy (61.4%, 27/44) and predicting insulin (39.6%, 17/44). The studies were published between 2016-2023, from a range of countries with different demographics and diagnosis guidelines; hence, the pooled result of this review is generalisable.

Overall, common clinical variables were used in the models, such as history of GDM, gestational week at GDM diagnosis, pregestational BMI and maternal age. After sensitivity analysis, these then included fasting and 1-hour glucose values from the 75g OGTT. This has some overlap with other reviews (Alvarez-Silvares et al., 2022; Benham et al., 2023a), which identified variables and risk factors for the need for pharmacological treatment among women with GDM.

Logistic regression was the most popular applied algorithm. Logistic regression is easy to implement, interpretable, can be used on small data sets and is not computationally expensive, hence it is often used in this medical setting.

The performance of the models was mostly evaluated by AUROC, which had a median of 0.75. AUROC was the only performance metric that was used across the different algorithms: logistic regression, CART, LASSO, simple super learner, and complex super learner. Differences in data population and the use of logistic regression outweigh the other algorithms, making it challenging to determine the best-performing algorithm. Nevertheless, logistic regression and CARTs performed well and could be a good starting point for future model development. In addition, it should be considered that for this particular clinical scenario, it may be better to have an algorithm that gives more false positives than false negatives, as these would be overseen by a clinician.

It is important to acknowledge the geographic and ethnic differences in the pathophysiology and clinical presentation of GDM, which may lead to variation in modelling performance, hence the need for model validation across different GDM populations. A reasonable number of the models were validated, however few were externally validated. The validation performance was good, indicating that the models are robust and generalisable. There does, however, need to be more external validation, particularly on data using different demographics, to ensure that the models perform sustainably.

Using PROBAST (Wolff et al., 2019) to assess the quality of the models, overall, there was a high risk of bias and concerns over applicability. The issues were mainly due to women with pre-diabetes being included in the modelling, unclear descriptions of approaches to missing data and the use of unbalanced and small datasets. These are areas that could be improved in future studies.

Machine learning has been shown to provide statistically significant improvements in health outcomes when incorporated into digital health interventions within real-life studies (Triantafyllidis and Tsanas, 2019). In a review by Sahota et al. (2011), of 36 randomised control trials of clinical decision support systems, 63% (22/35) of the studies showed an improvement in care; they found, however, that the improvements were in processes of care

rather than patient outcomes. Equally importantly, they found no significant reduction in major patient morbidity or mortality, demonstrating that when machine learning is incorporated into a care system, it is unlikely to cause harm to patients. GDM care could be personalised through the incorporation of a pharmacological therapy prediction model, which could also help streamline care. For example, it could reduce the number of follow-up appointments for women who have been identified as low risk for pharmacological therapy and allow for earlier, more targeted intervention for high-risk women. As predictions are not perfect, a model should be implemented alongside a monitoring system to ensure the safety of all patients. Future work needs to be done on the implementation of such predictive models in care.

3.4.7 Limitations

The heterogeneity of the included studies, including the different GDM diagnostic criteria, inclusion and exclusion criteria and datasets available for each model, limited the ability to make directional comparisons. These factors also contributed to simplified data extraction, and as such, the full complexities of the models may not be captured. Furthermore, due to the nature of a review, other relevant literature may have been published since the search.

3.4.8 Conclusion

The use of machine learning to predict pharmacological intervention in GDM could be easily implemented in clinics to risk-stratify patients and therefore personalise care and allocate resources more appropriately. From this review, it was found that a popular approach was logistic regression that had a median AUROC of 0.75 and used clinically available variables such as history of GDM, gestational week at GDM diagnosis, pregestational BMI, maternal age, HbA1c, fasting – and 1hr- glucose in the 75g OGTT. There were no multi-class models presented. Furthermore, there was a lack of external validation, which future models would benefit by incorporating.

3.4.9 Funding

JRK is funded by Medical Research Scotland (PHD-50224-2020). RMR acknowledges the support of the British Heart Foundation (RE/18/5/34216).

3.5 CHAPTER SUMMARY

This chapter presented a scoping review of 17 studies describing 44 machine learning models predicting pharmacological therapy in GDM. Models were binary classifiers, either predicting oral agents and insulin together (n=27, 61.4%) or insulin alone (n=17, 38.6%). The pooled AUROC was 0.75, with logistic regression being the most frequently used algorithm (n=26, 59.1%).

Studies had a median duration of 5 years (range 1-23 years), and a median of 1919 (range 37-30,474) participants for each model, with 64.7% of the studies diagnosing GDM using IADPSG.

Clinically available variables were used in the prediction models, including maternal characteristics and OGTT or GCT results. These results guided the selection of variables requested to develop my pharmacological prediction models, as described in the chapters 4 and 6.

Validation of 65.9% (n=29) of the models was reported. Despite 50% (n=22) of models being externally validated, only three (17.6%) studies conducted external validation. This highlights that while external validation was performed on several models, it was only done in a small number of studies, raising concerns about the generalisability of most of the models reported in the literature.

The quality of the models was assessed using PROBAST, which identified a high risk of bias and concerns about applicability. Mostly, the issues stemmed from data quality.

There was no evidence from the reviewed literature that the models had been implemented clinically or had plans for further pilot testing. This may be partly due to the lack of external validation.

CHAPTER 4

Identification and Assessment of Data Quality for a Glasgow Pregnancy Cohort

4.1 INTRODUCTION

Data cleaning involves inspecting, correcting, transforming and deleting data, ensuring that the correct methods can be applied for the analysis (Huxley et al., 2020). It is a vital step in removing inconsistencies, errors, and noise within the data (Sandfeld, 2024).

The data for this project are derived from a retrospective analysis of clinical care data; consequently, it was not created for statistical analysis and machine learning modelling. Therefore, preparing a clean dataset will increase the quality and reliability of the results.

In this chapter, using data from 30,666 pregnancies delivered in the Greater Glasgow and Clyde area between 2021 and 2023, I will describe how I selected and cleaned a cohort and identified pregnancy episodes complicated by GDM.

4.2 AIMS AND ENDPOINTS

This Chapter will focus on the first part of objective: II Identify a cohort of GDM-affected pregnancies in Glasgow. And to analyse and compare the outcomes of: (a) the cohort against treatments for GDM and (b) between GDM and non-GDM pregnancies. The specific aims and endpoints of this Chapter are:

4.2.1 Aims

1) To select and clean a cohort of singleton pregnancy episodes.
2) To identify pregnancy episodes complicated by GDM within the cohort.
3) Select a subset of pregnancy episodes complicated by GDM within the cohort that are suitable for model development.

4.2.2 Endpoints

1) A cohort of singleton pregnancies in Glasgow, cleaned such that statistical analysis can be performed.
2) A concise cohort of singleton GDM complicated pregnancies in Glasgow that is suitable for statistical analysis and developing a machine learning model.

4.3 DATA

4.3.1 Linkage and storage

Data were stored and analysed within the West of Scotland Safe Haven, which was accessed through a Virtual Private Network. Patients' data from NHS Greater Glasgow and Clyde were extracted from the electronic healthcare records and linked to be de-identified by the West of Scotland Safe Haven.

To ensure patient safety and confidentiality, the West of Scotland Safe Haven checked all outputs that are presented in this thesis and elsewhere in publications. No data were released if the pregnancy episode frequency was < 5 or could otherwise be inferred. R (version 4.3.2) was used for cleaning, and Microsoft Word (2016) for documentation.

4.3.2 Datasets

Clinical care data available from Greater Glasgow and Clyde described 30,666 pregnancy episodes with delivery dates between the 1st of January 2021 and the 31st of December 2023. Variables requested for the dataset were derived from the results of the scoping review, Chapter 3, and from discussion with supervisors. Therefore, antenatal booking data before the 1st of January 2021 were also included. A pregnancy episode is defined as any pregnancy, regardless of outcome. The study data contained three linked datasets (**Table 4.1**). These were 1) Badgernet (System C Healthcare, United Kingdom, https://www.badgernotes.net/), consisting of the pregnancy and antenatal data. Badgernet is a clinical information system widely used in obstetric care. 2) Deprivation, containing a single record for each patient in the study and including the gender, age and Scottish Index of Multiple Deprivation (SIMD), (Research Data Scotland, 2024). And 3) Pharmacy, of only diabetes-related prescriptions coded by the British National Formulary, Chapter 6.1, and includes details of strength, unit, and dose. All datasets had columns of variables and rows of pregnancy episodes.

Table 4.1 *Three linked study datasets*

Dataset	Number of pregnancy episodes (rows)	Number of variables (columns)
Badgernet (maternal and antenatal data)	30,666	107
Deprivation	28,724	16
Pharmacy (British National Formulary, Chapter 6.1)	25,546	23

4.3.3 Ethics

Research ethics approval was granted for linkage to Greater Glasgow and Clyde NHS patient data by the Local Privacy and Advisory Committee at NHS Greater Glasgow and Clyde under approval GSH23DI001 (**Appendix D**).

4.4 POPULATION

The data flow is depicted in **Figure 4.1**, which describes how the cohort was developed from the study data. The exclusion criteria for the cohort are outlined below in 4.4.1.1. The criteria used to identify GDM complicated pregnancies are outlined in 4.5.2.1.

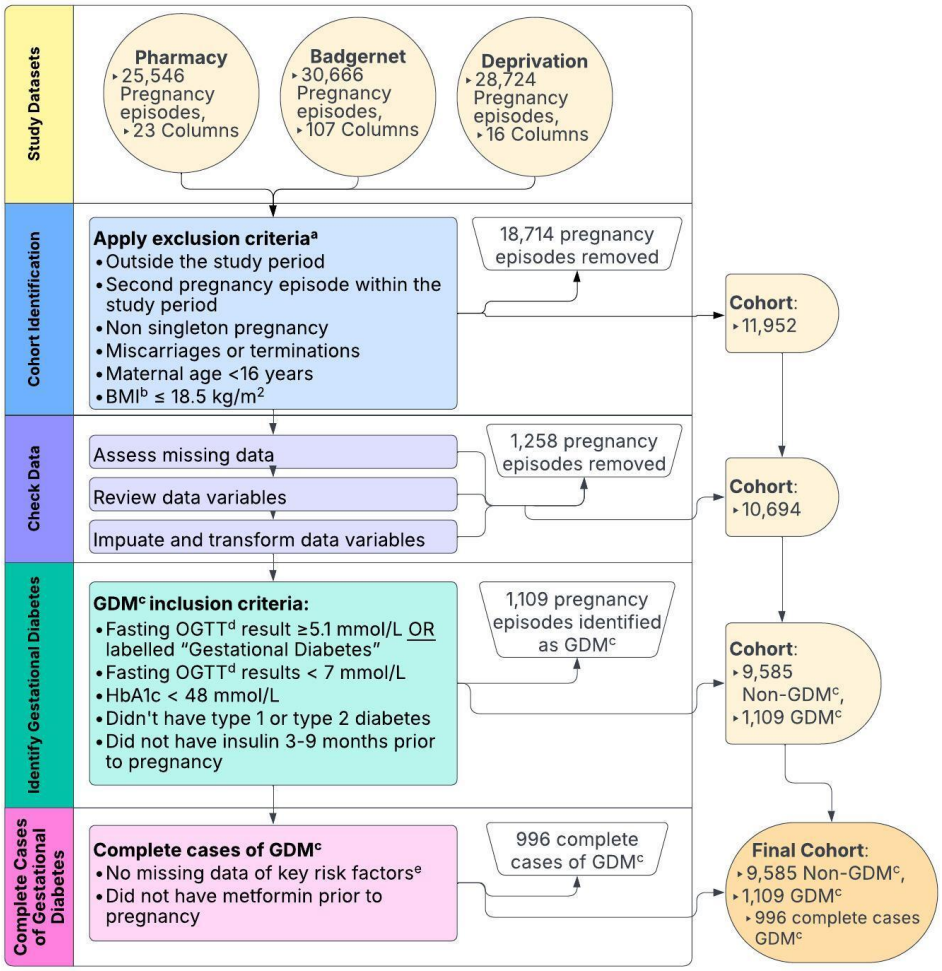


Figure 4.1 *Data flow diagram*

[a]Cohort exclusion criteria: Pregnancy episode had a booking date outside the study period (01/04/2022 – 31/12/2023), Second pregnancy episode within the study period, Multifetal pregnancy episode, Pregnancy episode ending in a miscarriage or termination, Maternal age < 16 years; Body mass index ≤ 18.5 kg/m².[b]BMI Body mass index. [c]GDM Gestational diabetes mellitus, [d]OGTT Oral glucose tolerance test. [e]Key risk factors: BMI, previous macrosomic baby, History of GDM, family history of diabetes and ethnicity.

4.4.1 Cohort identification

The cohort was selected from the study data using the following exclusion criteria:

4.4.1.1 Cohort Exclusion Criteria

1) Pregnancy episode with a booking date before 1st April 2022.
2) Multifetal pregnancy episode.
3) Pregnancy episode ending in a miscarriage or termination.
4) Maternal age < 16 years.
5) Maternal booking BMI ≤ 18.5 kg/m^2.
6) Second pregnancy episode within the study period.

Care was systematically different during the pandemic, and therefore, I excluded pregnancies that received care during the pandemic by selecting only pregnancies with a booking date after April 2022. Additionally, to allow for comparison between GDM and non-GDM complicated pregnancies, pregnancies that would require additional care were excluded. These included multifetal pregnancies, as they have an elevated risk and a different treatment compared to a singleton pregnancy. Similarly, being underweight while pregnant also comes with its own risk factors (Burnie et al., 2022); therefore, a BMI ≤ 18.5 kg/m^2 was excluded. It was not possible to identify other factors that would require additional care due to the limitations of the data. To comply with the ethics approval, miscarriages, terminations, and maternal age < 16 years were excluded. Furthermore, to ensure that the data were independent for statistical analysis, only the first pregnancy within the cohort period (01/04/22-31/12/23) was selected.

4.5 DATA CLEANING

The identification and cleaning of the cohort had three stages, shown in **Figure 4.2**, which were adapted from Sandfeld et al. (2024) Material Data Science Chapter 11, (Sandfeld, 2024). Approaches to dealing with missing data follow Gelman and Hill et al. (2006) Data Analysis Using Regression and Multilevel/Hierarchical Models, Chapter 25, (Gelman and Hill, 2006).

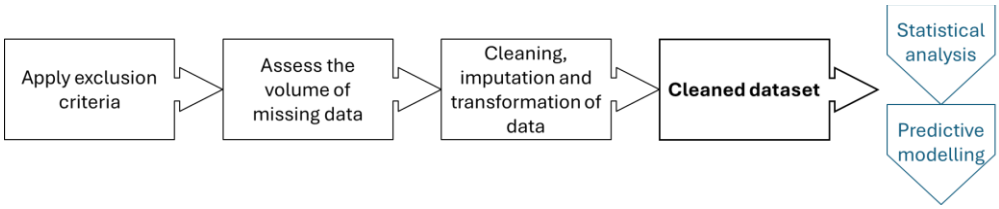


Figure 4.2 *Cohort identification and data cleaning schematic.*

4.5.1 Data cleaning process

4.5.1.1 Exploration

Firstly, the data were explored, and errors were corrected. There were several columns that contained high volumes of missingness and provided no information, which will be discussed in 4.5.1.3. Additionally, one key variable (identified in the scoping review, Chapter 3), the 2-hour OGTT result, was missing. The data also contained several structural errors, including dates, NULL and numerical variables being strings; these were converted to their appropriate structures. Similarly, spelling mistakes were corrected. To ensure that variables were within a biologically plausible range, fasting OGTT results >40 mmol/L, any HbA1c result <15 mmol/L or >200 mmol/L were replaced with NA (unknown). These thresholds were selected through discussion with clinical supervisors.

4.5.1.2 Applying exclusion criteria

The three study datasets (Badgernet, Deprivation and Pharmacy, see section 4.3.2 for more details), were merged with the exclusion criteria as described in 4.4.1.1, applied.

Of the 30,666 pregnancy episodes, 18,217 were outside the study period (01/04/2022 – 31/12/23), 124 were multifetal, 8 ended in miscarriage or termination, 21 had a maternal age <16 years, 784 had a BMI ≤18.5 kg/m^2 and 1,924 had a second pregnancy within the study period. These were excluded from the 30,666 pregnancy episodes, resulting in 11,952 pregnancy episodes that met the inclusion criteria.

4.5.1.3 Missing data

Most of the missing data were Missing Completely at Random (MCAR), that is, the data simply had not been entered, was not well recorded or recorded elsewhere. In all instances of missing data, there was no means to obtain it. MCAR data did not impact the rest of the data variables, as they were independent. I assumed that there was no bias involved during data input by HCPs. However, there are biases in the use of EHR data, which may arise from certain patient populations' access to care (Chen et al., 2021a), though this bias is likely to be negligible in Glasgow due to healthcare being free and accessible to all.

The volume of missing data for each variable (column) was assessed and grouped into four categories: ≤5%, 6-19%, 20-94% and ≥95%. For each variable, consideration on the volume of missingness, the type of missingness and how relevant it is to the overall analysis led to either column-wise (variable) deletion, row-wise (pregnancy episode) deletion or the variable being kept.

Overall, 52.8% (67/127) of the variables had some degree of missingness, ranging from 0.01 to 100%.

Table 4.2 contains the volume and action on the variables that had missing data. For some columns, (20.9%, 14/67), it was expected that we would encounter some missingness, for example, with one of the diagnostic tests for GDM it would not be expected that every pregnancy episode would have a result, only those at risk of GDM, therefore it is expected that there will be missing data for this variable. In these instances, missingness was kept. The variables and the volume of missingness from the Badgernet dataset can be seen in **Appendix E**.

Table 4.2 *Volume and action on missing variables*

Percentage of rows missing from the column	**Column-wise deletion**			**Row-wise deletion**	**Keep**		**Total columns for each missingness category (N, %)**
Reason for action:	**Multiple records**	**Not relevant to analysis**	**≥95% missing**	**Exclusion criteria**	**Maximise cohort**	**Expected missing**	
≤5%	2				12	2	**16 (23.9%)**
6-19%	6			1	3		**10 (14.9%)**
20-94%		1			3	9	**13 (19.4%)**
≥95%			25			3	**28 (41.8%)**
Total columns for each reason (N, %)	**8 (11.9%)**	**1 (1.5%)**	**25 (37.3%)**	**1 (1.5%)**	**18 (26.9%)**	**14 (20.9%)**	**-**
Total columns for each action (N, %)	**34, (50.7%)**			**1 (1.5%)**	**32 (47.8%)**		**-**

Variables that had missing data. The column describes the actions taken and reasons, and row describe the volume of the variables that were missing. Each action and volume has a total count and percentage. The percentage is out of the 67 variables that had missingness.

Of the missing variables, 50.7% (34/67) had column-wise deletion, most of which was because over 95% data for that variable were missing, and therefore it was impossible to gain any information from those variables. Other reasons were that they had a high volume of missingness (20-94%) and were not relevant to the analysis. And lastly, because there were multiple records for the same feature. For example, smoking status was recorded in five different variables, 'Smoker_InTheTwelveMonthsBeforeConception', 'EverSmoked', 'SmokingQuitDate', 'Smoker_AtBooking', 'Smoker_AtBooking_Ever'. Out of these variables, the one with the least amount of missing data was selected, and missing rows were interpreted from the other columns. To maximise the cohort and minimise bias introduced during cleaning, 26.9% (18/67) of the variables with missing data were retained.

The variable that described the ending of the pregnancy episode had 10.53% (n=1258) missing data; therefore, I could not be certain how the pregnancy had ended, and whether they met the inclusion criteria or not. Hence, row-wise deletion (pregnancy episode) was actioned on this variable. After addressing missingness in the data the cohort consisted of 10,694 pregnancy episodes.

4.5.1.4 Checking and adding additional variables

After applying the exclusion criteria and addressing the missingness in the data, the remaining variables were reviewed. Ten variables were removed as they were clearly for clinical use only (e.g. 'PregnancyStatus' was either closed or open). And a further nine variables were removed that were not relevant to the analysis (e.g. 'ExpectedNumberOfBabies', which was not needed for the analysis). Information provided by five columns was better presented elsewhere (e.g 'MEDSmetformin' came from the Badgernet data but was a bad record of metformin taken in pregnancy, in comparison to the records within the Pharmacy data, which also provided more depth, giving context regarding date of dispensation and units). And finally, five columns

were removed that had a category with a frequency < 5 to ensure that ethics were complied with.

It was necessary to determine an algorithm to identify pregnancies that were complicated by GDM as the description of GDM was poor and inconsistent within the data. Diabetes was described in six columns ('Diabetes_Display', 'Diabetes', 'DiabetesGestational', 'dtype1', 'dtype2' and 'dtype12', d stands for diabetes). Consistency between the diabetes columns was checked. Those describing type 1 and type 2 diabetes were consistent with each other; however, those describing GDM were not consistent with each other, nor were they consistent with the rest of the data. For example, 'Diabetes_Display' and 'DiabetesGestational' may indicate that there was no current GDM. However, if there was a fasting OGTT results taken between the booking and delivery dates that is above the diagnostic threshold and insulin was also dispensed during the same time period, this would strongly suggest that the pregnancy involved some form of diabetes. Given that the other diabetes related variables were consistent, it is most likely GDM. Hence, an algorithm to identify pregnancies that were complicated by GDM (discussed in section 4.5.2) was needed.

Additional variables were added; these were to either present the data in a more impactful or more concise way. Some maternal factors and pregnancy outcomes were altered. The type of Caesarean birth was pulled out of 'ModeOfDelivery'. Macrosomia variable was added, defined from 'BirthWeight_Grams' ≥ 4500g, similarly, LGA variable was added if 'BirthweightMinCentileWHO' ≥ 90th percentile. The ethnicity of the mother was described in detail in 'FamilyOrigin_Mother'; this was grouped into ethnicity groups defined by Public Health Scotland, specified in the Scottish Morbidity Record - maternity interaction/stay (SMR02) crib sheet (https://publichealthscotland.scot/media/24926/smr02_crib_270323.pdf)

For a pregnancy episode that required diabetic medication, the pharmacy data had multiple records for each time a drug was dispensed. The information I needed for a pregnancy episode in the cohort was: 1) was

medication dispensed, if yes, 2) what was it and 3) when was it first dispensed. For GDM, there are two drugs that are used, metformin or insulin (1.2.1.3); these are recorded in 'PI_Approved_Name' (drug approved name) within the Pharmacy dataset. This information was split into two columns: 'Metformin' [1/0] and 'Insulin' [1/0]. Where either of those was present (equal = 1), the first time it was dispensed was used to calculate the gestational week + day that medication was first administered during the pregnancy. Furthermore, it was recorded if insulin or metformin were dispensed 3-6 months prior to the booking date, which was then used for GDM identification.

4.5.2 Identifying GDM within the cohort

As mentioned above, in section 4.5.1.4, there were inconsistencies in the data's identification of GDM, and therefore, it was necessary to derive an algorithm to identify pregnancy episodes complicated with GDM. Out of the cohort of 10,694, 10% (1,109) were identified as GDM using the following criteria described below.

4.5.2.1 GDM inclusion criteria:

1) Had a fasting OGTT result ≥5.1 mmol/L taken between the booking date and delivery date OR had a label of "Gestational Diabetes" in 'DiabeteDisplay' variable.
2) Fasting OGTT < 7 mmol/L.
3) If present, HbA1c earliest, HbA1c at booking or HbA1c near booking < 48 mmol/L taken between the booking date and delivery date.
4) Was not labelled as type 1 or type 2 diabetes.
5) Did not have insulin dispensed 3-9 months prior to booking date.

Thresholds were selected in line with the SIGN 116 guidelines (Scottish Intercollegiate Guidelines Network, 2010) which were in place at the time of the study. To ensure that the women did not have type 1 diabetes, I excluded any pregnancies that had a record of insulin dispensed 3-9 months prior to their booking date.

There were 1,219 pregnancy episodes that met the first criterion, having a fasting glucose in the OGTT ≥5.1 mmol/L taken between the booking date and delivery date OR had a label of "Gestational Diabetes" in 'DiabeteDisplay' variable. Following, 110 pregnancy episodes did not meet the other inclusion criteria, leading to 1,109 pregnancy episodes being identified as GDM.

4.5.2.2 Complete cases of GDM

The primary goal of the data cleaning was to create a functional dataset for machine learning. Therefore, to create a dataset for machine learning, I chose a 'Complete-Case Analysis' (Gelman and Hill, 2006). A quite brutal approach to data cleaning, in which rows (pregnancy episodes) are removed if they do not meet the inclusion criteria. Pregnancy episodes that were missing risk factors described by SIGN 116 (Scottish Intercollegiate Guidelines Network 2010) were then removed to further ensure that the pregnancies were complicated with GDM.

Risk factors for GDM (SIGN 116, Table 4):

- BMI
- Previous macrosomic baby
- History of GDM
- Family history of diabetes
- Ethnicity

The only risk factor with missing inputs was ethnicity. I decided not to impute the missing variables, as ethnicity was the only variable with missingness and would be difficult to impute accurately without adding noise. The machine learning model is going to predict medication risk during GDM (Chapter 6); therefore, it was necessary to remove pregnancies that had metformin dispensed 3-9 months prior to booking data. It is acknowledged that this may exclude those on metformin treatment for polycystic ovary syndrome.

Having removed pregnancies that had missing key variables or had metformin dispensed prior to the booking date resulted in 996 complete cases of GDM,

which will be referred to as the modelling dataset, and was used to develop a machine learning model (Chapter 6).

4.6 CHAPTER SUMMARY

In this chapter, I described the process of identifying and cleaning a cohort of 10,694 singleton pregnancy episodes with a booking date between 1st April 2022 and 31st December 2023 from three linked study datasets of 30,666 pregnancies with delivery dates between 1st January 2021 and 31st December 2023 of five hospitals in the Greater Glasgow and Clyde area. Furthermore, within the cohort, I identified 1,109 pregnancies that were complicated by GDM. The cohort was cleaned to a level such that it was suitable for statistical analysis shown in Chapter 5. Applying a complete-case analysis to the pregnancies in the cohort that were complicated by GDM; 996 were selected for machine learning modelling, presented in Chapter 6.

A description of the cohort's variables, data type and volume of missing data is included in **Appendix F**.

CHAPTER 5

Descriptive Analysis of the Cohort's Maternal and Neonatal Characteristics and Outcomes

5.1 INTRODUCTION

Statistical analysis is a vital part of understanding the data and giving an insight into a population (Devore and Berk, 2012). This chapter will provide an understanding of how the cohort, described in Chapter 4, compares to national statistics, indicating how representative the cohort is of the general population. It will also offer an overview of the maternal characteristics and pregnancy outcomes of GDM and non-GDM complicated pregnancies, as well as compare the pregnancy outcomes based on treatment for GDM.

5.2 AIMS AND ENDPOINTS

This Chapter will focus on the second part of objective: II Identify a cohort of GDM-affected pregnancies in Glasgow. And to analyse and compare the outcomes of: (a) the cohort against treatments for GDM and (b) between GDM and non-GDM pregnancies. The specific aims and endpoints of this Chapter are:

5.2.1 Aims

1) To compare the cohort to the National statistics form Public Health Scotland (PHS)
2) To give an overview and statistical analysis of maternal characteristics and pregnancy outcomes between:
    a. GDM and non-GDM complicated pregnancies.
    b. GDM complicated pregnancies treated with diet, metformin or insulin.

5.2.2 Endpoints

1) A basic understanding of how representative the cohort is in comparison to National statistics.
2) An overview of statistically significant differences of maternal characteristics and pregnancy outcomes between GDM and non-GDM complicated pregnancies, and between different treatments for GDM.

5.3 METHODS

This is a retrospective analysis of clinical care data of singleton pregnancies in Greater Glasgow and Clyde with a booking date between 1st April 2022 to 31st December 2023, which met my inclusion criteria described in 4.4.1.1. Pregnancies complicated with GDM were identified if they met the criteria described in 4.5.2.1. Treatment for GDM was assumed using logic outlined in 4.5.2. Research ethics approval was granted for the analysis of Greater Glasgow and Clyde NHS patient data by the Local Privacy and Advisory Committee at NHS Greater Glasgow and Clyde under approval GSH23DI001 (**Appendix D**). All analysis was performed in R (version 4.3.2) and documented using Microsoft (2016) Word and Excel.

5.3.1 Public Health Scotland statistics

Reports on pregnancies and childbirth in Scotland are available from PHS (Public Health Scotland, 2024b). PHS mainly uses the SMR02 records to generate its statistics. SMR02 records are submitted by the maternity hospital to PHS whenever a woman is discharged from an episode of day case or in

patient maternity care. PHS then compares the SMR02 with the number of births registered by the National Records of Births Scotland to ensure completeness.

PHS runs analysis from 1st April to 31st March of each year. As my cohort runs from 1st April 2022 to 31st December 2023, the overlapping year for comparison was 1st April 2022 to 31st March 2023. Therefore, this year 22/23, was selected from the cohort. From the PHS published datasets, I selected the year 22/23, Greater Glasgow and Clyde and singleton pregnancies where possible, for basic comparative analysis.

5.3.2 Descriptive and statistical analysis

5.3.2.1 Descriptive analysis

Descriptive analysis of continuous variables included the mean, median, and standard deviation and boxplots including the mean, median, lower and upper quartiles. For categorical variables, descriptive analysis included the frequency and histograms.

5.3.2.2 Statistical analysis

Statistical tests were selected following the methodology described in Modern Mathematical Statistics with Applications by Devore et al., 3rd edition (Devore et al., 2021). A p-value of <0.05 was considered statistically significant.

5.3.2.2.1 Statistical analysis between two groups

For the statistical analysis between two groups (e.g. GDM vs non-GDM), categorical variables that had a sample size of less than 20 or a cell frequency less than 5, Fisher's Exact test was performed; otherwise, Pearson's Chi-Square test was performed.

For continuous variables, the distribution and variance were first assessed to ensure that the conditions for the chosen statistical test were satisfied. Normality was evaluated using the Shapiro-Wilk test or Q-Q (Quantile-Quantile) plot when the sample size was greater than 5000, and variance was evaluated with the Levene test. If the data exhibited normal distribution and

equal variance, the Student's t-test was applied. If the data showed normal distribution but unequal variance, Welch's t-test was used. If the data were not normally distributed, the Wilcoxon test (also known as the Mann-Whitney U-test) was performed.

5.3.2.2.2 Statistical analysis between three groups

For the statistical analysis between three groups (e.g. diet vs metformin vs insulin-treated GDM), the selection for statistical tests for categorical variables was the same as between two groups 5.3.2.2.1.

Conditions for statistical tests of continuous variables were assessed. The Shapiro-Wilk test was used for the normality of the variables, and for the variances, the Levene test was used. If the variable had a normal distribution and equal variance, then the One-Way Analysis of Variance (ANOVA) test was used. Where there was a normal distribution but unequal variance, then the Welch-ANOVA was performed. When the distribution was not normal, the Kruskal-Wallis Test was performed.

5.4 RESULTS

The cohort included 10,694 singleton pregnancy episodes with a booking date between 1st April 2022 to 31st December 2023 in Greater Glasgow and Clyde. Despite the higher BMI screening threshold of ≥35 kg/m^{2,} there remained a high prevalence of GDM, with 10.37% (1,109) of pregnancy episodes affected. Of those that were complicated by GDM, 72.32% (802) were managed with diet alone, 19.84% (220) with metformin and 7.84% (87) with insulin.

In this section, I will initially assess how representative the cohort identified in Chapter 4 is by comparing it to National statistics from PHS. I will then present the differences between pregnancies complicated by GDM and those that were not, as well as the variations in final treatment among those with GDM. To achieve this, I will first present maternal characteristics, followed by GDM-specific features, and finally present maternal and neonatal outcomes. Statistics are reported as mean [SD] or percentage (count) in the order of GDM

vs non-GDM or Diet vs Metformin vs Insulin. The counts of care location and ethnicity have been rounded to the nearest 10 to reduce the risk of disclosure and help maintain confidentiality.

5.4.1 Comparison to PHS births in Scotland

Greater Glasgow and Clyde managed 28.8% of all maternities in Scotland, between 1st April 2022 to 31st March 2023 (Public Health Scotland, 2024a). In comparison to the rest of Scotland, Greater Glasgow and Clyde has higher levels of ethnic minorities, deprivation (Scottish Index of Multiple Deprivation (SIMD) quintile 1) and rates of maternal diabetes (Public Health Scotland, 2024a).

To identify how representative my GDM cohort is of the Greater Glasgow and Clyde population, I compare the frequency of maternal characteristics and neonatal outcomes with published statistics by PHS Births in Scotland (Public Health Scotland, 2024b). This is presented in **Table 5.1**. In this section, when describing the PHS published data of maternities in Greater Glasgow and Clyde in the year 22/23, I will refer to it as PHS.

For the maternal characteristics, PHS reported the number of maternities (12,803) per year, which included both singleton and multiple pregnancies as well as multiple maternal episodes within the year; therefore, this comparison is very general, as my study cohort has excluded multiple pregnancies and the second pregnancy within the selected year. PHS reported 12,564 live singleton births, which is higher than the sample size in the cohort of 9,927; this is likely due to the cohort excluding underweight or the second pregnancies within the study period. While underweight women were excluded, the remaining BMI categories had a similar distribution between the cohort and PHS (PHS (all maternities): Healthy 40.4% (5119), Overweight 28.6% (3621), Obese 28.2% (3574), Unknown 1.1% (135), Cohort: Healthy 41% (3995), Overweight 29% (2845), Obese 31% (3004), Unknown 0.83% (83)). In addition, there was a similar age distribution between the cohort and PHS (PHS (all maternities): <20 years 2.2% (277), 20-24 years 11.0% (1409), 25-29 years 26.1% (3339), 30-34 years 34.8% (4460), 35-39 years 20.9% (2678), 40+ years 5.0% (640),

Cohort: <20 years 2% (206), 20-24 years 10% (1004), 25-29 years 25% (2457), 30-34 years 36% (3545), 35-39 years 22% (2186), 40+ years 5% (529)). There was a higher level of 'unknown' SIMD, and fewer SIMD of 2, 3, and 4 quantiles in the cohort than in PHS, but similar levels of the extremes, 1 (most deprived) and 5 (least deprived), (PHS (all maternities): SIMD 1 37.4% (4791), SIMD 2 18.1% (2323), SIMD 3 13.8% (1768), SIMD 4 15.2% (1942), SIMD 5 15.4% (1977), Unknown 1.1% (2), Cohort: SIMD 1 40% (3966), SIMD 2 15% (1513), SIMD 3 12% (1168), SIMD 4 12% (1186), SIMD 5 15% (1465), Unknown 6.3% (629). PHS used SIMD from 2020 (Public Health Scotland, 2024c), whereas the cohort used SIMD from 2016, which may explain the discrepancies. PHS had a higher frequency of Group A – White, Group B – Mixed or Multiple, lower Group C - Asian, Asian Scottish or Asian British and Group D – African, but similar frequency of Group F – other in comparison to the cohort, (PHS (all maternities, provisional data): A 82.6% (9194), B 2.9% (318), C 8.3% (923), D 2.9% (328), E 0.6% (64), F 2.8% (307), G 0.2% (22), H 13.0% (1663), Cohort: A 75% (7,418), B 0.4% (39), C 9.6% (951), D 5.2% (520), E 0.4% (38), F 2.5% (247), H 7.2% (714). However, PHS provided provisional data for ethnicity, and I had to group the ethnicity from the Badgernet records into the SMR02 groups, which may have influenced the differences seen.

In the cohort, diabetes was described in different ways (as explained in section 4.5.1.4). PHS diabetes data was from hard-coding and ICD-10 coding (International Classification of Diseases version 10), which I did not have access to. In the cohort, diabetes was recorded from Badgernet. However, there was a similar distribution of diabetes in both the cohort and PHS, (PHS (all maternities): No diabetes 87.1% (11,118), Yes, gestational 11.8% (1501), Yes, pre-existing 1.1% (142), Yes, status unknown 0.0% (2), Unknown0.3% (40), Cohort: No diabetes 89% (8806), Yes, gestational 10% (1024), Yes, pre-existing 1.0% (97)). The method of birth was quite different; PHS had a much higher frequency of spontaneous births and instrumental births, and fewer reported Caesarean births, both elective and emergency (PHS (live singleton births): Spontaneous 87.1% (11,118), Caesarean - Elective 11.8% (1501),

Caesarean – Emergency 1.1% (142), Instrumental 0.0% (2), Unknown 0.3% (40), Cohort: Spontaneous 45% (4510), Caesarean - 23% (2248), Caesarean – Emergency 19% (1910), Instrumental 11% (1130), Unknown 1.3% (129)). To calculate the methods of birth, I used 'ModeOfDelivery'; I had to group births that could be considered spontaneous or instrumental births, but I did not group the Caesarean births, as they were already coded. Therefore, there may be some human error in the grouping of spontaneous and instrumental, in addition to other factors within the underlying data or due to the exclusion criteria, that may have contributed to these differences. Additionally, there were fewer LGAs in the cohort than in PHS (PHS (12,934) LGA 13.1% (1684), Not LGA 86.8% (11,232), Cohort: LGA 9.5% (947), Not LGA 89% (8827)). I defined LGA in the cohort as any pregnancy episode that had a minimum WHO centile ≥ 90; this may have underestimated the true number of LGAs in the cohort. However, the frequency of preterm births (< 37 weeks of gestation) and admission to the neonatal care unit were similar in both the cohort and PHS (live singleton births): Premature (<37 Weeks) 7.2% (906), Full-term (37+ Weeks) 92.8% (11,658), Cohort: Premature (<37 Weeks) 8.5% (842), Full-term (37+ Weeks) 92% (9085)).

Overall, the cohort is a good representation of singleton pregnancies in Greater Glasgow and Clyde, especially regarding maternal characteristics. However, some birthing and neonatal outcomes might be underestimated within the cohort.

Table 5.1 *Comparison of Public Health Scotland and cohort births delivered between 1st April 2022 to 31st March 2023 in Greater Glasgow and Clyde*

	Public Health Scotland, Greater Glasgow and Clyde, %(N)	**Cohort, %(N)**
Number of maternities	12,803*[a]	
Singleton birth	12,612*[a]	-
Live births by singleton status	12,564*	9,927
Maternal age	12,803*[a]	9,927
… <20	2.2% (277)	2% (206)
… 20-24	11.0% (1409)	10% (1004)
… 25-29	26.1% (3339)	25% (2457)
… 30-34	34.8% (4460)	36% (3545)
… 35-39	20.9% (2678)	22% (2186)
… 40+	5.0% (640)	5% (529)
Scottish Index of Multiple Deprivation (SIMD)	12,803*[a,b]	9,927[b]
… 1 (most deprived)	37.4% (4791)	40% (3966)
… 2	18.1% (2323)	15% (1513)
… 3	13.8% (1768)	12% (1168)
… 4	15.2% (1942)	12% (1186)
… 5 (least deprived)	15.4% (1977)	15% (1465)
… Unknown	1.1% (2)	6.3% (629)
Ethnicity	12,819*[a,c]	9,927
… Group A – White	82.6% (9194)	75% (7,418)
… Group B - Mixed or multiple	2.9% (318)	0.4% (39)
… Group C - Asian, Asian Scottish or Asian British	8.3% (923)	9.6% (951)
… Group D – African	2.9% (328)	5.2% (520)
… Group E - Caribbean or Black	0.6% (64)	0.4% (38)
… Group F – other	2.8% (307)	2.5% (247)
… Group G - refused or not provided	0.2% (22)	-
… Group H – unknown	13.0% (1663)	7.2% (714)
BMI[d]	12,803*[a]	9,927
… Underweight	2.8% (354)	-
… Healthy	40.4% (5119)	41% (3995)
… Overweight	28.6% (3621)	29% (2845)
… Obese	28.2% (3574)	31% (3004)
… Unknown	1.1% (135)	0.83% (83)

Maternal diabetes	12,803*[a,e]	9,927
… No diabetes	87.1% (11,118)	89% (8806)
… Yes, gestational	11.8% (1501)	10% (1024)
… Yes, pre-existing	1.1% (142)	1.0% (97)
… Yes, status unknown	0.0% (2)	-
… Unknown	0.3% (40)	-
Live singleton births by method of birth[f]	12,564*	9,927
… Spontaneous	87.1% (11,118)	45% (4510)
… Caesarean - Elective	11.8% (1501)	23% (2248)
… Caesarean - Emergency	1.1% (142)	19% (1910)
… Instrumental	0.0% (2)	11% (1130)
… Unknown	0.3% (40)	1.3% (129)
Premature or full term by live singleton status (weeks of gestation)	12,564*[g]	9,927
… Premature (<37 weeks)	7.2% (906)	8.5% (842)
… Full-term (37+ weeks)	92.8% (11,658)	92% (9085)
Large for gestational age of live births	12,934*[h, i]	9,927[j]
… Large for gestational age	13.1% (1684)	9.5% (947)
… Not large for gestational age	86.8% (11,232)	89% (8827)
… Unknown	0.1% (18)	1.5% (153)
Admission to the neonatal care unit	10,592**[k]	9,927
… Baby admitted neonatal unit	10.1% (1073)	9.6% (951)
… Baby not admitted neonatal unit	89.9% (9519)	90% (8976)

*Public Health Scotland source is SMR02 (Scottish Morbidity Record - maternity interaction/stay), apart from neonatal unit admission, which is sourced from the Scottish Birth Record. *National Health Service (NHS) board of treatment, ** NHS board of residence. [a]An individual may have more than one maternity episode in a single year. [b]PHS used SIMD2020, cohort used SIMD2016. [c]Ethnicity data for 22/23 is provisional. [d]BMI Body mass index. [e]Diabetes identified from both hard-coding and ICD10 coding (International Classification of Diseases version 10). [f]Instrumental includes births by forceps, vacuum and breech. [g]Premature is defined as less than 37 weeks of gestation. [h]All live births are by appropriate weight for gestational age. [i]The unknown birthweight category are records with unknown sex, birthweight or gestation, or gestations outwith the range 24-42 weeks. [j]Large for gestational age is defined as a minimum World Health Organisation centile ≥ 90. [k]Includes live singleton and multiple births, stillbirths are excluded.*

5.4.2 Maternal characteristics

Here, I present the descriptive analysis of the maternal characteristics between GDM and non-GDM pregnancies, **Table 5.2**, and between the different treatments for GDM **Table 5.3**.

Of the 10,694 pregnancy episodes in the cohort, 10.4% (1,109) were complicated by GDM. The final treatment for GDM was mostly diet (802, 72.3%), followed by metformin (220, 19.8%) and then insulin (87, 7.8%). Queen Elizabeth University Hospital was the largest hospital of the five included and saw the most pregnancies (41.5%, 4,440), **Figure 5.1**.

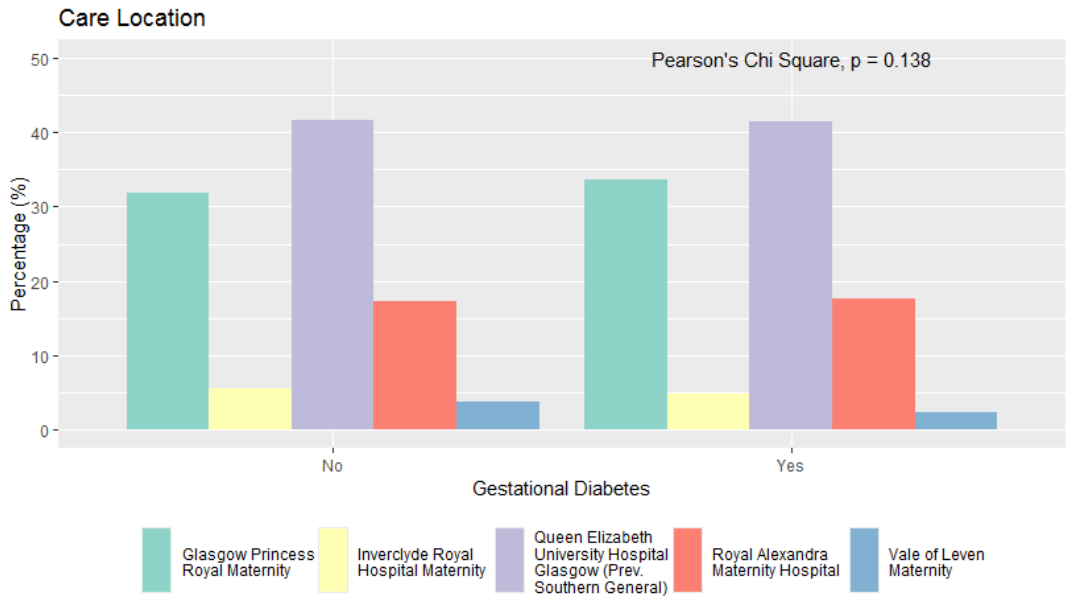
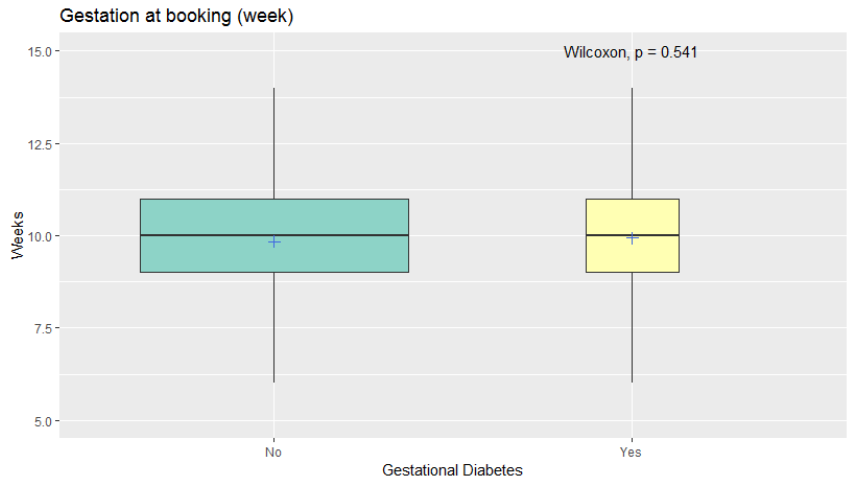


Figure 5.1 *Care location of the cohort grouped by GDM and non-GDM complicated pregnancies.*
Left: non-GDM, Right: GDM, Green: Glasgow Princess Royal Maternity, Yellow: Inverclyde Royal Hospital Maternity, Purple: Queen Elizabeth University Hospital, Glasgow (Prev. Southern General), Red: Royal Alexandra Maternity Hospital, Blue: Vale of Leven Maternity.
**Statistically significant, p<0.05.*

The mean gestational week at booking was not significantly different within the cohort (GDM: mean(SD) 10[3] gestational week, non-GDM:11[4.1] gestational week, p=0.541, treatment; Diet:11[3.1], Metformin: 10[2.7], Insulin: 10[1.8], p=0.230), **Figure 5.2**.

A



B

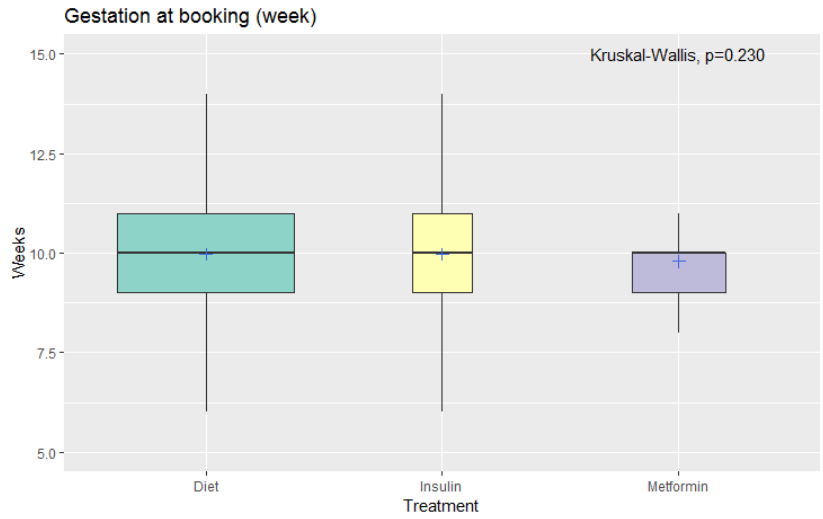


Figure 5.2 *Gestational week at booking of GDM and non-GDM complicated pregnancies, and treatment for GDM.*
Boxplot of gestational week at booking: (A) GDM vs non-GDM, Green: non-GDM, Yellow: GDM, (B) Diet vs Metformin vs Insulin, Green: Diet, Yellow: Insulin, Purple: Metformin. Black horizontal line median, blue cross mean, black vertical line range.
**Statistically significant, p<0.05.*
GDM: Gestational diabetes mellitus

Those with GDM were significantly older (32 [5.3] years vs. 31 [5.4] years, p<0.001), with those who required pharmacological therapy being older than diet-managed GDM (32 [5.4] years, 33 [4.8] years, 33 [5.3] years, p = 0.001). However, observing the boxplots in **Figure 5.3(A, B)**, it can be seen that the interquartile range (box) for each group overlaps, and the median line (horizontal black line) is not outside the other groups' boxes, indicating that the groups' ages are not very different, despite being statistically different.

The cohort was overweight with a mean BMI of 28 [6.1]kg/m^2, those with GDM had a higher BMI than those without (32 [7.4]kg/m^2, 27 [5.8]kg/m^2, p<0.001), with insulin-treated GDM having the highest (32 [7.2]kg/m^2, 32 [7.6]kg/m^2, 35 [7.5]kg/m^2, p<0.001). From the boxplots in **Figure 5.3(C, D)**, there is more of a distinction between the BMI of each group, in comparison to age, but still the boxes overlap, and the median line is not outside other groups' interquartile ranges.

Greater Glasgow and Clyde has high levels of ethnic minorities in comparison to the rest of Scotland, (Public Health Scotland, 2024b); despite this, 74.5% (7,970) of the cohort were White (Group A). GDM had higher diversity that the cohort, with metformin having a high number of Asian, Scottish Asian or British Asian (Group C), 23.0% (20).

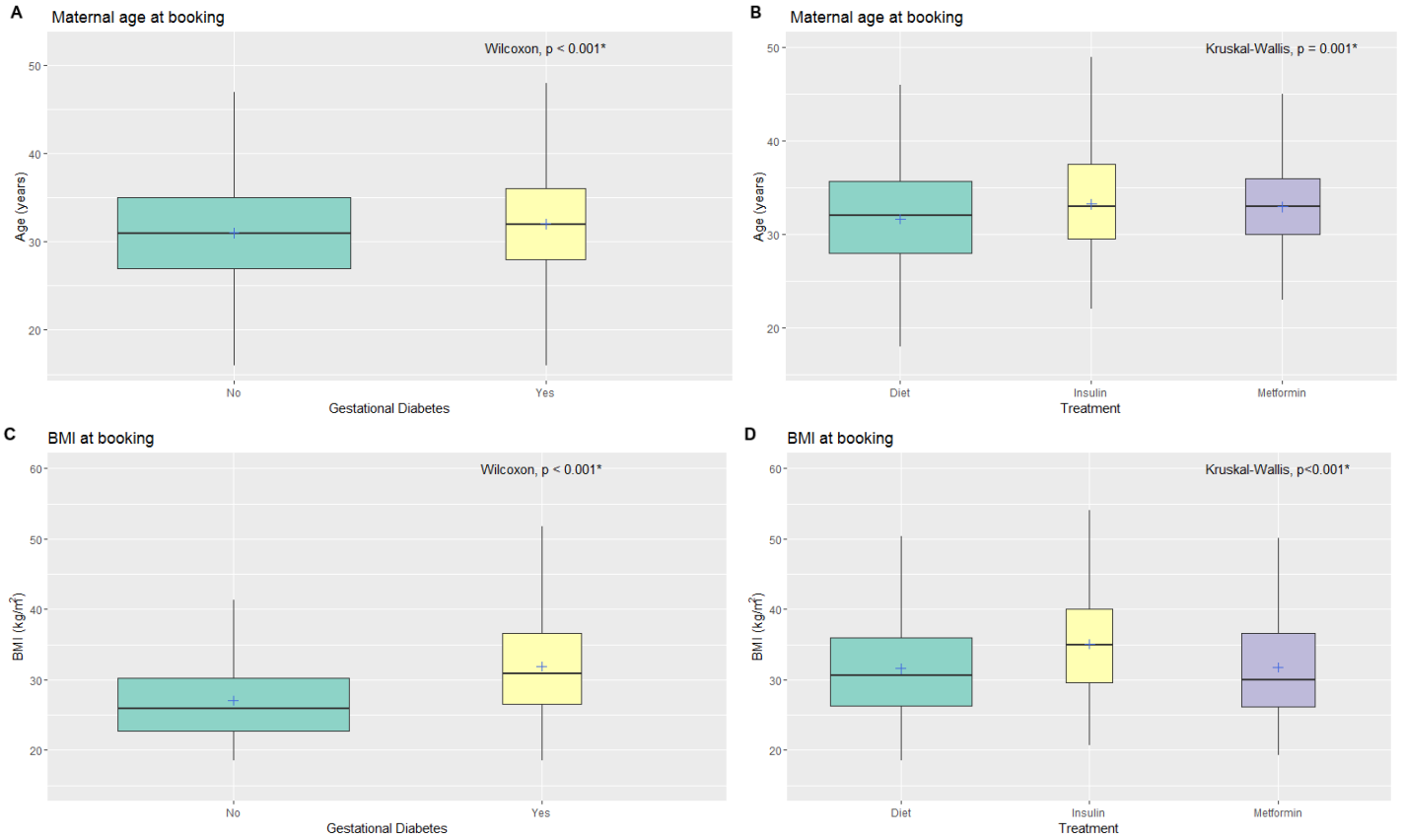


Figure 5.3 *Age and BMI at booking of GDM and non-GDM complicated pregnancies, and treatment for GDM.*
Boxplot of age (A-B) and BMI (C-D) at booking: (A, C) GDM vs non-GDM, Green: Non-GDM, Yellow: GDM, (B, D) Diet vs Metformin vs Insulin, Green: Diet, Yellow: Insulin, Purple: Metformin.
**Statistically significant, p<0.05. Black horizontal line median, blue cross mean, black vertical line range.*
GDM: Gestational diabetes mellitus, BMI: Body mass index

Similarly, Great Glasgow and Clyde had high levels of deprivation in comparison to Scotland (Public Health Scotland, 2024b). In the cohort, 40.1% (4,293) lived in the highest levels of deprivation (SIMD Quintile 1). There were higher rates of those with GDM living in SIMD Quintile 1 than those without GDM (45.4%(504), 39.5%(3789), p<0.001), **Figure 5.4**. Deprivation was not significantly different between the treatments of GDM (46.0%(369), 41.0%(90), 51.7%(45), p=0.726).

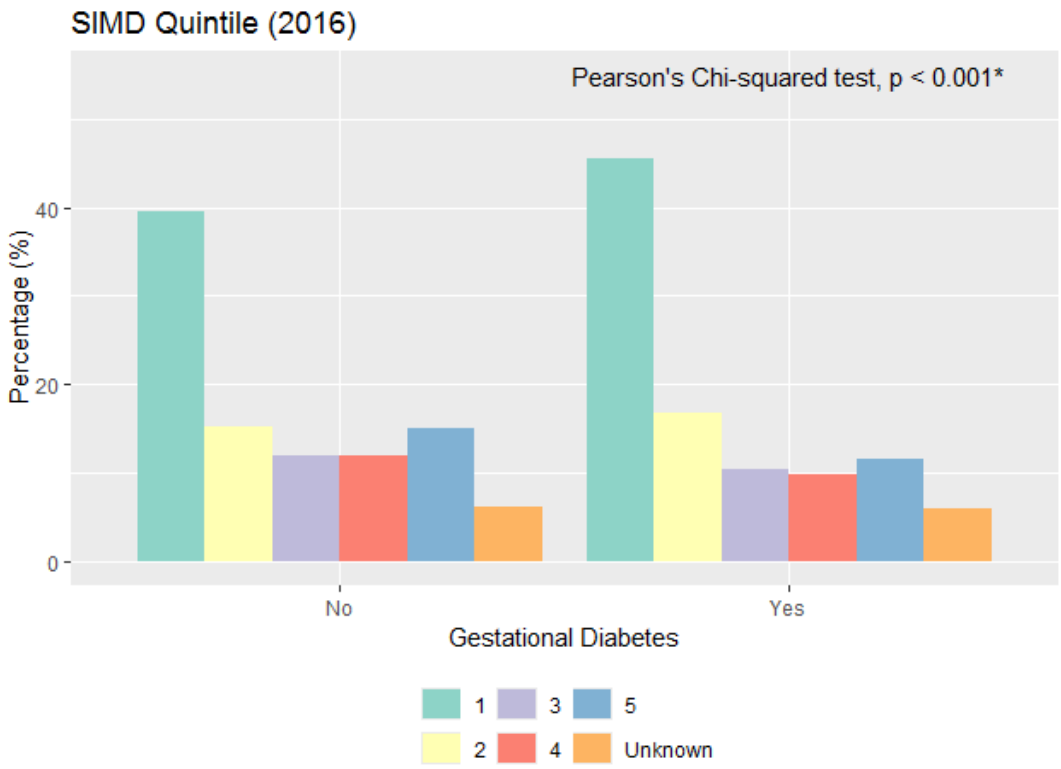


Figure 5.4 *SIMD Quintile (2016) of cohort grouped by GDM and non-GDM complicated pregnancies.*
Left: non-GDM, Right: GDM, SIMD Quintile 1: Green, 2: Yellow, 3: Purple, 4: Red, 5: Blue, Unknown: Orange.
**Statistically significant, p<0.05,*
SIMD: Scottish index of multiple deprivation.

The percentage of previous Caesarean births was significantly higher in GDM than in non-GDM complicated pregnancies (27.7% [307], 17.6% [1688], p<0.001), but there was no significant difference between treatment groups (25.7% [206], 32.3% [71], 34.5% [30], p=0.052). Similarly, gravida and parity were significantly higher within GDM complicated pregnancies (Gravida: 2.7 [1.7], 2.4 [1.5], p<0.001; Para: 2.1 [1.2], 1.8 [1], p<0.001). Between the treatment for GDM, those requiring insulin had the highest gravida and para (Gravida: 2.6 [1.7], 2.7 [1.5], 3.2 [2], p=0.011; Para: 2 [1.2], 2.1 [1.2], 2.3 [1.3], p=0.035). Furthermore, there were higher occurrences of previous macrocosmic babies among those with GDM (2.1% [23], 0.9% [87], p<0.001).

Within the cohort, 1.0% (107) were recorded to have type 1 or type 2 diabetes. Just over a third of non-GDM pregnancies had some family history (on either the baby's maternal or paternal side) of diabetes in comparison to almost 60% of those with GDM (58.2%(645), 34.2%(3,276), p<0.001), however other confounding factors, such as BMI, may attribute to this. As can be expected, previous GDM was significantly higher in those with GDM in their current pregnancy (17.0%(188), 1.5%(143), p<0.001), but this was not significantly different within the GDM treatment groups (16.7%(134), 16.4%(36), 20.7%(18), p=0.621).

Table 5.2 *Descriptive statistics of maternal characteristics between GDM and non-GDM complicated pregnancies*

Characteristic	N	Frequency	Mean [standard deviation]	N	Frequency	Mean [standard deviation]	p-value
	GDM[a]			Non-GDM[a]			
Total	1109	10.4%		9585	89.6%		
Care location[1]	1109			9585			0.138
… Glasgow Princess Royal Maternity	370	33.6%		3050	31.8%		
… Inverclyde Royal Hospital Maternity	50	4.8%		540	5.6%		
… Queen Elizabeth University Hospital, Glasgow (Prev. Southern General)	460	41.5%		3980	41.5%		
… Royal Alexandra Maternity Hospital	200	17.7%		1660	17.3%		
… Vale of Leven Maternity	30	2.4%		360	3.7%		
Gestation at booking (weeks)	1109		10 [3]	9585		11 [4.1]	0.541
Gestation at booking (days)	1109		3.1 [2]	9585		3 [2]	0.037**
Booking age (years)	1109		32 [5.3]	9585		31 [5.4]	<0.001**
Booking BMI[b] (kg/m^2)	1109		32 [7.4]	9585		27 [5.8]	<0.001**
Ethnicity[1]	1109			9590			<0.001**
… Group A White	610	54.8%		7360	76.8%		
… Group B Mixed or multiple ethnic groups	10	0.5%		40	0.4%		
… Group C Asian, Scottish Asian or British Asian	260	23.4%		780	8.2%		
… Group D African, Scottish African or British African	80	7.6%		480	5.0%		
… Group E Caribbean or Black	0	0.3%		40	0.4%		
… Group F Other ethnic group	50	4.1%		220	2.3%		
… Unknown	100	9.4%		670	7.0%		

SIMD[c] quintile (2016)	1109			9585			<0.001**
… 1	504	45.4%		3789	39.5%		
… 2	186	16.8%		1460	15.2%		
… 3	115	10.4%		1140	11.9%		
… 4	109	9.8%		1154	12.0%		
… 5	129	11.6%		1444	15.1%		
… Unknown	66	6.0%		598	6.2%		
Smoker at booking	1109			9585			0.661
… Not smoking	1014	91.4%		8721	91.0%		
… Smoking	95	8.6%		864	9.0%		
Previous Caesarean birth	1109			9585			<0.001**
… No	802	72.3%		7897	82.4%		
… Yes	307	27.7%		1688	17.6%		
Number of previous Caesarean births	1109		0.38 [0.71]	9585		0.22 [0.53]	<0.001**
Gravida	1109		2.7 [1.7]	9585		2.4 [1.5]	<0.001**
Para	1109		2.1 [1.2]	9585		1.8 [1]	<0.001**
Previous macrosomia	1109			9585			<0.001**
… No	1086	97.9%		9498	99.1%		
… Yes	23	2.1%		87	0.9%		
Previous GDM[a]	1109			9585			<0.001**
… No	921	83.0%		9442	98.5%		
… Yes	188	17.0%		143	1.5%		
Any family history of diabetes	1109			9585			<0.001**
… No	464	41.8%		6309	65.8%		
… Yes	645	58.2%		3276	34.2%		
Mother's family history of diabetes	1109			9585			<0.001**

… No	569	51.3%		7336	76.5%		
… Yes	540	48.7%		2249	23.5%		
Type 1 diabetes	1109			9585			-
… Not type 1	1109	100.0%		9534	99.5%		
… Type 1	0	0.0%		51	0.5%		
Type 2 diabetes	1109			9585			-
… Not type 2	1109	100.0%		9529	99.5%		
… Type 2	0	0.0%		56	0.5%		

*[a]GDM Gestational diabetes mellitus, [b]BMI Body mass index, [c]SIMD Scottish index of multiple deprivation. 1. Frequency rounded to the nearest 10. ** Statistically significant, p<0.05.*

Table 5.3 *Descriptive statistics of maternal characteristics between the final treatment for GDM complicated pregnancies*

Characteristic	N	Frequency	Mean [standard deviation]	N	Frequency	Mean [standard deviation]	N	Frequency	Mean [standard deviation]	p-value
	Diet			Metformin			Insulin			
Total	802	72.3%		220	19.8%		87	7.8%		
Care location[1]	802			220			87			<0.001**
… Glasgow Princess Royal Maternity	300	37.4%		40	18.2%		40	46.0%		
… Inverclyde Royal Hospital Maternity	40	5.0%		10	4.5%		0	0.0%		
… Queen Elizabeth University Hospital, Glasgow (Prev. Southern General)	300	37.4%		130	59.1%		30	34.5%		
… Royal Alexandra Maternity Hospital	140	17.5%		40	18.2%		20	23.0%		
… Vale of Leven Maternity	20	2.5%		10	4.5%		0	0.0%		
Gestation at booking (weeks)	802		11 [3.1]	220		10 [2.7]	87		10 [1.8]	0.230
Gestation at booking (days)	802		3.1 [2]	220		3.1 [1.9]	87		2.9 [2.1]	0.579
Booking age (years)	802		32 [5.4]	220		33 [4.8]	87		33 [5.3]	0.001**
Booking BMI[a] **(kg/m^2)**	802		32 [7.2]	220		32 [7.6]	87		35 [7.5]	<0.001**
Ethnicity[1]	802			220			87			0.191
… Group A White	450	56.1%		110	50.0%		50	57.5%		
… Group B Mixed or multiple ethnic groups	10	1.2%		0	0.0%		0	0.0%		

... Group C Asian, Scottish Asian or British Asian	180	22.4%		70	31.8%		20	23.0%		
... Group D African, Scottish African or British African	60	7.5%		10	4.5%		10	11.5%		
... Group E Caribbean or Black	0	0.0%		0	0.0%		0	0.0%		
... Group F Other ethnic group	30	3.7%		10	4.5%		10	11.5%		
... Unknown	80	10.0%		20	9.1%		10	11.5%		
SIMD[b] quintile (2016)	802			220			87			0.726
... 1	369	46.0%		90	41.0%		45	51.7%		
... 5	92	11.5%		28	12.7%		9	10.3%		
... Unknown	46	5.7%		14	6.4%		6	6.9%		
Smoker at booking	802			220			87			0.957
... Not smoking	733	91.4%		202	91.8%		79	90.8%		
... Smoking	69	8.6%		18	8.2%		8	9.2%		
Previous Caesarean birth	802			220			87			0.052
... No	596	74.3%		149	67.7%		57	65.5%		
... Yes	206	25.7%		71	32.3%		30	34.5%		
Number of previous Caesarean births	802		0.35 [0.69]	220		0.43 [0.73]	87		0.48 [0.76]	0.053
Gravida	802		2.6 [1.7]	220		2.7 [1.5]	87		3.2 [2]	0.011**
Para	802		2 [1.2]	220		2.1 [1.2]	87		2.3 [1.3]	0.035**
Previous GDM[c]	802			220			87			0.621
... No	668	83.3%		184	83.6%		69	79.3%		
... Yes	134	16.7%		36	16.4%		18	20.7%		

Any family history of diabetes	802			220			87			0.288
… No	343	42.8%		82	37.3%		39	44.8%		
… Yes	459	57.2%		138	62.7%		48	55.2%		
Mother's family history of diabetes	802			220			87			0.178
… No	420	52.4%		101	45.9%		48	55.2%		
… Yes	382	47.6%		119	54.1%		39	44.8%		

*[a]BMI Body mass index, [b]SIMD Scottish index of multiple deprivation. [c]GDM Gestational diabetes mellitus. 1. Frequency rounded to the nearest 10. ** Statistically significant, p<0.05.*

5.4.3 GDM-specific characteristics

This section explores whether the characteristics specific to GDM differed between those whose final treatment was diet, metformin or insulin; these are summarised in **Table 5.4**.

In the previous chapter, Chapter 4, I explained that GDM was not consistently recorded. Within the variable 'Diabetes Display', 67.2%(745) of those I defined as GDM had a label of "Current Gestational Diabetes" and 6.3%(70) were reordered in the variable 'Diabetes Gestational'.

The glucose levels recorded in the fasting OGTT were significantly higher in insulin treated GDM, (5.2 [0.5] mmol/L, 5.3 [0.46] mmol/L, 5.6 [0.57] mmol/L, $p<0.001$) and this test was taken significantly earlier (26 [6] gestational weeks, 24 [5.6] gestational weeks, 23 [5.6] gestational weeks, $p<0.001$). From the boxplots in **Figure 5.5(A)**, it can be seen that diet (green) is distinctly different from insulin (yellow), as diet's median is outside insulin's box (interquartile range), yet metformin (purple) overlaps with both diet and insulin, indicating that there is not a strong difference between metformin and the other treatment groups. Similarly, there was a significant difference between those who had a HbA1c taken at booking, with the insulin-treated group having higher results (35 [3.3] mmol/L, 35 [3.1] mmol/L, 36 [3.8] mmol/L, $p=0.038$). However, from the boxplot in **Figure 5.5(B)**, all treatments overlap; therefore, the groups may not be distinctly different, despite being statistically different.

Of those in which insulin was the final treatment, 59.8% (52) had metformin during their pregnancy, which was dispensed earlier than those that had a final treatment of metformin (metformin: 30 [4.4] gestational weeks, insulin: 27 [4.5] gestational weeks).

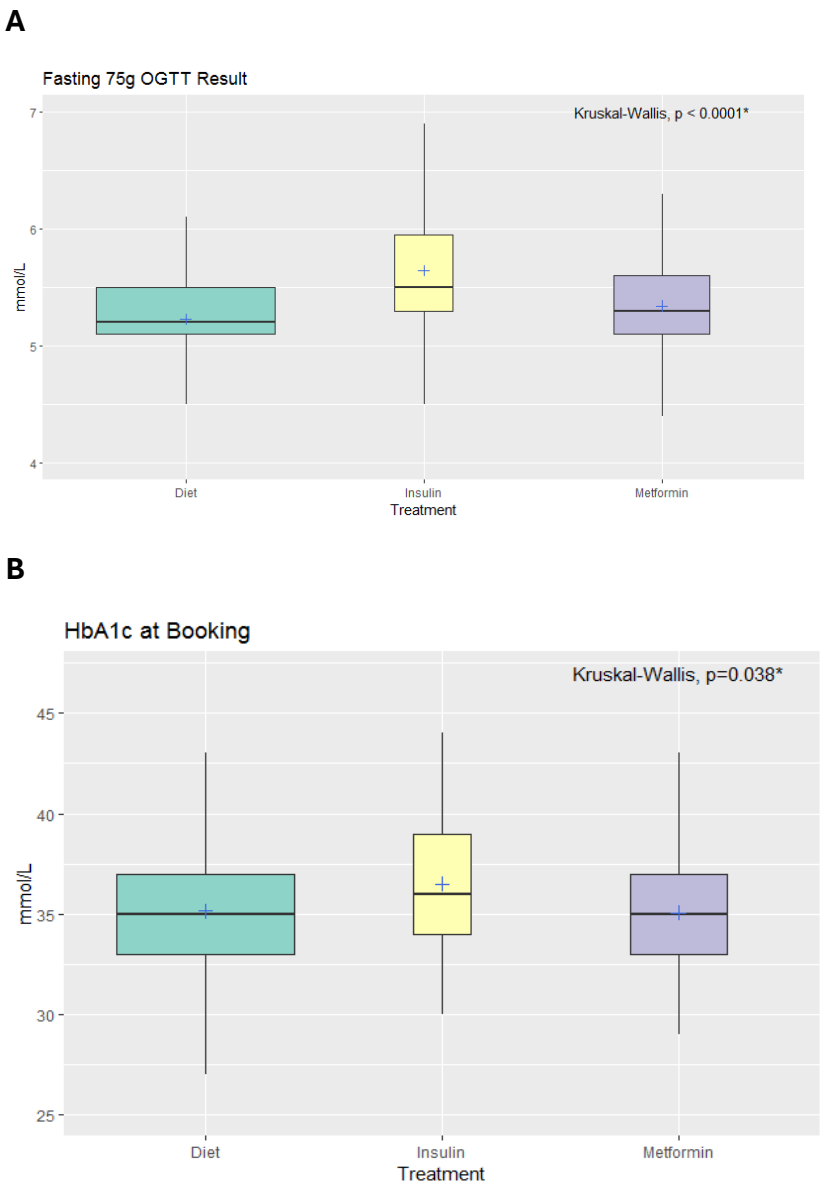


Figure 5.5 *GDM diagnostic tests*
Boxplot of (A) Fasting 75g OGTT result, (B) Result of HbA1c at booking
Green: Diet, Yellow: Insulin, Purple: Metformin.
Black horizontal line median, blue cross mean, black vertical line range.
**Statistically significant, p<0.05.*
GDM: Gestational diabetes mellitus, OGTT: Oral glucose tolerance test, HbA1c: glycosylated haemoglobin

Table 5.4 *Descriptive statistics of GDM-specific characteristics between the final treatment for GDM-complicated pregnancies*

Variable	N	Frequency	Mean [standard deviation]	N	Frequency	Mean [standard deviation]	N	Frequency	Mean [standard deviation]	p-value
	Diet			Metformin			Insulin			
Diabetes display[a]	802			220			87			-
… Current gestational diabetes	567	70.7%		122	55.5%		56	64.4%		
… No	101	12.6%		61	27.7%		13	14.9%		
Fasting glucose in OGTT[b]	800		5.2 [0.5]	220		5.3 [0.46]	87		5.6 [0.57]	<0.001**
Gestation OGTT[b] (week)	802		26 [6]	220		24 [5.6]	87		23 [5.6]	<0.001**
Gestation OGTT[b] (day)	802		3.3 [1.8]	220		3.4 [1.8]	87		3.5 [1.8]	0.285
Has an OGTT[b] result	793			220			85			0.270
… No	48	6.1%		14	6.4%		9	10.3%		
… Yes	745	93.9%		206	93.6%		76	87.4%		
Has a positive OGTT[b] result	793			220			85			<0.001**
… No	676	85.2%		177	80.5%		58	66.7%		
… Yes	117	14.8%		43	19.5%		27	31.0%		
Earliest HbA1c[c] result	158		35 [3]	48		35 [2.7]	13		38 [5]	0.033**
HbA1c[c] result near booking	552		35 [3.2]	163		35 [3.1]	67		36 [3.8]	0.092
HbA1c[c] result at booking	642		35 [3.3]	182		35 [3.1]	71		36 [3.8]	0.038**
HbA1c[c] result at 24-28 weeks of gestation	*		33 [2.8]	*		36 [4.4]	*		34 [5.1]	0.181
Metformin	802			220			87			<0.001**
… No	802	100%		0	0.0%		35	40.2%		

… Yes	0	0.0%		220	100%		52	59.8%		
Gestation at metformin dispense (week)	-	-		220		30 [4.4]	52		27 [4.5]	-
Gestation at metformin dispense (day)	-	-		220		3.3 [1.8]	52		3.5 [1.9]	-
Gestation at insulin dispense (week)	-	-		-	-		87		31 [4.6]	-
Gestation at insulin dispense (day)	-	-		-	-		87		3.1 [1.8]	-

*[a]Variable sourced from Badgernet, [b]OGTT Oral glucose tolerance test, [c]HbA1c glycosylated haemoglobin. *Redacted due to the potential risk of disclosure. ** Statistically significant, p<0.05.*

5.4.4 Maternal and neonatal outcomes

Finally, I present the maternal and neonatal outcomes of the cohort. **Table 5.5** summarises the outcomes of GDM and non-GDM complicated pregnancies, and **Table 5.6** summarises the outcomes of the different treatments for GDM.

The gestational week of birth was similar but significantly higher in non-GDM pregnancies (38 [2], 38 [3.8], p<0.001), but there was no difference in the frequency of premature births (defined as a pregnancy before 37 weeks of gestation) (9.5%(105), 9.5%(913), p=0.994). Although not statistically significant, metformin-treated GDM did have a marginally higher frequency of premature births (9.7%(78), 8.2%(18), 10.3%(9), p=0.754).

Just over half of GDM-complicated pregnancies had a Caesarean birth compared to two-fifths of non-GDM pregnancies (53.9%(593), 40.0%(3,834, p<0.001), **Figure 5.6(A)**. There were similar levels of emergency Caesarean births between the different treatment groups of GDM (21.3%(171), 21.8%(48), 28.7%(25), p=0.284). As it is recommended for women with GDM on insulin to have Caesarean births (Scottish Intercollegiate Guidelines Network, 2024), those that required insulin, did experience a higher frequency of elective Caesarean births (29.8%(239), 35.9%(79), 41.4%(36), p=0.033), **Figure 5.6(C)**. Furthermore, induction and instrumental births were both higher in GDM-complicated pregnancies, (induction: 39.0%(433), 34.4%(3,294), p=0.002, instrumental: 60.1%(667), 51.9%(4,975), p<0.001), again with insulin treated GDM experiences higher rates (induction: 39.9%(320), 37.7%(83), 34.5%(30), p=0.558, instrumental: 58.5%(469), 61.4%(135), 72.4%(63), p=0.038).

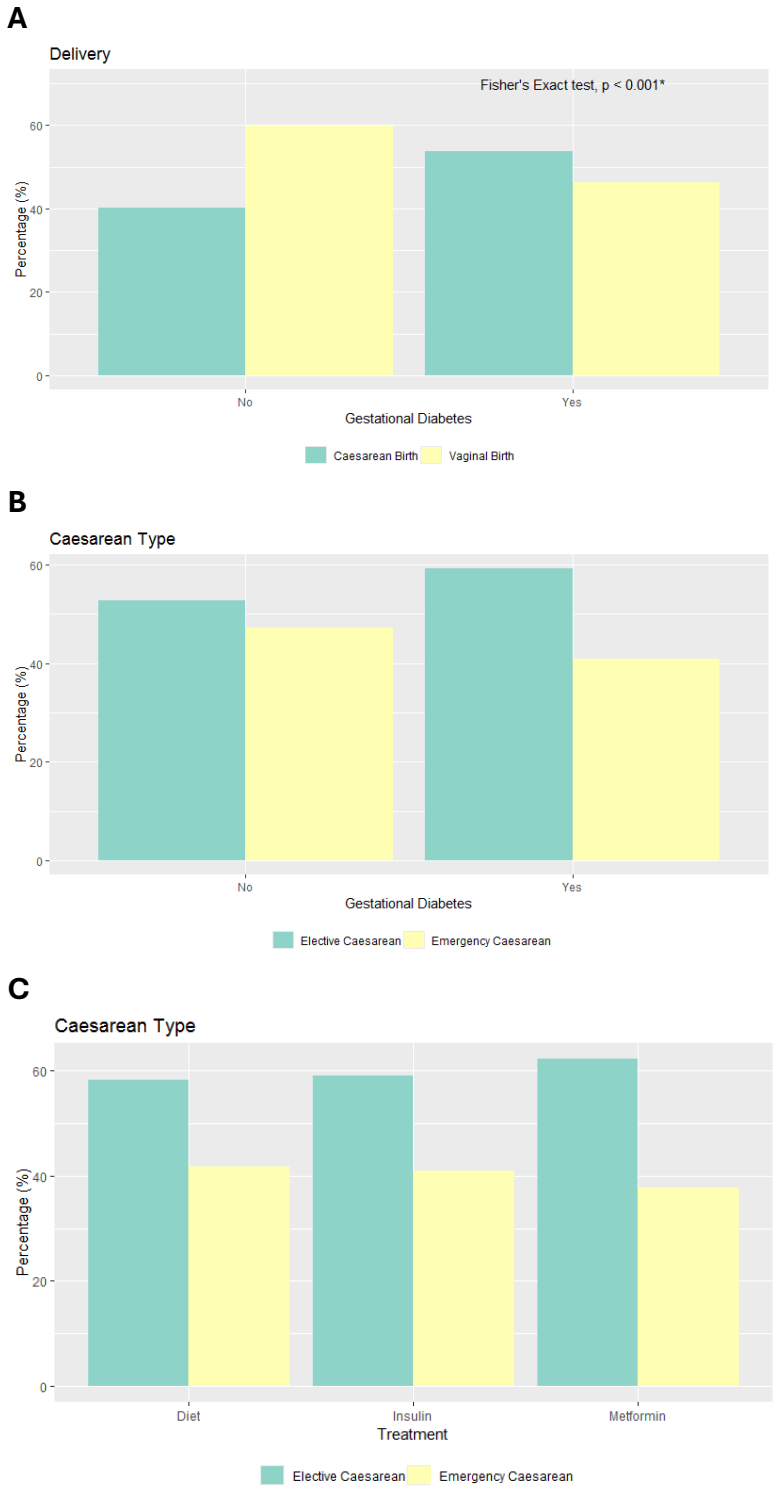


Figure 5.6 *Delivery and Caesarean type of GDM and non-GDM complicated pregnancies, and treatment for GDM.*
Boxplot of (A) Delivery type of GDM vs non-GDM, Left: non-GDM, Right: GDM, Green: Caesarean, Yellow: Vaginal, (unknown and other delivery types have been omitted due to the potential risk of disclosure), (B-C) Caesarean type, Green: Elective, Yellow: Emergency, (B) GDM (right) vs non-GDM (left), (C) Diet (left) vs Metformin (middle) vs Insulin (right)
**Statistically significant, p<0.05.*
GDM: Gestational diabetes mellitus

Baby sex was similar in both GDM and non-GDM complicated pregnancies (Female: 49.0%(542), 47.6%(4,559), p=0.001), with birthweight not being significantly different (3282 [556]g, 3271 [716]g, p=0.075). After adjusting for gestational age, LGA was significantly higher in GDM-complicated pregnancies compared to non-GDM (15.0%(166), 9.0%(841), p<0.001), as shown in **Figure 5.7**, but this was not influenced by GDM treatment (14.8%(119), 15.5%(34), 14.9%(13), p=0.980). Most 5-minute Apgar scores were ≥ 7 (96.6%(1,071), 94.0%(9,012), p=0.002). GDM-complicated pregnancies experienced higher rates of admissions to the neonatal unit than non-GDM pregnancies (41.2%(158), 9.4%(899), p<0.001), and although higher rates were observed in insulin-treated GDM cases, this difference was not statistically significant (13.8%(111), 14.1%(31), 18.4%(16), p=0.513).

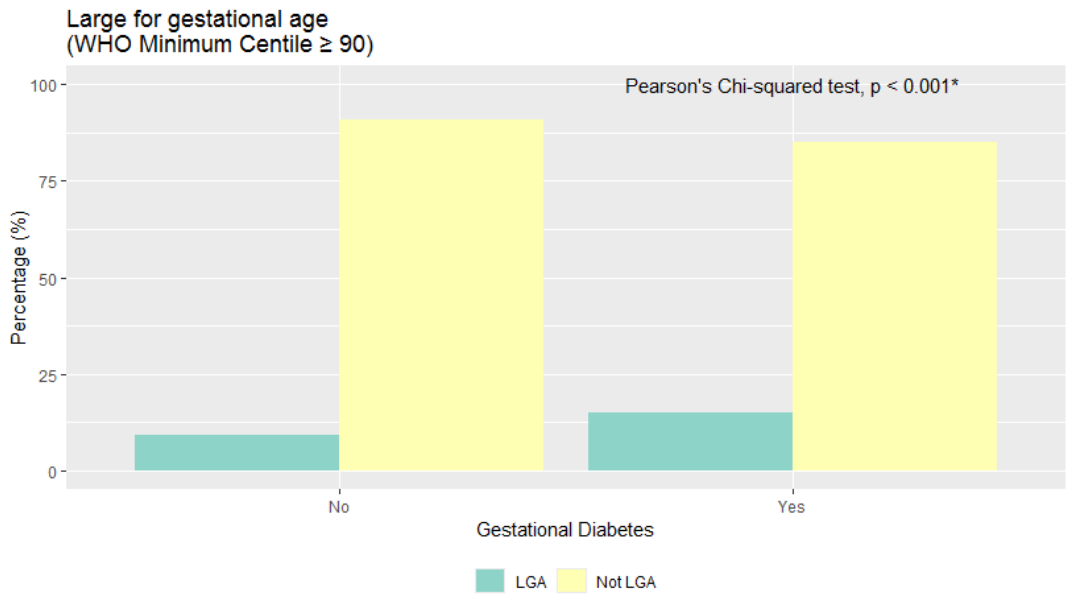


Figure 5.7 *Large for gestational age of GDM vs non-GDM complicated pregnancies.*
Left: non-GDM, Right: GDM, Green: LGA, Yellow: Not LGA. LGA is defined as having a WHO minimum birthweight centile ≥ 90.
**Statistically significant, p<0.05.*
GDM: Gestational diabetes mellitus, LGA: Large for gestational age, WHO: World Health Organisation.

Table 5.5 *Descriptive statistics of maternal and neonatal outcomes between GDM and Non-GDM complicated pregnancies*

Outcome	N	Frequency	Mean [standard deviation]	N	Frequency	Mean [standard deviation]	p-value
	GDM[a]			Non-GDM[a]			
Gestation at birth (weeks)	1109		38 [2]	9585		38 [3.8]	<0.001**
Gestation at birth (days)	1109		2.7 [2]	9585		2.7 [2]	0.220
Premature[b]	1109			9585			0.994
… Full term (≥ 37 weeks)	1004	90.5%		8672	90.5%		
… Premature (<37 weeks)	105	9.5%		913	9.5%		
Final type of delivery	1109			9585			
… Caesarean birth	593	53.5%		3834	40.0%		<0.001**
… Vaginal birth	511	46.1%		5710	59.6%		
… Unknown	5	0.5%		41	0.4%		
Elective Caesarean birth	1109			9585			<0.001**
… No	755	68.1%		7545	78.7%		
… Yes	354	31.9%		2040	21.3%		
Emergency Caesarean birth	1109			9585			0.022**
… No	865	78.0%		7756	80.9%		
… Yes	244	22.0%		1829	19.1%		
Induced birth	1109			9585			0.002**
… Induced	433	39.0%		3294	34.4%		
… Unknown	676	61.0%		6291	65.6%		
Instrumental birth	1109			9585			<0.001**
… No	442	39.9%		4610	48.1%		
… Yes	667	60.1%		4975	51.9%		
Any tear	1109			9585			0.023**
… No	529	47.7%		4224	44.1%		

… Yes	580	52.3%		5361	55.9%		
Total blood loss (ml)	1109		647 [560]	9585		602 [520]	<0.001**
Baby sex	1105			9404			0.001**
… Female	542	49.0%		4559	47.6%		
… Male	563	51.0%		4845	50.5%		
Birth weight (g)	1109		3282 [556]	9554		3271 [716]	0.075
Macrosomia[c]	1109			9554			0.422
… No	1102	99.4%		9466	99.1%		
… Yes	7	0.6%		88	0.9%		
Large for gestational age[d]	1105			9377			<0.001**
… No	939	85.0%		8536	91.0%		
… Yes	166	15.0%		841	9.0%		
5-minute Apgar score	1109			9585			0.002**
… <7	13	1.2%		153	1.6%		
… ≥ 7	1071	96.6%		9012	94.0%		
… Unknown	25	2.3%		420	4.4%		
Shoulder dystocia	1109			9585			<0.001**
… No	1090	98.3%		9260	96.6%		
… Yes	14	1.3%		112	1.2%		
… Unknown	5	0.5%		213	2.2%		
Baby admitted to neonatal unit	1109			9585			<0.001**
… No	951	85.8%		8686	90.6%		
… Yes	158	14.2%		899	9.4%		

[a]GDM Gestational diabetes mellitus [b]Premature birth is defined as a birth before 37 weeks of gestation, [c]Macrosomia is defined as birthweight ≥ 4500g, [d]Large for gestational age is defined as a minimum World Health Organisation centile ≥ 90.

Table 5.6 *Descriptive statistics of maternal and neonatal outcomes between the final treatment for GDM-complicated pregnancies*

Outcome	N	Frequency	Mean [standard deviation]	N	Frequency	Mean [standard deviation]	N	Frequency	Mean [standard deviation]	p-value
	Diet			Metformin			Insulin			
Gestation at birth (weeks)	802		38 [2.2]	220		38 [1.4]	87		38 [1.3]	<0.001**
Gestation at birth (days)	802		2.8 [2]	220		2.4 [1.9]	87		2.5 [2.2]	0.043**
Premature[a]	802			220			87			0.754
… Full term (≥ 37 weeks)	724	90.3%		202	91.8%		78	89.7%		
… Premature (<37 weeks)	78	9.7%		18	8.2%		9	10.3%		
Elective Caesarean birth	802			220			87			0.033**
… No	563	70.2%		141	64.1%		51	58.6%		
… Yes	239	29.8%		79	35.9%		36	41.4%		
Emergency Caesarean birth	802			220			87			0.284
… No	631	78.7%		172	78.2%		62	71.3%		
… Yes	171	21.3%		48	21.8%		25	28.7%		
Induced birth	802			220			87			0.558
… Induced	320	39.9%		83	37.7%		30	34.5%		
… Unknown	482	60.1%		137	62.3%		57	65.5%		
Instrumental birth	802			220			87			0.038**
… No	333	41.5%		85	38.6%		24	27.6%		
… Yes	469	58.5%		135	61.4%		63	72.4%		
Any tear	802			220			87			0.002**
… No	361	45.0%		128	58.2%		40	46.0%		
… Yes	441	55.0%		92	41.8%		47	54.0%		

Total blood loss (ml)	802		631 [478]	220		674 [786]	87		725 [575]	0.334
Birth weight (g)	802		3286 [595]	220		3277 [481]	87		3262 [494]	0.543
Macrosomia[b]	802			220			87			0.638
… No	*			*			*			
… Yes	*			*			*			
Large for gestational age[c]	798			220			87			0.980
… No	679	84.7%		186	84.5%		74	85.1%		
… Yes	119	14.8%		34	15.5%		13	14.9%		
5-minute Apgar score	802			220			87			0.245
… <7	7	0.9%		*			*			
… ≥ 7	777	96.9%		*			*			
… Unknown	18	2.2%		*			*			
Baby admitted to neonatal unit	802			220			87			0.513
… No	691	86.2%		189	85.9%		71	81.6%		
… Yes	111	13.8%		31	14.1%		16	18.4%		

*[a]Premature birth is defined as a birth before 37 weeks of gestation, [b]Macrosomia is defined as birthweight ≥ 4500g, [c]Large for gestational age is defined as a minimum World Health Organisation centile ≥ 90. *Redacted due to the potential risk of disclosure. ** Statistically significant, p<0.05.*

5.5 DISCUSSION

From the comparison with PHS data, the cohort is a good representation of the Greater Glasgow and Clyde pregnant population, particularly regarding the maternal characteristics.

It is established that older maternal age, obesity, ethnicity and socioeconomic deprivation are associated with GDM and linked to adverse outcomes, (Vounzoulaki et al., 2024; Farrar et al., 2016; Lawrence et al., 2008) and these were also observed in this cohort. I found that GDM in Greater Glasgow and Clyde, with a booking date between 1st April 2022 and 31st December 2023, involved older women and those with higher BMI compared to those without GDM. The GDM cases requiring pharmacological therapy were older, and those needing insulin had a higher BMI than other GDM treatment groups. Despite a predominantly white population (74.5%), GDM-complicated pregnancies involved diverse ethnicities. Greater Glasgow and Clyde exhibits high levels of deprivation, with 40.0% of the cohort living in the most deprived SIMD quintile 1, which was notably higher among women with GDM (45.4% in quintile 1) than those without (39.5%). Diagnosis tests for GDM were more frequently performed in women receiving insulin treatment and were conducted earlier.

Though GDM-complicated pregnancies experienced a higher frequency of Caesarean births, these were mostly elective, with insulin-treated experiencing the highest frequency. LGA babies were higher in GDM-complicated pregnancies, but this was not affected by treatment. GDM-complicated pregnancies experienced high levels of admission to the neonatal unit, and although this was higher in the insulin treatment group, this was not significant. Similar results have been found in a retrospective cohort study of 1,319 women with GDM in which 752 were treated with GDM. It was reported that insulin treated GDM were more likely have Caesarean births (OR 1.67, 95% CI 1.25-2.23), LGA (19.7% vs 12.5%; $p < 0.01$) and admission to the NICU (OR 4.88, 95%CI 3.54-6.73), (Bogdanet et al., 2018).

5.5.1 Strengths and limitations

This descriptive analysis has strengths, as I selected a large cohort of 10,694 pregnancy episodes of real-world, granular data, with good statistical power and reliability. Alongside investigating the differences between GDM and non-GDM complicated pregnancies, the effect of GDM treatment was also explored. However, there are some limitations due to the difference in sample size between the GDM and non-GDM populations, which reduced the power to detect actual differences within the groups; hence, there were instances where there were significant p-values, but no visible differences with the boxplots. In addition, where there was statistical significances, this does not directly transfer to clinical significance. Furthermore, as the aim of the analysis was to establish a general understanding of the cohort, I did not account for any cofounding factors.

5.6 CHAPTER SUMMARY

This chapter has presented the descriptive analysis of the cohort of singleton pregnancies that had a booking date from 1st April 2022 to 31st December 2023 in Greater Glasgow and Clyde. Comparison of maternal characteristics and pregnancy outcomes of GDM and non-GDM complicated pregnancies were investigated, as well as looking into the differences within treatment groups of GDM.

In comparison to the PHS data of maternities in Greater Glasgow and Clyde that were delivered between 1st April 2022 and 31st March 2023, the cohort is a reasonable representation of the general population.

A tenth of the pregnancies were complicated with GDM; of these, 72.3% were managed through diet, 19.8% through metformin and 7.8% with insulin. Non-GDM-complicated pregnancies were women who were younger and had a lower BMI than GDM-complicated pregnancies. Within the GDM-complicated pregnancies, those that required pharmacological therapy were older, and those that required insulin had higher BMI. There was higher ethnic diversity

within the GDM pregnancies, and a significant proportion of GDM-complicated pregnancies were living in the most deprived areas.

Insulin-treated GDM had higher GDM diagnosis test results (higher blood glucose levels) and experienced a higher frequency of partially elective Caesarean births and admission to the neonatal unit, but not significantly, in comparison to other GDM treatments and non-GDM complicated pregnancies.

This analysis has given a general understanding of the cohort and the differences within it compared to general population statistics. The results of the analysis influenced the modelling described in the following chapter, Chapter 6.

CHAPTER 6

Development of an Insulin Risk Prediction Model for Gestational Diabetes

6.1 INTRODUCTION

As detailed in 1.2.1.3, current care for GDM is managed through lifestyle or pharmacological therapy (metformin or insulin). Women are required to check and record their blood glucose approximately four times a day, which is reviewed by HCPs at an antenatal-metabolic appointments, typically fortnightly (National Institute for Health and Care Excellence (NICE), 2015, 25 February ; Scottish Intercollegiate Guidelines Network, 2024). Pharmacological therapy is then initiated if blood glucose levels are outside the target range. This is usually administered in a stepwise fashion; firstly, metformin is started and then escalated to insulin if further support is needed (Mayne et al., 2025; Scottish Intercollegiate Guidelines Network, 2024).

Those treated with insulin often experience higher rates of adverse outcomes, such as LGA, Caesarean birth and admission to the neonatal unit, as shown in Chapter 5 and the literature (de Souza et al., 2024; Bogdanet et al., 2018).

From a systematic review of 156 studies with 7,506,061 pregnancies investigating the adverse pregnancy outcomes associated with GDM, after adjusting for confounding factors, it was reported that those treated with insulin had increased odds of having a LGA (odds ratio 1.61, 1.09 to 2.37) or a baby admitted to NICU (2.29, 1.59 to 3.31), in comparison to non-GDM pregnancies (Ye et al., 2022). It is thought that achieving the target blood glucose ranges could reduce these adverse outcomes. Therefore, there is a need to identify and monitor those at a higher risk of insulin, so that appropriate care can be implemented sooner, improving the achievement of target blood glucose levels. Similarly, by risk-stratifying care of women with GDM, it could also improve the allocation of hospital resources and move towards a personalised model of GDM care and management.

Previous literature (shown in Chapter 3) has used machine learning to classify treatments for women with GDM; these models predicted either insulin or pharmacological therapy (grouping oral agents and insulin) and used logistic regression, CART, LASSO, response-mean, random forest, and extreme gradient boosting. Some models were validated, internally through cross-validation, and some externally, either geographically or temporarily. There was a varying degree of performance; however, Liao et al. (Liao et al., 2022) reported that logistic regression performed similarly to more complex machine learning algorithms (response-mean, LASSO regression, CART, random forest, and extreme gradient boosting) on an external validation set, showing that treatment prediction for GDM is a simple classification problem.

No models have been developed for a Scottish population, and few have been developed for clinical implementation. In this Chapter, I will present the development and comparison of two models predicting insulin risk for GDM. This will use a train-test split of 996 pregnancies complicated by GDM that had a booking date between 1st April 2022 and 31st December 2023 in Glasgow and will be evaluated during training through 10-fold cross-validation and then validated on unseen test data.

6.2 AIMS AND ENDPOINTS

This Chapter will focus on objective: III Develop and validate a machine learning model to predict insulin therapy in women with GDM in Glasgow. The specific aims and endpoints of this Chapter are:

6.2.1 Aims

1) To develop one or more supervised machine learning model(s) that predicts at the time of GDM diagnosis, the risk of insulin therapy to manage GDM.
2) To validate the supervised machine learning model on unseen data and assess its performance.

6.2.2 Endpoints

1) A risk prediction model that is suitable to use at the time of GDM diagnosis, that can predict the risk of insulin therapy to manage GDM.

6.3 METHODS

To develop the models, I followed the machine learning workflow described by Sandfeld et al. (Sandfeld, 2024), shown in **Figure 6.1**. I have based the reporting of the results on the Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis with Artificial Intelligence (TRIPOD + AI), (Collins et al., 2024).

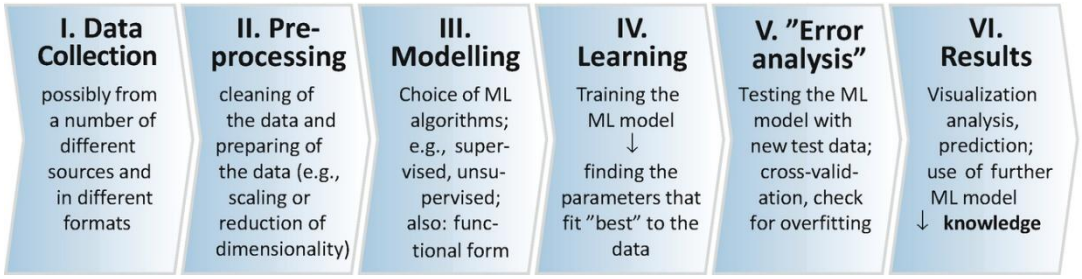


Figure 6.1 *Modelling Workflow. Reproduced with permission from Springer Nature of Sandfeld, S. (2024). Introduction and General Concepts of Machine Learning and Data Science. Chapter 11, Figure 11.2.*

All pre-processing and modelling were completed within the West of Scotland Safe Haven. R (version 4.3.2) was used for feature selection and Python (version 3.11) for modelling. Microsoft Word (2016) was used for documentation.

6.3.1 Study population and design

The model was developed on 996 clinically collected complete cases of singleton GDM pregnancy episodes from five hospitals in Greater Glasgow and Clyde, with booking dates between 1st April 2022 and the 31st December 2023. The full details of the data cleaning and identification of GDM complicated pregnancies, alongside the ethics, are presented in Chapter 4. To summarise, pregnancies that were complicated by GDM were identified and each included in the modelling data if it:

1) Had a booking date between 1st April 2022 and 31st December 2023.
2) Was a singleton pregnancy episode.
3) Did not end in a miscarriage or termination.
4) Had a maternal age > 16 years.
5) Had a maternal booking BMI > 18.5 kg/m^2.
6) Was a first pregnancy episode within the study period.
7) Had a fasting OGTT result ≥ 5.1 mmol/L taken between the booking date and delivery date OR had a label of "Gestational Diabetes" in 'DiabeteDisplay' variable.
8) Fasting OGTT < 7 mmol/L.
9) If present, HbA1c earliest, HbA1c at booking or HbA1c near booking < 48 mmol/L and taken between the booking date and delivery date.
10) Was not labelled as type 1 or type 2 diabetes.
11) Did not have insulin dispensed 3-9 months before booking date.
12) Had no missing data in the following variables: Fasting OGTT result, Gestation (week + day) at fasting OGTT results, Ethnicity, History of GDM, BMI, Age, Previous macrosomic baby, Family history of diabetes
13) Did not have metformin dispensed 3-9 months before booking date (suspected polycystic ovary syndrome).

To be able to validate the model's performance, I chose to hold out 30% (299) of the data in an unseen test set and build and train the model using 10-fold cross-validation on the remaining 70% (697) (Maleki et al., 2020), shown in **Figure 6.2**. A 70:30 train-test split is common for small data sizes (Gupta and Sedamkar, 2020). Furthermore, as reported in Chapter 5, the majority (92.1%) of all GDM cases in the cohort were managed through diet or metformin, so I chose a 70:30 split to ensure that there were enough samples in the test data to make a validation assessment of the models' performances.

The target end-user for the model is HCPs using it within a GDM clinic. It has been reported that 'black box' models reduce HCPs' trust in models (Wang et al., 2023), therefore, I selected easily interpretable algorithms. Logistic regression and CART were selected because they produce simple classification which is good for use within healthcare (Gupta and Sedamkar, 2020). Model development and evaluation were similar for the two models, with only minor differences, which is explained later in this section.

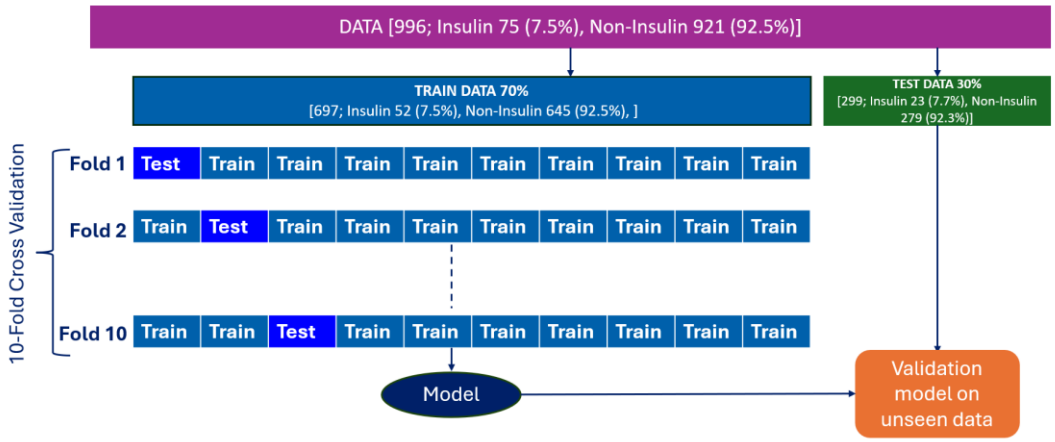


Figure 6.2 *Data split. 70:30 train-test split with 10-fold cross-validation to develop the model and validate on unseen test data.*

6.3.2 Outcome

Insulin was chosen to be the predicting outcome, as those who require insulin to manage GDM experience higher rates of elective Caesarean and instrumental births, and a higher frequency of admission to the neonatal unit (shown in Chapter 5) alongside having significantly different maternal

characteristics. Overall, of the 996 pregnancy episodes, 7.5% (75) required insulin. The other 92.5% (921) of pregnancies were managed through diet and lifestyle changes or metformin (supplementary statistical analysis of the diet versus metformin is in **Appendix G**).

6.3.3 Candidate predictors

Variable or feature selection followed the steps outlined in Chowdhury et al. 2020. Step 1: Reduce the number of variables in the dataset by selecting variables through literature search, clinical knowledge and experience and/or consultation with experts. Step 2: Apply a variable selection model to identify the variables to include in the final model.

Common routinely collected variables available at the time of GDM diagnosis were selected for potential predictors for insulin; these included:

- Gestation at booking (weeks)
- Gestation at booking (days)
- Booking age
- Booking BMI
- Family origin of mother (SMR02 grouping)
- SIMD 2016 quintile
- Smoker at booking
- Previous Caesarean birth
- Number of previous Caesarean births
- Gravida
- Para
- Previous macrosomia baby
- Previous GDM
- Any family history of diabetes
- Mother's family history diabetes
- Fasting OGTT result
- Gestational week of fasting OGTT result
- Gestational day of fasting OGTT result
- HbA1c at booking result

Binning continuous variables into discrete variables can improve the fitting power of the model and reduce overfitting (Gupta and Sedamkar, 2020). Additionally, algorithms cannot handle unknowns (NAs). Variables with

missing data were assessed and dichotomised by using the Youden index to identify the threshold. The Youden index identifies an optimal cutoff point for a predictive test (Ruopp et al., 2008).

Candidate predictors were selected for either having known clinical importance or a p-value above a threshold. The p-value threshold should be between 0.15-0.25; a higher p-value indicates that a variable has predictive potential rather than if the variable has significance ($p<0.05$) (Hosmer et al., 2013; Chowdhury and Turin, 2020). Variables available at GDM diagnosis were selected as candidate predictors by using univariate analysis with a $p<0.20$. Correlation between variables was assessed, and those highly correlated were removed. As CARTs have an embedded feature selection, the candidate variables selected by univariate analysis were then used in the model. Logistic regression requires an additional set on top of the filter method (Hosmer et al., 2013); therefore, LASSO was applied to identify predictors for logistic regression. LASSO regularises and adds penalties to the variables; it shrinks the variables' coefficients down, reducing some to zero, and leaving only important variables with coefficients, (Gupta and Sedamkar, 2020). Those remaining with coefficients were then used as predictors for the logistic regression model.

6.3.4 Sample size calculation

The sample size was calculated using the simple event per variable ratio with the common rule of ten (Riley et al., 2020). Given the 7.5% prevalence of insulin therapy and the training data consisting of 697 pregnancy episodes, the maximum number of predictions that could be included in the model was five.

6.3.4.1 Pre-processing

The logistic regression model was scaled using sklearn 'StandardScaler', which standardises the features by removing the mean and scaling to the unit variance. This reduces the negative influences of outliers on the logistic regression model results; tree-based algorithms, such as CART, are not

affected by this (Gupta and Sedamkar, 2020). Both models had the categorical predictors encoded using sklearn 'OneHotEncoder'. Categorical variables can be misinterpreted during the learning process of an algorithm; therefore, encoding reduces these mistakes and speeds up the learning process (Gupta and Sedamkar, 2020).

6.3.4.2 Class imbalance

Class imbalances are common in healthcare data (Gupta and Sedamkar, 2020) and the data had a class imbalance, with the minority class being 7.5% (insulin). CART and logistic regression are often biased towards the majority class (e.g. non-insulin) and show skew and inaccurate performance (Gupta and Sedamkar, 2020). To address this, I resampled the training data using the Synthetic Minority Oversampling Technique (SMOTE) (Fernández et al., 2018). SMOTE, rather than just replicating the data, introduces synthetic samples of the minority class through interpolation. This then matches the sample number of the majority class (e.g. if there are 100 in the majority and 5 in the minority, SMOTE will resample the minority so there are 100 in both classes). It is particularly good when positive cases (insulin) are underrepresented.

6.3.4.3 Hyperparameter tuning

To optimised the algorithm's performance, the algorithms' hyperparameters were tuned using sklearn 'GridSearchCV' on the training data, which performs an extensive search for each combination of parameters to optimise the performance during 10-fold cross-validation. To ensure that there was no data leakage between the training and testing splits within the folds of the cross-validation, imblearn 'Pipeline' was used.

Logistic regression candidate hyperparameters were selected from Sklearn documentation (https://scikit-learn.org/stable/api/sklearn.linear_model.html), these were Plenty: L1, L2, Elasticnet or none, C:-10^4 – 10^4, Solver: 'liblinear' or 'saga', and Maximum iteration: 100, 200 or 300. CART's candidate hyperparameters were selected

from the literature (Gomes Mantovani et al., 2024), these were maximum depth: 1-5, minimum sample leaf: 1-21 and minimum sample split: 2-41.

6.3.4.4 Performance metrics

Both models were built on the training data and evaluated through 10-fold cross-validation, and metrics are reported as the mean [SD] of the 10-folds. The model developed from the training data was then validated on unseen test data. To assess the models' performance, the AUROC, precision, recall, accuracy, F1-score and confusion matrix were calculated.

6.3.5 Statistical analysis

The same approach to statistical analysis as described in 5.3.2 was used to compare the non-insulin and insulin groups in the modelling data.

6.4 RESULTS

The modelling data had 996 singleton pregnancies complicated by GDM; these were split, with 649 (70%) used to build and train the model, and 299 (30%) held out as test data to validate the model.

6.4.1 Participants

Comparing the treatment of women who received insulin with those who did not, **Table 6.1**, I observed in the maternal characteristics of the modelling data that those who required insulin had significantly higher age (non-insulin: mean[SD] 32 [5.3] years, insulin 33 [5.1] years, p=0.029), BMI (32 [7.3]kg/m^2, 35 [7.6]kg/m^2, p<0.001) and previous pregnancies (Gravida: 2.6 [1.6], 3.3 [2], p=0.003. Para: 2 [1.2], 2.3 [1.3], p=0.041). While other maternal characteristics were comparable. The diagnostic tests for GDM were significantly different; the fasting OGTT was higher in the insulin group (5.2 [0.5] mmol/L, 5.6 [0.56] mmol/L, p<0.001) and taken earlier (25 [5.9] weeks of gestation, 23 [5.8] weeks of gestation, p<0.001). HbA1c measured at booking was higher for those that required insulin (35 [3.3], 36 [3.8]mmol/L, p=0.008).

Table 6.1 *Modelling data participant characteristics*

Characteristic	**Non-Insulin** (mean[standard deviation] or N(%))	**Insulin** (mean[standard deviation] or N(%))	**p-value**
Total	921 (92.5%)	75 (7.5%)	
Gestational week at booking **Gestational day at booking**	10 [2.9] 3.1 [2]	10 [1.9] 3 [2.1]	0.782 0.736
Booking age	32 [5.3]	33 [5.1]	0.029**
Booking BMI[a] (kg/m^2)	32 [7.3]	35 [7.6]	<0.001**
Family origin of mother[1] **... Group A – White** **... Group F – Other ethnic group**	 560 (60.4%) 40 (4.3%)	 50 (62.7%) 10 (8.0%)	0.559
SIMD 2016 quintile **... 5** **... 1** **... Unknown**	 411 (44.6%) 108 (11.7%) 53 (5.8%)	 38 (50.7%) 8 (10.7%) 6 (8.0%)	0.616
Smoker at booking	81 (8.8%)	8 (10.7%)	0.737
Previous Caesarean birth	247 (26.8%)	25 (33.3%)	0.279
Number of previous Caesarean births	0.36 [0.68]	0.47 [0.74]	0.176
Gravida	2.6 [1.6]	3.3 [2]	0.003**
Para	2 [1.2]	2.3 [1.3]	0.041**
Previous macrosomia baby	*	*	1
Previous gestational diabetes	152 (16.5%)	15 (20.0%)	0.536
Any family history of diabetes	537 (58.3%)	42 (56.0%)	0.789
Mother's family history diabetes	449 (48.8%)	36 (48.0%)	0.996
Fasting glucose in OGTT[b] (mmol/L)	5.2 [0.5]	5.6 [0.56]	<0.001**
Gestational week of OGTT[b] **Gestational day of OGTT[b]**	25 [5.9] 3.3 [1.8]	23 [5.8] 3.5[1.8]	<0.001** 0.352
HbA1c[c] at booking (mmol/L)	35 [3.3]	36 [3.8]	0.008**

[a]BMI Body mass index, [b]OGTT Oral glucose tolerance test, [c]HbA1c glycosylated haemoglobin. 1. Frequency rounded to the nearest 10. Redacted due to the potential risk of disclosure.** Statistically significant, p<0.05.*

6.4.2 Feature selection

HbA1c at booking had 20% missing in the training data. As logistic regression cannot handle missing data, I dichotomised the variable using the Youden index. For HbA1c at booking, the optimal threshold was 37 mmol/L for predicting insulin with a Youden index of 0.21.

Variables available at the time of GDM diagnosis, listed in 6.3.3, and the dichotomised HbA1c at booking, were filtered using univariate logistic regression and selected if $p < 0.20$, **Table 6.2**, to identify potential insulin predictors. Candidate predictors that had high correlation with each other were removed.

Table 6.2 *Candidate variables selected by univariate logistic regression with a p-value <0.20*

Candidate Variable	p-value
Booking age	0.005
Booking BMI[a]	0.004
Family origin of mother (grouped)[b]	0.116
Gravida	0.064
Fasting glucose in OGTT[c]	<0.001
Gestational week of OGTT[c]	0.100
HbA1c[d] at booking < 37 mmol/L	0.051
HbA1c[d] at booking ≥ 37 mmol/L	0.034

[a]BMI Body mass index, [b]Ethnicity grouping (SMR02): Group A – White, Group B – Mixed or multiple ethnic groups, Group C – Asian, Scottish Asian or British Asian, Group D – African, Scottish African or British African, Group E – Caribbean or Black and Group F – Other ethnic group, [c]OGTT Oral glucose tolerance test, [d]HbA1c glycosylated haemoglobin.

As CART has an embedded feature selection, these candidate variables were then used in the model development. CART selected booking age, fasting glucose in the OGTT and the gestational week of the OGTT (detailed in section 6.4.3.2). For logistic regression, LASSO was applied to the candidate variables to identify the predictors. LASSO helps to identify the most important predictors and aims in improving the model's performance by reducing overfitting. The best lambda was 0.011, and as shown in **Table 6.2**, the

predictors for logistic regression, identified by LASSO, were: booking age, booking BMI, ethnicity of other ethnic group, fasting glucose in the OGTT and the gestational week of the OGTT.

Table 6.3 *Predictors for logistic regression selected after LASSO*

Variable	Lambda (regularisation parameter)
(Intercept)	-10.809
Booking age	0.039
Booking BMI[a]	0.019
Family origin of mother - Group B Mixed or multiple ethnic groups	0.00
Family origin of mother - Group C Asian, Scottish Asian or British Asian	0.00
Family origin of mother - Group D African, Scottish African or British African	0.00
Family origin of mother - Group E Caribbean or Black	0.00
Family origin of mother - Group F Other ethnic group	0.186
Gravida	0.00
Fasting glucose in the OGTT[b]	1.217
Gestational week of OGTT[b]	-0.009
HbA1c[c] at booking < 37 mmol/L	0.00
HbA1c[c] at booking ≥ 37 mmol/L	0.00

[a]BMI Body mass index, [b]OGTT Oral glucose tolerance test, [c]HbA1c glycosylated haemoglobin. Note that Family origin of mother - Group A White is left out and serves as a reference category.

The predictive power of the individual predictors of both models was assessed by the receiver operator curve (ROC) derived from univariate logistic regression predicting insulin, **Figure 6.3** . The fasting glucose in the OGTT had the greatest predictive power with an AUROC of 0.730 (95% CI: 0.658-0.802), followed by booking age and BMI with similar predictive power, 0.608 (0.529-0.689) and 0.633 (0.560-0.705), respectively. Other ethnic group had a predictive power almost as good as guessing, 0.526 (0.485-0.568), and the gestational week of the OGTT had a weak predictive power, 0.395 (0.324-0.465). However, this only looks at the predictors in isolation and does not account for confounding factors between the predictors.

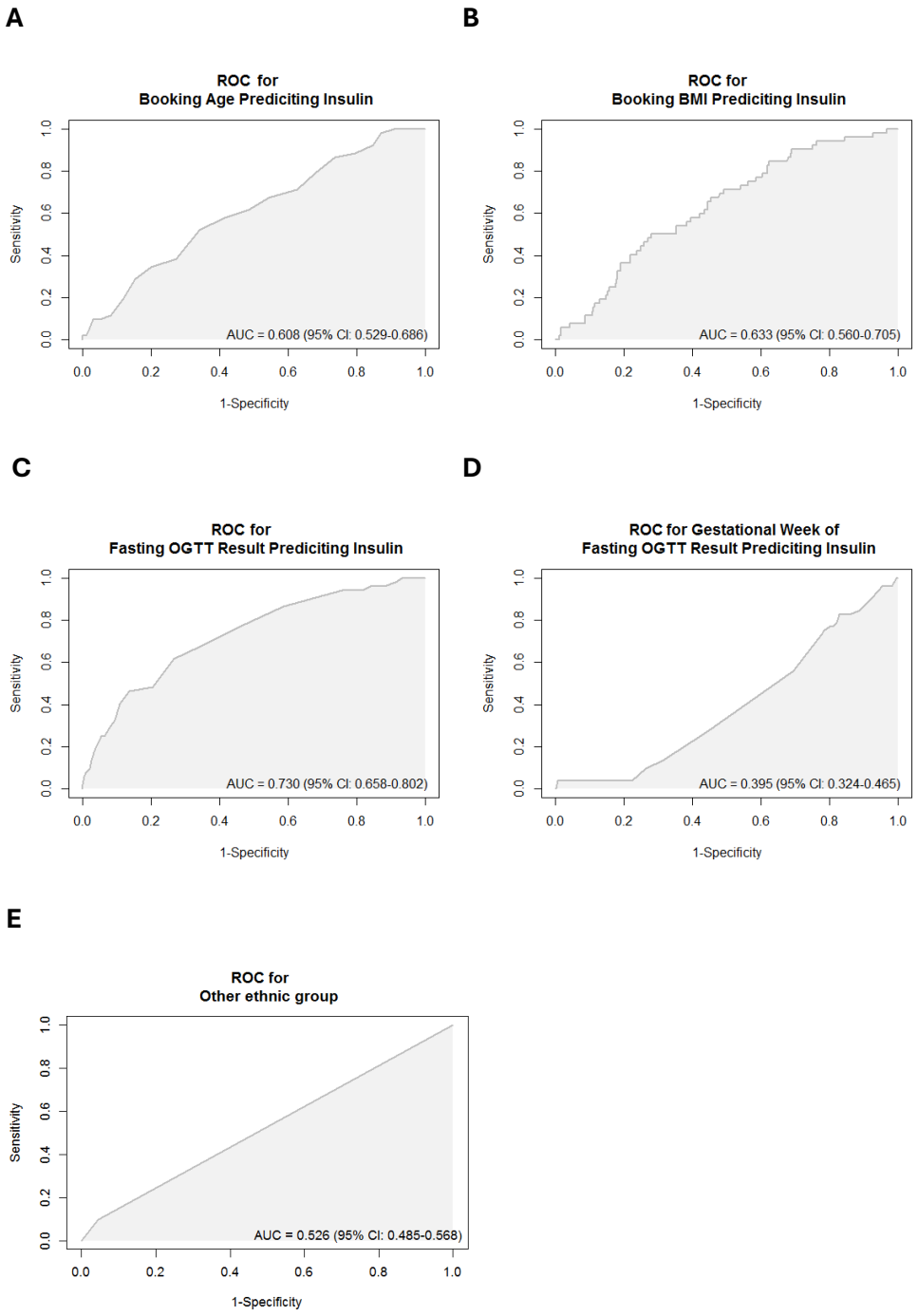


Figure 6.3 *Univariate logistic regression predicting insulin area under the receiver operator curve of selected predictors.*
Predictors A-E predictors for logistic regression model, predictors A, C and D predictors for classification and regression tree model.
AUC Area under curve, BMI Body mass index, CI Confident interval, OGTT Oral glucose tolerance test, ROC, Receiver operator curve.

6.4.3 Model development

6.4.3.1 Logistic regression

The predictors, booking age, booking BMI, other ethnic group, fasting glucose in the OGTT and the gestational week of the OGTT, were included in the development of the model. Optimal hyperparameters were identified as: penalty = 'l2', C = 0.00026366508987303583, solver = 'liblinear', and maximum iteration = 100. The final model was developed by evaluating the performance of 10-fold cross-validation, detailed in section 6.4.4. The logistic regression formula, shown in **Equation 6.1**, produces the percentage risk of requiring insulin to manage GDM; above 50% would be considered high risk of requiring insulin (Hosmer et al., 2013). An example of its use is shown in section 6.4.5.

$$Percentage\ insulin\ risk = \left(\frac{1}{1 + e^{-z}}\right) \times 100$$

$$\begin{aligned} z = -0.9510 &+ (0.0061 \times Booking\ age) \\ &+ (0.0040 \times Booking\ BMI) \\ &+ (0.1305 \times Fasting\ OGTT\ results) \\ &+ (-0.0027 \times Gestational\ week\ of\ the\ fasting\ OGTT\ result) \\ &+ (0.0919 \times Other\ ethnic\ group) \end{aligned}$$

Equation 6.1 *Logistic regression formula to give the percentage risk of requiring insulin therapy to manage GDM at the time of GDM diagnosis.*
BMI Body mass index, OGTT Oral glucose tolerance test.

6.4.3.2 CART

The candidate predictions, booking age, booking BMI, Family origin of mother Group A – White, Family origin of mother Group B – Mixed or multiple ethnic groups, Family origin of mother Group C – Asian, Scottish Asian or British Asian, Family origin of mother Group D – African, Scottish African or British African, Family origin of mother Group E – Caribbean or Black, Family origin of mother Group F – Other ethnic group, Gravida, Fasting glucose in the OGTT, gestational week of OGTT, HbA1c at booking < 37 mmol/L and HbA1c at booking ≥ 37 mmol/L were used to build a CART model. Optimal

hyperparameters were identified as maximum depth = 2, minimum sample leaf = 1 and minimum sample split = 2. The final model was developed by assessing the performance of 10-fold cross-validation, detailed in section 6.4.4. The final CART model, shown in **Figure 6.4,** included booking age, fasting glucose in OGTT and gestational week of OGTT as predictors. Gini impurity criterion indicates how well the model has split the data. The closer to zero the better the model is at distinguishing between classes on the training set, however, often when a model has a Gini of zero it is actually overfitting (Rebala et al., 2019). On the leaf node of the CART model, Gini is above zero I have pruned the tree early to prevent overfitting. An example of its use is shown in section 6.4.5.

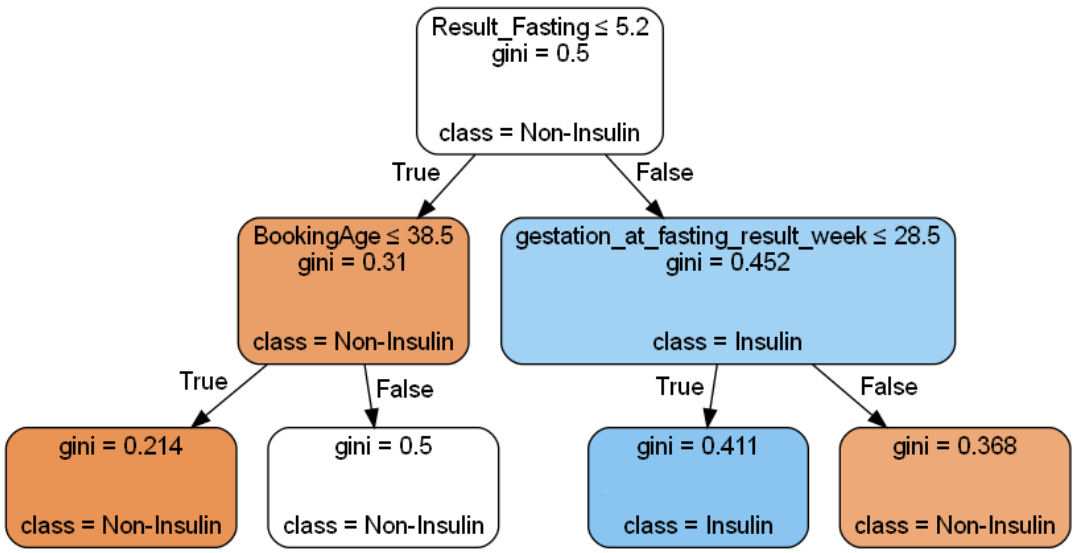


Figure 6.4 *CART Decision Tree for predicting insulin therapy to manage GDM at the time of GDM diagnosis.*

OGTT Oral glucose tolerance test. CART Classification and regression tree.

6.4.4 Model performance and evaluation

The performance of the models during the training 10-fold cross-validation, and validated on unseen test data, is summarised in **Table 6.4**.

During training, the logistic regression model has a mean AUROC of 0.77 [0.07], mean precision of 0.13 [0.04], mean recall of 0.88 [0.11] and a mean F1

score: 0.23 [0.06]. In comparison, during training, the CART model had a mean AUROC of 0.70 [0.10], a mean precision of 0.14 [0.06], a mean recall of 0.68 [0.22] and a mean F1 score: 0.23 [0.09].

Table 6.4 *Models' training and validation performance.*

Performance metric	**Logistic regression**		**CART[a]**	
	Training 10-fold cross-validation (mean [standard deviation])	**Validation on unseen test data**	**Training 10-fold cross-validation** (mean [standard deviation])	**Validation on unseen test data**
AUROC[b]	0.77 [0.07]	0.71	0.70 [0.10]	0.70
Accuracy	0.57 [0.04]	0.57	0.65 [0.08]	0.65
Precision	0.13 [0.04]	0.91* Non- Insulin: 0.97 Insulin: 0.13	0.14 [0.06]	0.91* Non-insulin: 0.97 Insulin: 0.15
Recall	0.88 [0.11]	0.57*	0.68 [0.22]	0.65*
Sensetivity (recall Insulin)		0.83		0.78
Specificity (recall non-insulin)		0.55		0.64
F1-score	0.23 [0.06]	0.67* Non- Insulin: 0.70 Insulin: 0.23	0.23 [0.09]	0.73* Non- insulin: 0.77 Insulin: 0.26

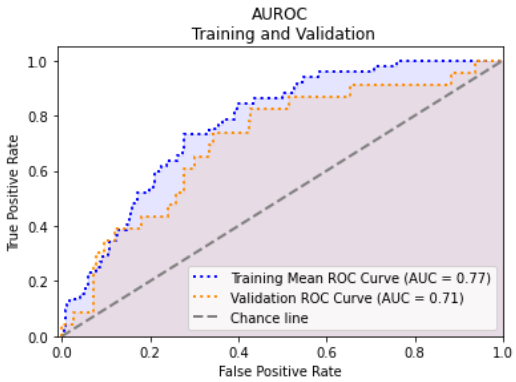
**Weighted average. [a]CART Classification and regression tree. [b]AUROC Area under the receiver operator curve.*

To validate the models, they were tested on unseen data. The AUROC of logistic regression was 0.71, which is a reduction from the training 10-fold cross-validation AUROC of 0.77, **Figure 6.5-1A**. The model achieved a weighted average precision of 0.91, with a precision for the non-insulin class of 0.97 and the insulin class of 0.13. A weighted average recall of 0.57, with a specificity (recall for the non-insulin class) of 0.55 and sensitivity (recall of the insulin class) of 0.83. And a weighted average F1-score of 0.67, with an F1-score for the non-insulin class of 0.70 and the insulin class of 0.23. The confusion matrix, **Figure 6.5-1B**, highlights that logistic regression correctly predicted 83% of insulin therapy cases, and only missed 17%. However, it does show that only 55% of those not requiring insulin were correctly identified, with 45% being incorrectly identified as requiring insulin. The model is suboptimal at distinguishing between the two classes but is effective at identifying those at risk of insulin therapy.

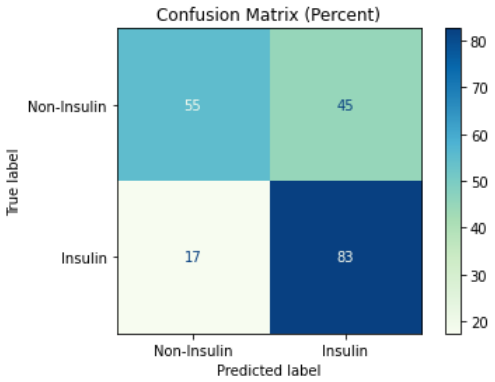
Testing the CART's performance on unseen data, its performance did not reduce and had an AUROC of 0.70, **Figure 6.5-2A**. A weighted average precision of 0.91, with a precision for the non-insulin class of 0.97 and the insulin class of 0.15. A weighted average recall of 0.65, with a specificity (recall for the non-insulin class) of 0.64 and a sensitivity (recall of the insulin class) of 0.78. And a weighted average F1-score of 0.73, with an F1-score for the insulin class of 0.77 and the non-insulin class of 0.26. From the confusion matrix, **Figure 6.5-2B**, CART correctly predicted 78% of the insulin class, missing 22%. However, it was better than the logistic regression in distinguishing between insulin and non-insulin and correctly identifying 64% of non-insulin.

Logistic regression

1A

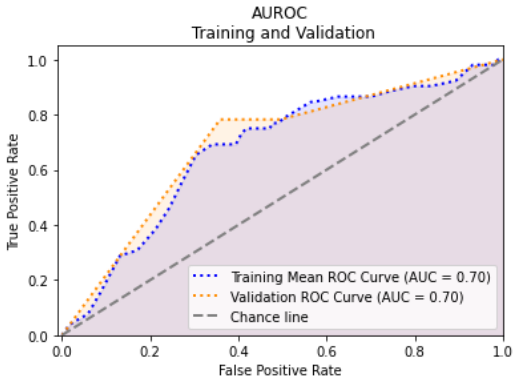


1B



CART

2A



2B

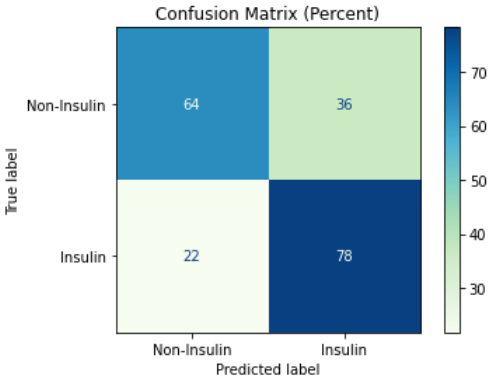


Figure 6.5 *Model performance.*
(1A-B) Logistic regression model, (2A-B) CART model.
(A) AUROC during training 10-fold cross-validation and validation testing on unseen data. Blue dotted line – AUROC of the training 10-fold cross-validation. Orange dotted line – validation AUROC of unseen test data. Grey dash line – chance line (AUROC = 0.5). The shaded area is the AUC (area under the curve).
(B) Confusion matrix with percentages and colour grading; the darker the colour, the higher the percentage. Taking insulin to be the positive class; Top left quadrant: true-negatives, top right quadrant: false-positives. Bottom left quadrant: false-negatives, bottom right quadrant: true-positives.
CART Classification and regression tree. AUROC Area under the receiver operator curve.

6.4.5 Example cases and clinical use

The logistic regression model and the CART model can give different levels of information; logistic regression can give a percentage, and then, commonly, the threshold of 50% used to classify the cases, whereas CART gives only a classification. An example of how the information could be imputed by an HCP is shown in **Figure 6.6**. To show the use of the models, let's look at two hypothetical cases.

Given woman A, who has just been diagnosed with GDM, is 37 years old with a BMI of 40kg/m^2, she has an SMR02 ethnicity of Group D – African, Scottish African or British African and has a fasting OGTT result of 6.7 mmol/L taken at 29 weeks of gestation. The logistic regression formula would say that woman A has a 52% chance of requiring insulin to manage her GDM, and given the threshold of 50%, would be classified as requiring insulin. Similarly, the CART model would also classify her as at risk of needing insulin.

Women B, who has also just been diagnosed with GDM, is 30 years old with a BMI of 29kg/m^2. She has an SMR02 ethnicity of Group C – Asian, Scottish Asian or British Asian, and a fasting OGTT of 5.2mmol/L taken at 28 weeks of gestation. The logistic regression model says she has a 49% chance of requiring insulin and would be classified as non-insulin. Similarly, the CART model would also classify woman B as non-insulin.

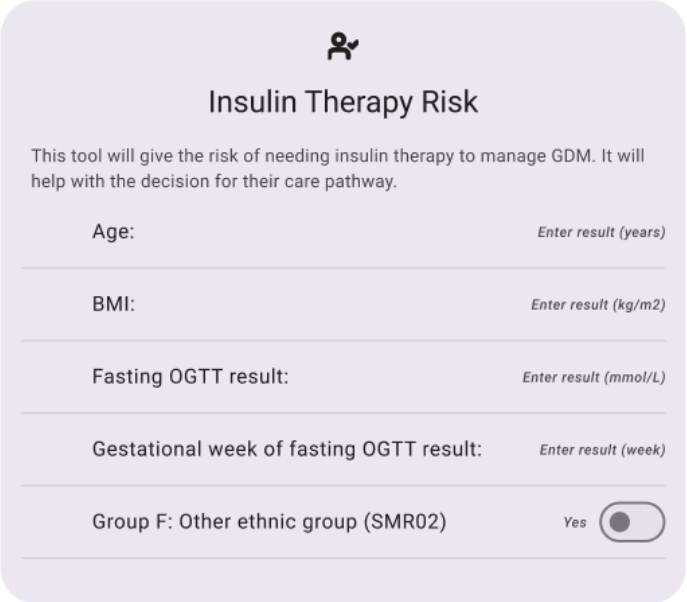


Figure 6.6 *Example of insulin therapy risk prediction input.*

BMI Body mass index, OGTT Oral glucose tolerance test, SMR02 Scottish morbidity record - maternity interaction/stay.

6.5 Discussion

The models are designed to be used to provide HCPs with additional information at the time of GDM diagnosis, so that they can make a more informed decision on their patient care pathway; whether they require a more intensive monitoring and appointments, similar to current care, or could be seen primarily through remote monitoring and have fewer in-person appointments.

The variables included in the logistic regression model were booking age, booking BMI, fasting OGTT result, gestational age at the fasting OGTT result and SMR02 Group F other ethnic groups. And for the CART model, they were fasting OGTT result, booking age and the gestational age at fasting OGTT. These were mostly expected variables, apart from other ethnic groups. It is possible that there were some confounding factors within the data which caused other ethnic groups to be predictive variables.

As the model aims to support HCPs treatment decisions, recall is a particularly important metric to assess the model's performance. Recall

describes how successful the model was at identifying a particular class; in this instance, I wanted to maximise the recall for the insulin class (sensitivity). Logistic regression had a better sensitivity, 0.83, than CART, 0.78.

However, although the recall was high for the insulin class (sensitivity), the precision was poor for both models, which affected the F1-score negatively. The low precision indicates that the models generate many false positives, as shown in the confusion matrix (**Figure 6.5**), with 45% of non-insulin cases being misclassified by logistic regression and 36% by CART. Since the data was unbalanced with a 7.5% prevalence of insulin, SMOTE was employed to balance it. This improved the models' sensitivity to the minority class, insulin, but also introduced noise, which could explain why the models misclassify the non-insulin class and consequently have low precision.

Examining the overall performance of the models, logistic regression shows a greater decline in performance on the unseen data, with the AUROC decreasing from 0.77 to 0.71, while CART remained more stable, with an AUROC of 0.70 in both training and testing. The AUROCs for both models on the unseen data are reasonable and indicate that the models were able to generalise well to the unseen data.

Compared to the previous literature within this area, presented in Chapter 3. I have also used simple interpretable algorithms; however, I had a smaller dataset of 996 pregnancies collected for 1.75 years in comparison to the literature, with a median (range) of 1919 pregnancies (37 - 30,474), and a data collection period of 5 years (1-23 years). The median AUROC of the literature was 0.75 (0.61-0.93). Logistic regression predicting insulin had a median AUROC of 0.76 (0.71-0.87), and CART predicting insulin had an AUROC of 0.74 (only one model). The logistic regression model that I built had a similar predictive value during training (logistic regression 0.77 [0.07]), whereas the CART did not perform as well (0.70). Furthermore, the validation AUROC obtained here (Logistic regression: 0.71, CART: 0.70) is comparable to the literature, 0.72 (range 0.59-0.82). None of the models in the literature accounted for their class imbalance, whereas I did, and used SMOTE.

However, my cohort had a more severe class imbalance, 7.5%, than those in the literature, which had a median of 38.8% (10.8-69.8%) in the predictor class. Similarly to the literature, my models have not been clinically implemented; given the time frame of this PhD and the requirement for ethical approval, it would not have been possible. But I did build the models to integrate them into a digital tool (Chapter 7).

In conclusion, both models are sensitive (recall for the insulin class is high) but not specific (precision is low); they can generalise reasonably well, with similar AUROC on the unseen data of both models. Given that the purpose of the model is to risk-stratify care, I believe that the logistic regression model is the better of the two. The AUROC was similar, and CART was better at distinguishing between the two classes. However, the logistic regression model had a higher recall and, therefore, was better able to identify the risk of insulin, which is important as missing those that do require insulin is more costly than overclassifying those that do not end up requiring it. In addition, the logistic regression model gives a risk percentage rather than a binary yes/no. This provides more information and a better understanding of how high or low the risk of requiring insulin would be, and therefore, I believe it is more appropriate for clinical use, where many confounding factors need to be considered when providing care and not just the risk of insulin therapy.

6.5.1 Strengths and limitations

The models' strengths lie in being developed using clinically collected data from EHR, which reduces data collection bias; however, I had to estimate pregnancies complicated by GDM due to inconsistent recording of its instances (Chapter 4). The models use readily available clinical data, requiring no additional processing steps, and are therefore easy to implement in clinics. The data imbalance was considered, and SMOTE was used to rebalance the data during training. The models are limited by a small dataset, potentially embedding historical bias into the model. Despite planning to externally validate them temporally with a later cohort of pregnancies from Glasgow and geographically with pregnancies in Edinburgh, the data were not available

within the timeframe of my PhD. And hence were limited in that they were only internally validated within the same location (Glasgow). Without external validation, it limits the confidence in the generalisability of the models. Therefore, the findings highlight the promise and the immaturity of the models. Since collecting the data, the guidelines for diagnosing GDM have changed from a fasting OGTT threshold of 5.3mmol/L to 5.6mmol/L (Scottish Intercollegiate Guidelines Network, 2024). As a result, the models may not generalise to these new guidelines and would need recalibration.

6.6 CHAPTER SUMMARY

In this chapter, I have presented the development of two machine learning models that predict, at the time of GDM diagnosis, the risk of needing insulin therapy to manage GDM. These models, logistic regression and CART, were developed from a data set of 996 singleton pregnancy episodes complicated by GDM in Greater Glasgow and Clyde that had a booking date between 1st April 2022 and the 31st December 2023, with 7.5% requiring insulin. The data was split; 697 (70%) was used to develop and train the models, and 299 (30%) was held out to validate the models by testing on unseen data. During training, the models' performance was evaluated through 10-fold cross-validation. SMOTE was applied to address the data imbalance. Both models generalised reasonably well on the unseen data; logistic regression had a validation AUROC of 0.71 and CARTs had a validation AUROC of 0.70. Logistic regression was better able to identify the insulin class and correctly classified 83% of the insulin class in comparison to 78% by CART.

The model aims to risk-stratify GDM care by identifying women with GDM who are at high risk of needing insulin therapy. If there is a high risk of insulin therapy, then care could involve more intensive monitoring and appointments, similar to current practice, whereas if there is a low risk, care could mainly consist of remote monitoring with fewer in-person visits. This monitoring and sharing of blood glucose readings would be conducted through a digital tool,

comprising a patient-facing app and a clinical dashboard, which is described in the next chapter, Chapter 7.

CHAPTER 7

The User-Centred Design and Evaluation of a Digital Tool For GDM

7.1 Introduction

As found in Chapter 3, the models reviewed from the literature often did not allude to how they could be clinically implemented or used. One (Souza et al., 2019), described a nomogram for clinical interpretation and another (Velardo et al., 2021), indicated that it could potentially be integrated into an existing mHealth app. It is envisioned that the model presented in Chapter 6 could streamline GDM care by being clinically implemented through a digital tool. In addition, a digital tool could also be used in the care and management of GDM. It has the potential to help HCPs provide care while also empowering the self-management of women with GDM. In this Chapter, I will describe the development of 'MyGDM', a clinical dashboard and patient-facing app.

Digital tools, such as mHealth or telemedicine, have been reported to have the potential to streamline GDM care, whilst not compromising maternal and neonatal outcomes (Rasekaba et al., 2015), by reducing in-person appointments (Rasekaba et al., 2015) and facilitating better and easier self-management of blood glucose and diabetes (Birati et al., 2022; Sakamoto et al., 2022; Khalil, 2019). Positive satisfaction levels of HCPs and GDM women

who use smartphone apps for GDM management have been reported (Smyth et al., 2022). mHealth can remove barriers for pregnant women such that they can access services and health information (Sakamoto et al., 2022). The application of artificial intelligence in digital tools can improve care in resource-poor settings (Saif-Ur-Rahman et al., 2023).

Personalisation of care has a more beneficial effect on achieving target blood glucose levels in GDM (Guo et al., 2023). However, there are a distinct lack of digital tools that have been developed for GDM self-management that incorporate computer decision support or artificial intelligence (Daley et al., 2021).

There are few established mHealth tools for GDM that include end-users from the early stages of development, though there are some in development (Safiee et al., 2024; Shanmugavel et al., 2024; Duan et al., 2024). Including users early improves the effectiveness of the results (Chamberlain et al., 2022).

Therefore, I aimed to design a digital tool for GDM through a user-centred approach, by first assessing the end-users' (HCPs and women with GDM) needs for a digital tool through semi-structured interviews. Following this, an initial design of a prototype was then evaluated through feedback sessions with HCPs and researchers, and questionnaires with women with GDM.

7.2 AIMS AND ENDPOINTS

This Chapter will focus on two objectives: IV Assess the needs, acceptability and barriers around digital tools in GDM with women who have GDM and HCPs who care for them. And V Design and evaluate a user-centred digital tool for GDM care and management. The specific aims and endpoints of this Chapter are:

7.2.1 Aims

1) The user needs assessment:

   a. To understand the current management of GDM from an end-user's perspective and to identify current barriers and facilitators of care.
   b. To understand what technology is currently being used in care and what end-users would want from a GDM-specific digital tool.
2) Design a digital tool that meets end-users' needs and desires as identified in the user-needs assessment.
3) Evaluate the design of the digital tool and identify areas for improvement with end-users.

7.2.2 Endpoints

1) An understanding of what end-users want from a digital tool in GDM.
2) A digital tool prototype that has been initially evaluated and can be further developed into a clinically viable product.

7.3 CHAPTER STRUCTURE

This chapter presents the journal article entitled 'The User-Centered Design of a Clinical Dashboard and Patient-Facing App for Gestational Diabetes' (Kirkwood et al., 2024), published in the Journal of Diabetes Science and Technology in November 2024. This work was completed with MyWay Digital Health. The user needs assessment study is further detailed in the publication entitled 'User needs assessment for gestational diabetes care and management digital tool: a qualitative study', (Kirkwood et al., 2025a), published at the 35th Medical Information Europe Conference 2025, which focused on end-users' experience of care and what they would want from a digital tool (**Appendix H**, including author contribution).

To aid with readability, the accepted manuscript text of the published article (Kirkwood et al., 2024) has been formatted into the style of this thesis (section 7.4), and the supplementary material can be found in **Appendix I**. The structure and content of the article have not been altered and were written with American English. Permission has been granted for the re-use of this

article and its content by the publishers Sage and MyWay Digital Health. The ethical approval letter for the user-needs assessment study is in **Appendix J**. The authors' contributions to the publication are provided in **Table 7.1** using CRediT (CRediT, 2022).

Table 7.1 *Author's contributor role, following CRediT, for the publication of The User-Centered Design of a Clinical Dashboard and Patient-Facing App for Gestational Diabetes, Journal of Diabetes Science and Technology.*

Contributor Roles	**Author's initials**
Conceptualisation	JRK, RMR, DJW, RSL, AM
Data curation	JRK
Formal analysis	JRK, JD
Funding acquisition	RMR, DJW, RSL, AM
Investigation	JRK
Methodology	JRK
Project administration	JRK
Resources	RMR, DJW, RSL, MS
Software	Not applicable
Supervision	RMR, DJW, RSL, AM
Validation	JD
Visualisation	JRK, MS
Writing – original draft	JRK
Writing – review & editing	RMR, DJW, RSL, AM, JD, MS

JRK Jasmine R Kirkwood, JD Jane Dickson, MS Marryat Stevens, AM Areti Manataki, RSL Robert S Lindsay, DJW Deborah J Wake, RMR Rebecca M Reynolds

7.4 The User-Centered Design of a Clinical Dashboard and Patient-Facing App for Gestational Diabetes [Published in the Journal of Diabetes Science and Technology]

The User-centered Design of a Clinical Dashboard and Patient-facing App for Gestational Diabetes

Jasmine R Kirkwood, BSc (Hons)[1], Jane Dickson, PhD[2], Marryat Stevens, PhD[3], Areti Manataki, PhD[4], Robert S Lindsay, BSC, MBChB, PhD, FRCP (Glasg)[5]Deborah J Wake, PhD, MBChB, BSc (Hons), Dip (Med Ed)[3,6], Rebecca M Reynolds, MA, PhD, BM BCh, FRCP, FRCPE, FMedSci[1]

Author Affiliations: [1]Centre for Cardiovascular Science, University of Edinburgh, [2] University of Dundee Medical School, [3]My Way Digital Health, Dundee, [4]School of Computer Science, University of St Andrews, [5]School of Cardiovascular and Metabolic Health, The University of Glasgow, [6]Usher Institute, The University of Edinburgh.

List of abbreviations: (GDM) Gestational diabetes mellitus, (HCPs) Healthcare professionals, (SMBG): Self-Monitoring Blood Glucose.

Keywords: Digital health, gestational diabetes mellitus, health education, mHealth, user-centered design.

Corresponding Author: Jasmine R Kirkwood, Centre for Cardiovascular Science, Queen's Medical Research Institute, The University of Edinburgh, EH16 4TJ, UK, J.R.Kirkwood@sms.ed.ac.uk

Funding Sources: JRK is funded by Medical Research Scotland (PHD-50224-2020). RMR acknowledges the support of the British Heart Foundation (RE/18/5/34216).

Conflict of Interest Disclosures

DW is the CEO, co-founder, and shareholder of MyWay Digital Health. MS is the Product Manager at MyWay Digital Health and own Share Options.

Acknowledgments

JRK is grateful for being able to complete this work though the Medical Research Scotland PhD Scholarship. JRK would also like to thank MyWay Digital Health, who part sponsored this work. Experts at MyWay Digital Health supported with the design development, the development of GDM-specific educational resources, language options and culturally inclusive meal ideas, to which JRK is most appreciative of.

Abstract

Background: The number of pregnancies affected by gestational diabetes mellitus (GDM) is growing. With the increased use of smartphones and predictive modelling, a mobile health (mHealth) solution could be developed to improve care and management of GDM whilst streamlining care through risk stratification.

Methods: A user-centered mHealth tool was designed from ethnographic observations and 11 semi-structured interviews (6 healthcare professionals and 5 women with GDM), followed by iterative changes and evaluation from three feedback group with 31 participants (17 healthcare professionals, 14 researchers) and 13 questionnaires with women with GDM.

Results: "MyGDM" includes a clinical dashboard that centralizes the clinic's patients, highlighting off-target blood glucose and predicting the need for pharmacological intervention. It is linked with a patient-facing app, that includes structured education, culturally inclusive language options, and meal ideas. Through the feedback sessions, iterative changes were made around visualization and patient safety, and participants were positive towards the potential user experience. In the 13 questionnaires with women with GDM, 100% said it would fit into their lifestyle and help them manage GDM. Educational resources and the 'request a call' functions were well

received with 61.5% (8/13) and 69.2% (9/13) saying they were very likely or likely to use these, respectively.

Conclusion: A user-centered mHealth tool consisting of a clinical dashboard linked with a patient-facing app for GDM care and management has been designed. Evaluation of the interactive design by end-users was positive and showed that it met their needs.

7.4.1 Introduction

Gestational diabetes mellitus (GDM) is hyperglycaemia that is first recognized during pregnancy (World Health Organisation (WHO) & International Diabetes Federation (IDF), 2021, 13 April; National Institute for Health and Care Excellence (NICE), 2015, 25 February ; 2021). GDM is one of the most common pregnancy complications (Hivert et al., 2024), affecting approximately 5.8% of pregnancies in Europe (Zhu and Zhang, 2016) and up to 30% globally (White et al., 2023). It is expected to increase due to the rise in obesity (Ferrara, 2007; Johns et al., 2018) and maternal age (Johns et al., 2018). Women who had GDM have a ten-fold increase in developing type 2 diabetes (Vounzoulaki et al., 2020), and a two-fold increase in developing early cardiovascular disease (Carr et al., 2006; Daly et al., 2018), with their offspring at a greater risk of childhood obesity and diabetes (National Institute for Health and Care Excellence (NICE), 2015, 25 February ; 2021; Buchanan et al., 2012).

There is a short period between diagnosis of GDM and birth where a woman with GDM must comprehend the consequences of the diagnosis, whilst managing the additional burden it places on their pregnancy. This places considerable stress on the women during their pregnancy (Fraser et al., 2023; Draffin et al., 2016; Nazarpour et al., 2024). In addition, care is time-consuming, with women with GDM spending over 2 hours attending a multidisciplinary antenatal diabetes clinic, every 1-2 weeks, excluding travel time (Alqudah et al., 2019). As such, remote monitoring and remote consultation could potentially be beneficial. Clinical needs are diverse and therefore there is the potential to use risk stratification to identify those likely

to need pharmacological therapy to streamline care to those most in need of regular face-to-face clinic review (Benham et al., 2023b).

mHealth is any service provided through a mobile device within a healthcare setting. mHealth could be a good addition to standard care for GDM (Mackillop et al., 2018), as it allows patients to be more involved with their treatment and is not restricted to a time or place (Van Den Heuvel et al., 2018). The use of smartphone apps and online dashboards for GDM remote monitoring has been found to increase women's satisfaction, confidence, and knowledge regarding GDM, and to optimize clinical time without worsening glycaemic control (Bertini et al., 2022). Such tools could lead to personalization of care with the potential for rapid integration into the clinic, as smartphone use is high among pregnant women with diabetes (Alqudah et al., 2019).

Currently, most mHealth apps for GDM mainly report patient information, alerts around medication and appointments or education, (Daley et al., 2021), however, there is a lack of GDM-specific information or courses (Garg et al., 2022), and a limited number including artificial intelligence or machine learning (Daley et al., 2021). Furthermore, to our knowledge, none of the clinically used apps have been designed with their end-users (Garg et al., 2022). Hence, there is a gap, to develop with end-users, an mHealth tool for GDM care and management, to help women to self-manage their GDM, and with the potential to personalize care through pharmacological therapy risk stratification.

We aimed to conduct a user-centered design and evaluation of a potential mHealth tool to facilitate the management of GDM.

7.4.2 Methods

We used a user-centered approach (Farao et al., 2020), in which we included the end-users (healthcare professionals (HCPs) and women with GDM) in the design and evaluation of a new mHealth tool. The interactive design mock-up

is a clinical dashboard linked to a patient-facing app that was developed through four stages, shown in **Figure 7.1**.

7.4.2.1 User needs assessment

The aim of the user needs assessment was to understand the views, options, and acceptance of digital technologies within GDM care and management, and what end-users would want from such technologies. The results then informed the initial design. To familiarize with the setting, firstly, four antenatal diabetes clinics in three hospitals in Edinburgh and Glasgow were observed between September - December 2022. This was followed by 11 (HCP n=6, women with GDM n=5), semi-structured interviews that took place between January - July 2023 (topic guide in supplementary material 1, **Appendix I**). Semi-structured interviews allow for an in-depth exploration of the research topics and a greater understanding of the users' wants (Gubrium et al., 2012), hence a minimum of six to ten participants is deemed sufficient (Terry and Hayfield, 2021).

Participants were recruited from flyers and emails in the antenatal diabetes clinics in Edinburgh and Glasgow. Written and verbal consent was obtained before interviews which were transcribed by a third-party, 1st Class Secretarial Services (https://www.1stclass.uk.com/) or a researcher. Transcripts were coded using NVivo (12 software V3). Thematic analysis of transcripts was carried out using the 6-step framework described by Braun and Clarke (Braun and Clarke, 2006): (1) Familiarization with the data, by reading through transcripts; (2) coding the transcripts with 10% double-coded (JRK and JD); (3) searching for themes by reviewing the codes and summarizing the interviews; (4) reviewing themes and (5) defining and (6) naming themes. Codes and themes were refined through discussions throughout.

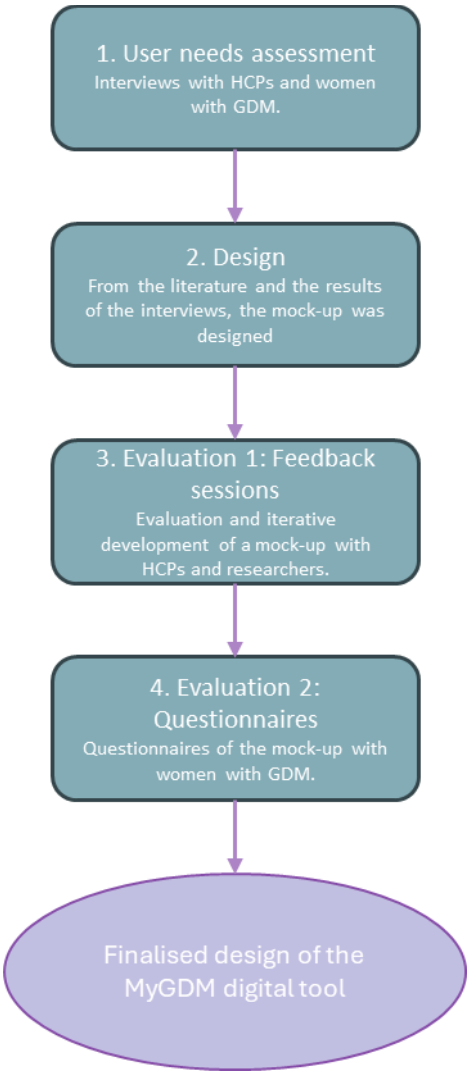


Figure 7.1 *Study design schematic*

(HCPs: Health care professionals, GDM: Gestational diabetes mellitus)

7.4.2.2 Design

From the literature (Safiee et al., 2023; Dalton et al., 2018; Duan et al., 2022; Blair et al., 2022; Akalpler and Bağriaçik, 2023; Ghaderi et al., 2022; Lazarevic et al., 2023; Safiee et al., 2022; Shanmugavel et al., 2024) and the results of the user needs assessment, an interactive design, consisting of a clinical dashboard for HCPs linked with a patient-facing app was developed on Figma (www.figma.com/prototyping/). The aim was to design a tool that met our end-users' needs and desires.

7.4.2.3 Evaluation

The mHealth tool was evaluated through feedback sessions with HCPs and researchers and questionnaires with GDM women. The aim was to appraise and identify areas of improvement with end-users.

7.4.2.4 Feedback sessions

Three feedback sessions of 31 participants (17 HCPs and 14 researchers) from Edinburgh and Glasgow were recruited through the researcher team's network (RMR and RSL). Feedback sessions took place between January - February 2024 and were either hybrid or online with a mean (standard deviation) duration 32 (3.9) minutes. Verbal consent was gained before each session. During the session, the clinical dashboard and patient-facing app were demonstrated, followed by an open discussion, (topic guide in supplementary material 2, **Appendix I**). After each session, the automatic Teams transcripts were reviewed against the video, alongside notes taken. Iterative changes to the interactive design mock-up were made and then presented to the next feedback session. After the final session, transcripts were thematically analyzed, again using the Braun and Clarke framework (Braun and Clarke, 2006).

7.4.2.5 Questionnaires

Women with GDM were recruited between February – April 2024, at antenatal diabetes clinics in Edinburgh (n=12, 92.3%) or, if they had been previously interviewed and agreed to participate in the questionnaire stage (n=1, 7.7%). Data saturation was reached after 13 questionnaires, where no more themes emerged. Verbal consent was gained, and then participants were shown a short video (2:48 minutes) demonstrating the patient-facing app and asked to complete a questionnaire (Online Surveys, https://www.onlinesurveys.ac.uk/) (supplementary material 3, **Appendix I**). The questionnaire enquired about the app's potential usefulness, options on features and areas of improvements. A short questionnaire was chosen to reduce the burden on participants. Themes of free-text questions were identified by reading through answers and grouping

into similar topics and then naming themes. Basic descriptive statistics were completed on Microsoft Excel (version 2405).

7.4.2.6 Participant inclusion criteria

Participants were included if they were (1) ≥16 years old, (2) had given consent to take part in the study, (3) able to read and write, (4) willing to use their email address for correspondence. Women with GDM (a) had GDM/previous GDM and were treated in Edinburgh or Glasgow within the last two years, (b) were not pre-diabetic or had pre-gestational diabetes. HCPs had (a) worked in GDM care in Edinburgh/Glasgow in the last two years, and (b) at least one year of experience within GDM care. Participants in the feedback sessions had worked or researched pregnancy care in Edinburgh or Glasgow in the last two years and had an understanding of GDM.

7.4.2.7 Ethics

The study received ethical approval from the Research Ethics Committee (reference number 22/NW/0164). The study was co-sponsored by NHS Lothian and the University of Edinburgh through the Academic and Clinical Central Office for Research and Development (ACCORD).

7.4.3 Results

7.4.3.1 User needs assessment

The characteristics of the 11 participants who were interviewed for the user-needs assessment are described in **Table 7.2**.

Table 7.2 *Characteristics of Participants in the User Needs Assessment*

Health Care Professionals (N=6)		
Variable	Value	Number (percentage response, %)
Location	Edinburgh Glasgow	4 (66.7%) 2 (33.3%)
Role	Consultant endocrinologist/ physician Consultant obstetrician Dietitian Diabetes specialist nurse	3 (50.0%) 1 (16.7%) 1 (16.7%) 1 (16.7%)
	Women with Gestational Diabetes (N=5)	
Variable	Value	Number (percentage response, %)
Location	Edinburgh	5 (100.0%)
Age (years)	21-25 31-35 36-40	1 (20.0%) 1 (20.0%) 3 (60.0%)
Ethnicity	White Mixed Black/African/Caribbean/Black British	2 (40.0%) 2 (40.0%) 1 (20.0%)
Gestational diabetes diagnosis before the interview	8-14 days ago 3-6 months ago 7 to 12 months ago	1 (20.0%) 3 (60.0%) 1 (20.0%)
Gestational diabetes in a previous pregnancy	Yes No, first pregnancy	4 (80.0%) 1 (20.0%)
Treatment for gestational diabetes at the time of the interview	Metformin Insulin Metformin and insulin	3 (60.0%) 1 (20.0%) 1 (20.0%)

Four themes with 13 subthemes emerged from the interviews, shown in **Table 7.3** alongside illustrative quotes. Participants expressed a desire for greater availability of information on GDM and diet, including the need for it to be culturally inclusive and credible. Additionally, a need to improve the identification of issues through highlighting and notification was expressed, with the potential to streamline care through risk prediction. However, concerns over the safety of predictions were also raised. Finally, time constraint was a main barrier; it is time-consuming for HCPs to manage the growing number of women with GDM, whilst women with GDM are required to fit in and attend additional appointments on top of their other commitments.

The results of the interviews were then used to develop features for the initial design, in accordance with available literature (Safiee et al., 2023; Dalton et al., 2018; Duan et al., 2022; Blair et al., 2022; Akalpler and Bağriaçik, 2023; Ghaderi et al., 2022; Lazarevic et al., 2023; Safiee et al., 2022; Shanmugavel et al., 2024).

Table 7.3 *Themes, subthemes, and illustrative quotes from the semi-structured interviews with healthcare professionals and women with gestational diabetes.*

Quotes have been identified with participant type and number, and nonrelevant sections have been denoted […]. (HCPs: Healthcare professional, GDM: Gestational diabetes mellitus).

Theme **Frequency, percentage (%)**	**Subtheme** **Frequency, percentage (%)**	**Illustrative quote**
Time constraints **Overall: 10 (90.9%)** **HCPs: 6 (100.0%)** **Women with GDM: 4 (60.0%)**	Time management and workload of HCPs Overall: 4 (36.4%) HCPs: 4 (66.7%) Women with GDM: 0	"Something that could be done in a completely different way and free up time to allow us to work more efficiently and utilize the time in a different way." (HCP 3)
	Patients' parental, work, and life commitments Overall: 8 (72.7%) HCPs: 4 (66.7%) Women with GDM: 4 (80.0%)	"I've no time. Most of the time I have just enough time in my brain to remember to do my blood sugar, and then it's back into the kids." (GDM Woman 2)
Barriers to current GDM care **Overall: 9 (81.8%)** **HCPs: 6 (100.0%)** **Women with GDM: 3 (60.0%)**	Technical issues Overall: 7 (63.6%) HCPs: 5 (83.3%) Women with GDM: 2, (40.0%)	"(blood glucose monitor) sometimes there's IT issues. So not everybody shares their data." (HCP 6)

	Appointment attendance Overall: 7 (63.6%) HCPs: 4 (66.7%) Women with GDM: 3 (60.0%)	"With my job, it's difficult to take time off in general. So, it's quite difficult when they're like, this one is going to be on a computer, and you're like, well, I need to get cover, [...] It's just an added in extra." (GDM Woman 1) "We get lots of DNAs (did not attend) for people that just go like, well, I am just going to ignore this." (HCP 4)
	Information, resources, and education Overall: 4 (36.4%) HCPs: 2 (33.3%) Women with GDM: 2 (40.0%)	"The problem is that at booking so much information has to be given out at booking and they (woman with GDM) are bombarded." (HCP 4)
	Cultural, social, and language Overall: 6 (54.5%) HCPs: 6 (100.0%) Women with GDM: 0	"There may be cultural barriers as well in terms of what they're happy to do and understanding sharing of information." (HCP 1)
Suggestions for digital tool **Overall: 11 (100%)** **HCPs: 6 (100%)** **Women with GDM: 5 (100%)**	Highlighting off-target blood glucose readings Overall: 5 (45.5%) HCPs: 3 (50.0%) Women with GDM: 2 (40.0%)	"Dashboard and then you just see who needs to be contacted, that would be amazing." (HCP 5)

	Notification and prompts Overall: 9 (81.8%) HCPs: 6 (100.0%) Women with GDM: 3 (60.0%)	"Kind of prompt or troubleshoot it quickly, and be like, try this. Because as soon as I put in evening snack, right away it helped." (GDM Woman 1)
	Predictions and risk stratification Overall: 8 (72.7%) HCPs: 6 (100.0%) Women with GDM: 2 (40.0%)	"I'm sure you could be predicting that (medication) a bit earlier. It's not that we don't get there, we do, but we could probably be narrowing the time and improving the delivery of care in that way." (HCP 2)
	Diet management Overall: 5 (45.5%) HCPs: 2 (33.3%) Women with GDM: 3 (60.0%)	Recipes. [...] Some little quick and easy, like, here's something that you can change in your cooking. Take that recipe that you love and if you just do this." (GDM Woman 1)
	Information and education Overall: 5 (45.5%) HCPs: 4 (66.7%) Women with GDM: 1 (20.0%)	"It would be really good to have lots of really good information all in the one place." (HCP 4)
Safety and trust in data-driven solutions	Information produced by digital tools	"A process where you were very confident that the information going back to the patient was accurate and appropriate." (HCP 2)

Overall: 4 (36.4%) **HCPs: 4 (66.7%)** **Women with GDM: 0**	Overall: 3 (27.3%) HCPs: 3 (50.0%) Women with GDM: 0	
	Predictions made by digital tools Overall: 3 (27.3%) HCPs: 3 (50.0%) Women with GDM: 0	"Any predictive tool should be in conjunction to clinical care and that really the last signoff should be with the doctors to oversee whether or not the algorithms are making sense almost to each individual patient." (HCP 1)

7.4.3.2 Feedback sessions

There were three feedback sessions, with a total of 31 participants (17 HCPs and 14 researchers). After each feedback session iterative changes were made and presented to the next group, (feedback session participant characteristics and iterative changes, supplementary material 2, **Appendix I**).

Clinical adaptation and the practicality of using the digital tool were discussed. This included the ability to set up clinic-wise Self-Monitoring Blood Glucose (SMBG) targets for all women with GDM, whether these readings would be taken pre- or post-prandial, and the option to remove features that do not suit the clinic's practices. Further, there were discussions about who would be responsible for checking the dashboard and how a patient's request for a call would be notified. For example, there was the suggestion to have an email notification or to be able to filter the dashboard for call requests. Patient safety is key and was raised in many different ways throughout the feedback sessions. An indicator was added to identify if SMBG had been edited or added manually by patients, this was to allow transparency around SMBG and to ensure that clinical decisions could be made accurately. Additionally, there were concerns over misinformation spreading easily in the 'GDM Community' forum, which was incorporated following recommendations from literature (Safiee et al., 2023), as it would be unregulated. Visualization of the patient's SMBG through graphs and log-book formatting to simplify reviewing was discussed. All feedback sessions were positive regarding the aesthetics and potential user experience.

7.4.3.3 Questionnaires

In total, 13 women with GDM completed the questionnaires, most were 31-40 years old, white, diagnosed 7-9 months ago, had children, did not have previous GDM, and were being treated using diet or metformin, (characteristics of questionnaire participants, supplementary material 3, **Appendix I**).

The responses to the quantitative questions are shown in **Table 7.4**. All of the participants thought that the patient-facing app would fit into their lifestyle and help them manage their GDM. Using a 5-point Likert scale from very likely to very unlikely, 69.2% (9/13) and 61.5% (8/13) of participants indicated that they were likely or very likely to use the 'request a call' function or educational materials, respectively, while the eCourse had a more varied response.

Table 7.4 *Quantitative questionnaire responses evaluating the patient-facing app*

Question	Response	Number (percentage response, %)
Do you think the app presented would have helped you manage your gestational diabetes?	Yes	13 (100)
	No	0
Could you see the app fitting in with your lifestyle to help manage your gestational diabetes?	Yes	13 (100)
	No	0
On a scale of 1 to 5, 1 being very unlikely and 5 being very likely. How likely, if you needed to, would you have used the 'request a call' feature?	Very likely (5)	5 (38.5)
	4	4 (30.8)
	3	3 (23.1)
	2	1 (7.7)
	Very unlikely (1)	0
On a scale of 1 to 5, 1 being very unlikely and 5 being very likely. How likely would you have used the 'Know more' educational resources?	Very likely (5)	5 (38.5)
	4	3 (23.1)
	3	4 (30.8)
	2	1 (7.7)
	Very unlikely (1)	0
On a scale of 1 to 5, 1 being very unlikely and 5 being very likely. How likely would you have used the eCourse/quiz?	Very likely (5)	3 (25)
	4	4 (33.3)
	3	3 (25)
	2	2 (16.7)
	Very unlikely (1)	0

There were five free-text questions (supplementary material 3, **Appendix I**), in which four themes with eight subthemes emerged (**Table 7.5**). Participants expressed how helpful meals and exercise ideas would be for managing their GDM, as these were areas that they struggled with. There was appreciation for the comprehensive and accessible information, this was also shown in the Likert responses. Easy access to SMBG readings with visualizations was seen as beneficial, especially with the ability to manually add or edit readings which could compensate for technical issues. Participants particularly valued the ability to contact their healthcare team through 'request a call' and that the information was tailored to GDM, all of which would help them self-manage their GDM. When asked what they liked least or was missing from the patient-facing app, participants responded with nothing and that it had everything they wanted.

Table 7.5 *Themes, subthemes, and illustrated quotes from the free-text responses in the questionnaire on the patient-facing app.*

Quotes have been identified by participant number, along with the question Q1: Do you think the app presented would have helped you manage your gestational diabetes? [Comments], Q2: How likely, if you needed to, would you have used the 'request a call' feature [Comments]?, Q3: What did you like most about the app?

Theme	Sub-theme	Illustrative quote
Managing lifestyle changes	Diet	"Suggestion for meal prep. As I struggle with it every day." P2 Q3
	Exercise	"Meal ideas would have been very helpful and exercise too!" P1 Q1
Facilitating self-management	Education resources	"It was full of information that is useful." P6 Q3 "Helpful information is great." P13 Q3
	Blood glucose tracking	"Looks easy to understand and find your readings." P12 Q3
	General gestational diabetes self-management	"My glucose levels have been up and down and a request for a call button would help with the anxiety of calling someone." P3 Q2
Contact and communication with the clinical team	-	"Ease of getting in touch with hospital team/receiving messages from them." P9 Q3
General app usability	User experience and interface	"Good layout and lots of information." P6 Q8
	Useful and helpfulness	"It's easy to navigate and has everything I need." P2 Q3
	Tailored to gestational diabetes	"Relevant to my current situation, I feel [current blood glucose monitoring app] isn't aimed at me." P8 Q3

7.4.3.4 Final design

The mHealth tool, "MyGDM", was designed from results of the interviews and refined through feedback sessions and questionnaire. A description of the key features of the final design alongside the rationale can be seen in **Table 7.6**.

The finalized interface design consists of a clinical dashboard linked with a patient-facing app. It is intended that SMBG readings will be uploaded from the women's blood glucose meter to their app via Bluetooth, and once connected to the internet, readings are then shared with the clinic.

At the clinical set up, clinical glycemic targets are set, and features can be removed. Patients would be sent a clinic code to download and register on the app before their appointment.

The clinical dashboard's homepage has all the women with GDM currently in the clinic, with prompts regarding if any women have requested a call and a simple traffic-light system showing if a woman has uploaded their SMBG readings and how in-target they are. The dashboard can also be filtered and sorted by location, medication, call requests and warnings. The patient page has visualization, including a scatter graph of SMBG and a pie chart of in-target SMBG, to simplify the identification of trends in women's SMBG. And a logbook to aid with reviewing and making clinical decisions. Basic patient information including medication is displayed in a sidebar.

Additionally, there is a function at booking, to predict if a woman would potentially need pharmacological intervention. The clinic could use this information to streamline care, by identifying high-risk patients who would need more clinical time compared to low-risk patients who may be able to manage through diet therapy and remote monitoring.

The patient-facing app for women with GDM contains visualization and notices regarding their SMBG, reminders of targets and medication, the ability to add notes, log food/meals, or manually add or edit SMBG.

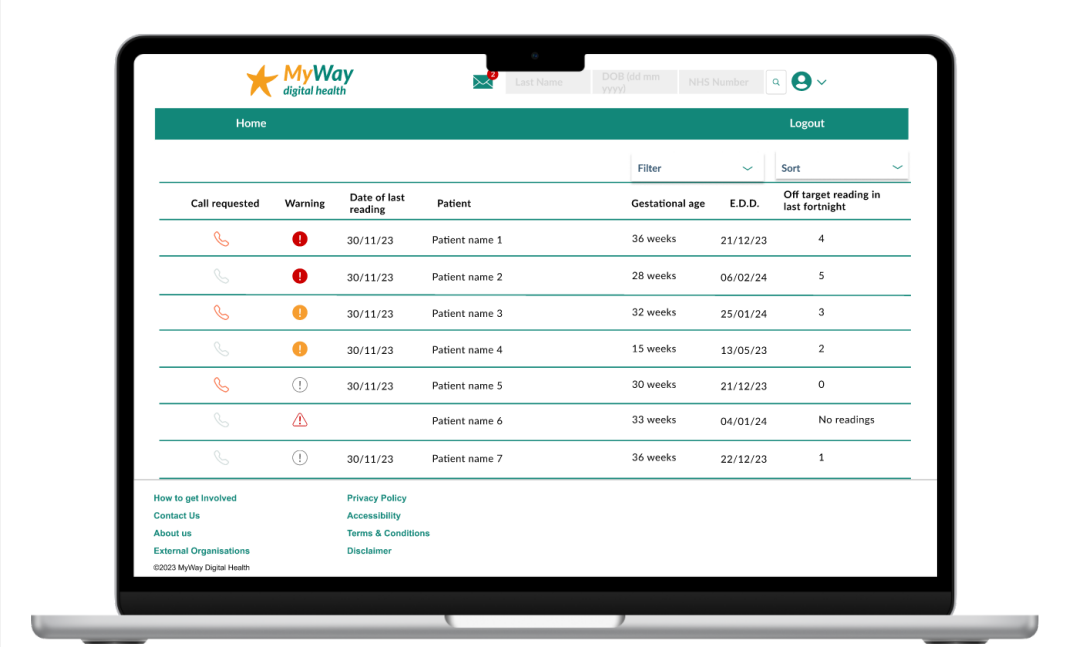
Educational articles include videos and text on GDM topics, an eCourse, and meal ideas. Language options, at app set-up, were included because of the diverse GDM population. Finally, there is an ability to request a call and view any messages from the clinical team.

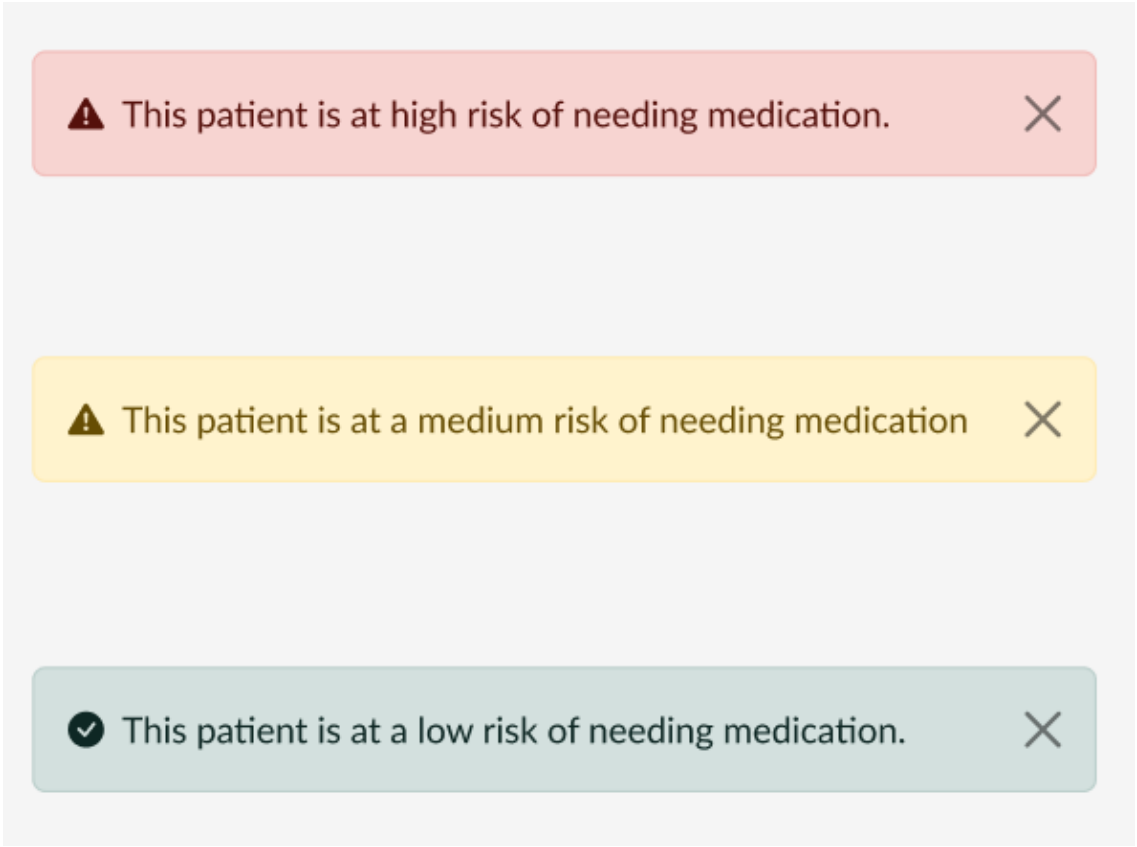
A GDM community forum which was initially included from the results of (Safiee et al., 2023), was removed from the final design because concerns over patient safety with the potential to spread misinformation was raised during the feedback sessions.

Furthermore, during the interviews and feedback sessions it was considered how the app could be used to improve attendance of postpartum HbA1c, so education articles and a postpartum reminder email/notification was added.

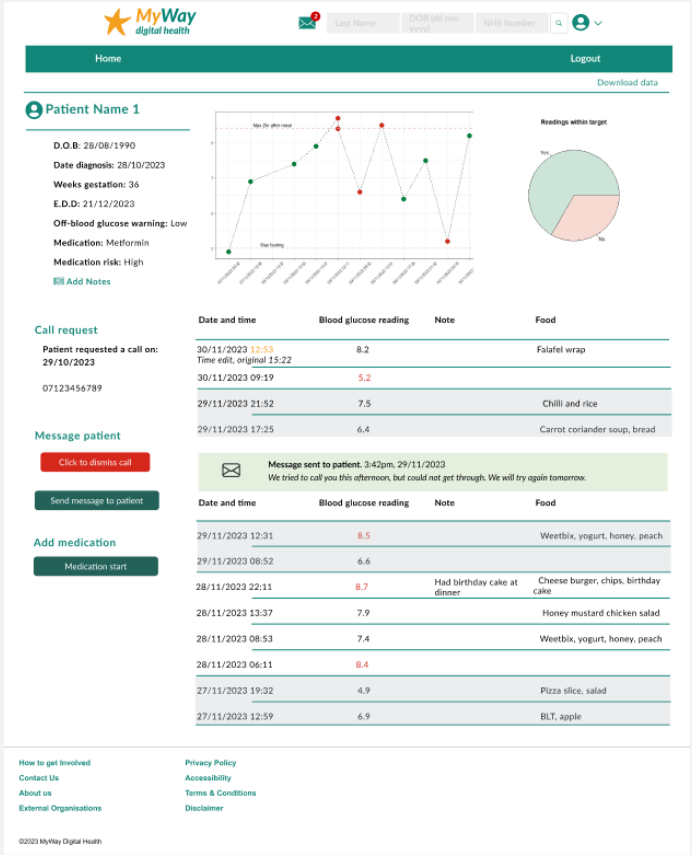
Table 7.6 *mHealth tool key features with rationale and sources*

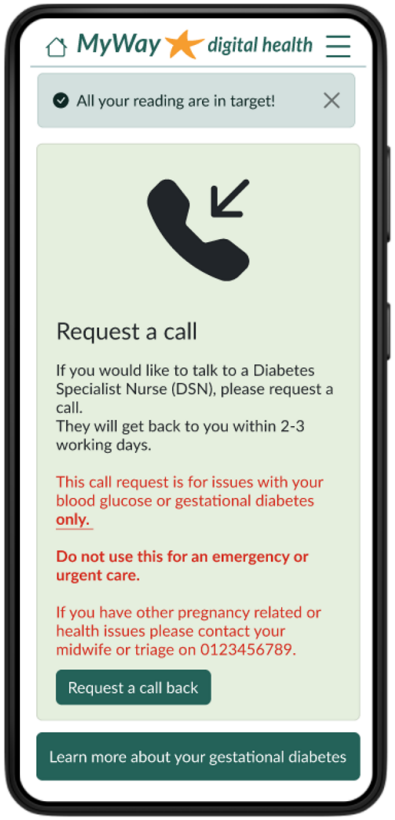
(HCPs: Health care professionals, GDM: gestational diabetes mellitus, SMBG, self-monitoring blood glucose)

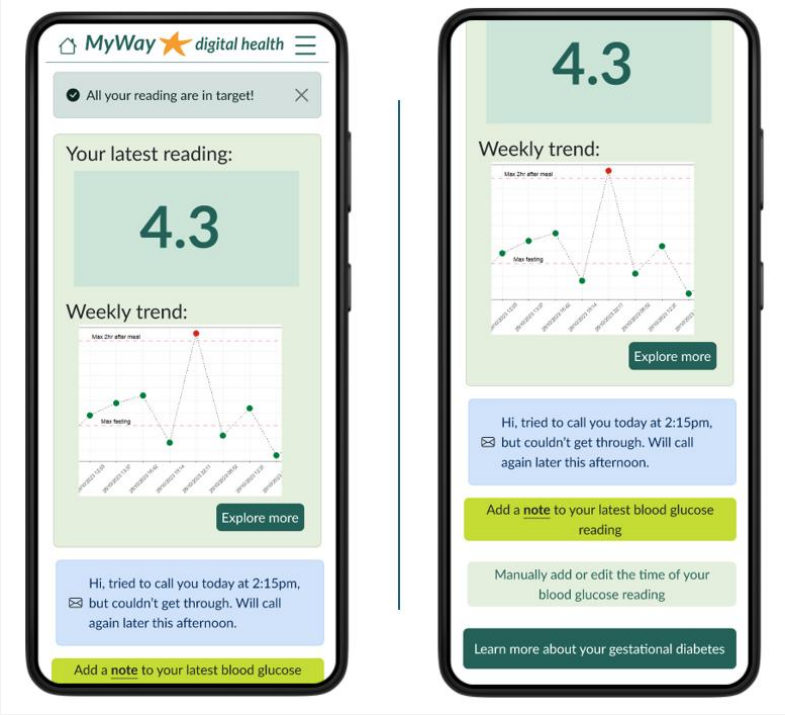
mHealth Tool Key Features	Rational and Source
Clinical dashboard homepage 	A major issue that arose during the clinical observation and interviews is that HCPs must search each patient's name to review their SMBG readings. There was a clear need for a dashboard that held all patients currently in the clinic centrally. Furthermore, highlighting off-target SMBG was a theme from the interviews, and adapting a simple traffic light system to achieve this was implemented on the clinical dashboard homepage. The addition of filtering and sorting came from feedback sessions 2 and 3.

Medication risk prediction 	During the interviews, HCPs spoke of an increase in the GDM population, and that care is time-intensive for both HCPs and patients. Therefore, streamlining care by predicting high-risk patients would allow for rapid treatment escalation and could reduce the burden on the care system. In addition, during Feedback session 2 it was said that it could help identify which patients would need closer follow-up.

Viewing a patient's page	The patient's page is a solution to include all the patient's inputs. The addition of visualization and logbook formatting came from feedback sessions 1 and 2. The ability for a woman with GDM to manually add or edit their SMBG was included because of (Safiee et al., 2023). It was then in feedback session 3 that highlighting those readings that had been edited was added.



Communication **(Request-a-call, and messages from HCPs)** 	To increase app engagement (Dalton et al., 2018), and to make contacting HCPs for support more convenient (Duan et al., 2022), 'request a call' function, and messages from HCPs were included in the digital tool. From the questionnaires, the ability to request a call was well-received.

App homepage 	The key information needed to self-manage GDM is displayed on the homepage. This aims to be as efficient as possible and to fit into women with GDM's lifestyle which was a theme that emerged from the interviews. Additionally, women with GDM wished to have all their data summaries onto one page (Safiee et al.). Visualization and messages of SMBG reading using a simple traffic-light system allows for real-time feedback. The message about the latest readings has been adapted from (Blair et al., 2022).

Education: Know more' and eCourse 	During the interviews, there was a clear voice for more information and education on GDM from both women with GDM and HCPs, which was consolidated in the questionnaires. In addition, it has been found that education can improve quality of life and self-efficacy (Akalpler and Bağriaçik, 2023). Structured smartphone-based education for GDM is a low-cost way to improve self-efficacy (Ghaderi et al., 2022). But reliable information on GDM is hard to acquire (Safiee et al., 2023), which was also emphasized in our interviews with HCPs. Therefore, to include reliable information, we included educational articles.

Food diary and log 	A food log and diary to help facilitate women to self-manage GDM. Tracking meals could help women identify which foods spike their blood glucose that should then be avoided. In addition, it has been expressed that participant interviewed for a theoretical pregnancy app wanted to monitor diet (Lazarevic et al., 2023).

Meal ideas

Snack Item	Kilo-calories (kcal)	Carbohydrate (grams)
1 small banana	60	15
1 Medium apple	62	15
2 Apricots	34	8
2 Satsumas/ kiwis/ plums	44	10
1 Medium peach	46	11
10 Strawberries	38	8
2 Prunes	48	10
1 Rice cake	30	7

During the interviews with women with GDM, meal, and recipe ideas were regularly discussed as desirable. During the questionnaires, the inclusion of meal ideas was reported as beneficial. Culturally inclusive meal options were included.

7.4.4 Discussion

We designed an interactive interface of a user-centered mHealth tool consisting of a clinical dashboard linked to a patient-facing app for GDM care and management, through key stakeholder interviews and a literature review. This was evaluated positively through feedback sessions and questionnaires.

Key features of the mHealth tool included a centralized clinical dashboard; promoting and highlighting off-target SMBG; credible education; meal ideas; data visualization; facilitating communication; and predictive risk stratification.

A time-consuming task that was highlighted by HCPs was the current method of inputting each patient to review SMBG. To combat this, a centralized clinical dashboard using a traffic light system to indicate priority patients was developed. This allows for remote monitoring which is time-efficient for both clinics (Rigla et al., 2018; Poulter et al., 2022) and women with GDM (Harrison et al., 2017; Alqudah et al., 2019) and reduces the need for in-person appointments (Zork, 2022). From the questionnaires, the ability to request a call was well received by women with GDM, with 69.4% likely to use it. In the feedback session HCPs could see that this would help women manage themselves better but were concerned about who would monitor the requests. Remote monitoring in combination with 'request a call' could reduce the number of appointments and therefore fit better into a GDM woman's lifestyle by reducing the burden of traveling and attending appointments, which can often be a barrier (Zork, 2022; Safiee et al., 2023; Harrison et al., 2017; Poulter et al., 2022; Duan et al., 2022).

Women with GDM in the interviews requested education and meal ideas, and so these were included and then well received with 61.5% of women completing the questionnaires indicating that they would be likely to use them. Education can improve the quality of life and self-efficacy for women with GDM (Akalpler and Bağriaçik, 2023), especially as women with GDM use the internet as a supplementary resource for information (Xu et al., 2024), suggesting that there is a need for credible information sources. Hence,

evidence-based GDM-specific education resources was developed and included. Structured smartphone-based education for GDM has been shown to improve self-efficacy (Ghaderi et al., 2022) and so an eCourse was included. In addition, language options and culturally inclusive meal ideas were included to reduce cultural and language barriers that have been reported (Garnweidner-Holme et al., 2018; Birati et al., 2022).

The clinical dashboard could identify patients' risk of need for medication which could streamline care through risk stratification. Models have been developed to predict pharmacological intervention (Alvarez-Silvares et al., 2022; Ali et al., 2018; Barnes et al., 2016; Ducarme et al., 2019; Eleftheriades et al., 2021; Ford et al., 2022; Liao et al., 2022; Pertot et al., 2011; Spaulonci et al., 2013; Velardo et al., 2021; Watanabe et al., 2016; Weschenfelder et al., 2021; Zhang et al., 2016), but to our knowledge, none have been included within a digital tool or implemented clinically. The use of predictive modelling in GDM could be a useful adjunct as a clinical decision support tool for HCPs.

A review of 11 mHealth apps for GDM that use artificial intelligence (Daley et al., 2021) found only one app (GDmHealth) that has been clinically implemented in England (Lu et al., 2023). GDmHealth (Mackillop et al., 2014) is GDM specific and helps with the management of SMBG, enables two-way communication between HCPs and patients but does not provide any education. Whilst, a scoping review including 30 studies on the use of mHealth for GDM, found that half (n=14) were for SMBG, and other common features included, real-time feedback, communication with professionals, and education (Edwards et al., 2022), all of which we have included in mHealth tool, in addition to developing with our end-users.

Diabetes technology is diversifying (Rhee and Rhee, 2024; Murphy et al., 2008) and as such, the mHealth tool will be adaptable for the inclusion of new technologies for GDM. Future work includes development of predictive models for pharmacological intervention, in conjunction with back-end building and integration into MyWay Clinical and obtaining necessary regulatory standards for medical devices.

7.4.4.1 Strengths and Limitations

One strength is the user-centered design and evaluation used. This prioritizes both women with GDM and HCPs needs and wishes and produces a more usable and effective digital tool. However, there was a lack of ethnic diversity and with self-selecting sampling increases the bias toward those who were more engaged with their care and, hence, may not be representative of the whole GDM population.

7.4.5 Conclusion

We have designed a user-centered mHealth tool consisting of a clinical dashboard and a patient app for GDM care and management. Evaluation of the interactive design by end-users was positive and showed that it met their needs. Future work includes the development of predictive modeling and back-end, and then piloting to ensure that it can function and improve outcomes within a clinical setting.

7.5 EXPLANATION OF QUALITATIVE METHODOLOGY

The user needs assessment study had a focused ethnographic approach to explore the experience and challenges of GDM care and management with HCPs and women with GDM, namely, to understand current care and the views, options, and acceptance of digital technologies within GDM care and management. Focusing on specific aspects of GDM care through observations and semi-structured interviews allowed for in-depth qualitative insights, which informed the design of MyGDM.

Prior to conduction interview and to familiarise myself with GDM clinics, I observed four antenatal diabetes clinics in three hospitals in Edinburgh (two at The Royal Hospital and one at The Western General Hospital) and Glasgow (The Princess Royal) were observed between September - December 2022. During these observations I would sit within clinical appointments and watch how HCPs and patients interacted with each other and what the flow of the clinic was. I would write brief notes, mainly on if the patient has GDM or other type of diabetes, so that I could remember key event when I wrote up. After the

observation, I would then write up my notes, these include what patients had been seen, how many were there, what were their emotional state, how did HCP interact with them, how many HCPs did a patient have to see in one clinic, how many people were in the room at one time, did people come and go, who attended the appointment with the women, what was they layout of the room, what technology and software was used during appointment, how did HCP and patients use and interact with the technology and software, did it cause a physical barrier or open up conversation, what direction were HCPs facing when looking at notes or reviewing blood glucose levels e.g where they turn away of patient, were they looking over the top of their screen. After observing the four antenatal clinics and reviewed my observation notes and reflected on them. I then wrote up my overall reflections on the observations, focusing on the use of technology and clinical workflow within GDM antennal clinics.

Thematic analysis of transcripts from the semi-structured interviews, the feedback sessions and the free-text answers of the questionnaires, was carried out using the 6-step framework described by Braun and Clarke (Braun and Clarke, 2006). This included (1) Familiarisation with the data, where I read through the transcripts; (2) coding the transcripts using NVivo, with 10% double-coded by Dr Jane Dickson; (3) searching for themes by reviewing the codes and summarising the transcripts; (4) reviewing themes with Dr Jane Dickson and (5) defining and (6) naming themes. Codes and themes were refined through discussions throughout with Dr Jane Dickson.

GDM care is multidisciplinary; to gain concise information about GDM care, key HCPs were targeted during the interviews and feedback sessions. These included consultants, dietitians and diabetes specialist nurses or midwives who have vital roles through regular contact, education, prescribing and monitoring within GDM care.

During this piece of research, as I became more familiar with the settings, I was able to better understand how digital tools fit into the care and what aspects of the currently used digital tools were missing. In conjunction to the

literature, my own observations and direct information from HCPs and women with GDM, MyGDM was design to meet the needs of the end-users while being able to capture the key feature and integrate smoothly into clinical workflow.

7.5.1 Reflective statement

I am a white British female who completed this qualitative research for my PhD. I have a background in mathematics and no prior qualitative or interview experience, therefore, I held an 'outsider's' view when conducting this research. I had no personal relationship with any of the participants prior to the study. I received formal training during the study through the University of Edinburgh's Clinical Research Facility (Introduction to Qualitative Methods). I was also guided by Dr Jane Dickson, a medical anthropologist at the University of Dundee, who has many years of research experience in understanding people's experiences to improve healthcare, particularly within diabetes, and works with MyWay Digital Health.

7.6 CHAPTER SUMMARY

This chapter presented the user-centred design and evaluation of an interactive prototype, 'MyGDM', which consists of a clinical dashboard and a patient-facing app. MyGDM was developed from 11 semi-structured interviews with end-users (six HCPs and five women with GDM) and the literature. It received a positive evaluation through three feedback sessions with 31 participants (17 HCPs and 14 researchers) and questionnaires with 12 women with GDM.

Key features of the clinical dashboard are that it centralises all the women within the GDM clinic, highlights women with consistently off-target readings and those who have requested a call and can also provide pharmacological therapy prediction that could be used to risk-stratify care. The patient-facing app's key feature improves the self-management of GDM through colour-coded blood glucose reading, graphs and visualisations of SMBG, notifications and prompts about SMBG, bi-directional communication with the GDM clinic, reliable multi-media education resources, self-paced

eCourse, meal logging, culturally diverse recipes and meal ideas and alternative language options.

This has provided the initial steps in the development of a Scottish GDM-specific digital tool that could integrate into Scotland's diabetes database, SCI-Diabetes, described in more detail in the next Chapter, Chapter 8.

CHAPTER 8

Discussion

8.1 Introduction

This thesis presents an interdisciplinary, mixed-methods approach to data-driven personalised care and management for GDM, through a digital tool, "MyGDM", and risk-stratified care through identifying which patient might need treatment with insulin by machine learning insulin prediction models.

In this chapter, I will firstly highlight the main findings from each chapter. I will then discuss the implications of the thesis's findings. Following this, I will summarise strengths and limitations, while reflecting on the methods I have used. I will finish with the future directions for the work done in this PhD that can be taken forwards.

8.2 Main findings

The thesis starts with a comprehensive scoping review of the use of machine learning to predict pharmacological therapy in GDM, Chapter 3. The review included 17 studies describing 44 models, all of which were binary classifiers, and often used logistic regression. Overall, the models had a median AUROC of 0.75 (range 0.61-0.93) and used common clinical variables. Some were internally validated, but there was a lack of external validation. The results of the scoping review set the basis for variable retrieval for the study datasets, which are then cleaned in the following chapter and the starting points for the modelling.

In the following chapter, Chapter 4, three study datasets consisting of 30,666 pregnancy episodes with a delivery date between 1st January 2021 and 31st

December 2023 in Greater Glasgow and Clyde were provided for research. I cleaned and selected a cohort of 10,694 singleton pregnancy episodes that received care after the pandemic, with a booking date after 1st April 2022. Additionally, I identified 1109 (10.4%) pregnancies complicated by GDM by applying SIGN 2010 guidelines. For modelling purposes, I removed missing risk factors for GDM and any cases with prior metformin use, resulting in a modelling dataset of 996 singleton pregnancy episodes complicated by GDM.

The cohort of 10,694 singleton pregnancy episodes was analysed in Chapter 5. Compared to the PHS data of maternities in Greater Glasgow and Clyde between 1st April 2022 and 31st March 2023, the cohort was found to have good representative maternal characteristics. Furthermore, from descriptive analysis, it was identified that those with GDM had significantly higher age and BMI in comparison to women without GDM. While those requiring insulin experiencing higher rates of Caesarean birth and LGA babies, and, while not significant, had higher rates of neonatal unit admissions.

The final quantitative chapter, Chapter 6, presents the development of two machine learning models, logistic regression and CART, to predict the need for insulin at the time of GDM diagnosis using the modelling dataset (n=996) described in Chapter 4. The models were built and developed on 70% (n=667) of the data and evaluated using 10-fold cross-validation. They were then validated on 30% (n=299) of unseen test data. Logistic regression had a training AUROC of 0.77 [0.07] and a validation of 0.71, whereas CART had an AUROC of 0.70 [0.10] and 0.70, respectively. Both models had good recall on the insulin class during validation (logistic regression: 0.83, CART: 0.78), showing that they are sensitive, but were not specific, with poor precision (logistic regression: 0.13, CART: 0.15).

The final results chapter, Chapter 7, details the user-centred design and potential use of MyGDM, a clinical dashboard and patient-facing app for GDM care and management. It was designed from ethnographic observations, 11 semi-structured interviews with both HCPs (n=6) and women with GDM (n=5), followed by iterative design changes from three feedback groups with 31

participants (17 HCPs, 14 researchers) and 13 questionnaires with women with GDM. MyGDM highlights off-target SMBG reading, has bi-directional communication and centralises women with GDM within the clinic. In addition, it provides structured education, language options and culturally appropriate meal ideas. Overall, potential end-users were enthusiastic about the design and prospect of its use in the clinic.

8.3 IMPLICATIONS OF THE RESULTS

This thesis proposes an alternative approach to GDM care that integrates predictive modelling with a user-centred digital tool, which can be embedded in SCI-Diabetes through MyWay Digital Health. The implications of these findings span digital health design, predictive modelling, service delivery and health policy.

Digital health interventions have many benefits; they can aid with better communication, increase flexibility of care, improve understanding through easy data access and education, aid in decision making and remove geographical barriers (Odendaal et al., 2020). mHealth within GDM has high user satisfaction among both women and HCPs (Mackillop et al., 2018; Smyth et al., 2025; Peleg et al., 2017a; Butten et al., 2024; Skar et al., 2018; Caballero-Ruiz et al., 2017). It is often reported to improve self-management, recording compliance and reduce in-person appointments, while not compromising care provided or maternal and neonatal outcomes (Caballero-Ruiz et al., 2017; Kytö et al., 2024; Mackillop et al., 2018; Smyth et al., 2025; Peleg et al., 2017a; Butten et al., 2024; Borgen et al., 2019).

MyGDM was developed with a user-centred design, which places it within the lived experience of HCPs and women with GDM. In addition, it builds on current literature, and it provides structured education and language accessibility. It addresses gaps within the literature, especially around health literacy, fragmented communication and information, and postpartum follow-up.

Despite the benefits of increased access to information and virtual visits that allow remote monitoring, these also come with barriers, (Cao et al., 2025). Increasing access to information does not necessarily improve care that is provided, particularly if the data recorded by mHealth is not integrated into the health care systems, or there are concerns around the safety and accuracy of the information. In addition, the uptake of such tools into clinics often requires changes to workflow, which may be met with resistance, particularly if it will cause an increase in workload, for example, from the duplication of tasks. Emphasising again that mHealth tools need to be integrated fully into existing systems.

GDM requires additional appointments where SMBG values are reviewed and adjustments to treatment are made (Scottish Intercollegiate Guidelines Network, 2024), placing a strain on women with GDM's time, which was also highlighted in the user needs assessment (Kirkwood et al., 2025a). Providing remote monitoring of SMBG with virtual appointments can relieve some of this burden. Consistently, mHealth apps for GDM have reduced the number of in-person appointments (Mackillop et al., 2018; Peleg et al., 2017a; Borgen et al., 2017; Caballero-Ruiz et al., 2017), and it has been reported to increase the number of interactions per patient (Hyams et al., 2024). However, it appears that the apps in the literature mainly focus on the management of GDM and lack personalisation (Daley et al., 2021). Therefore I proposed care to be personalised such that those with a low risk of requiring insulin therapy, e.g. they will be able to manage their GDM through diet or metformin, could have their blood glucose remotely monitored and have monthly check-ins during their joint antenatal-diabetic clinic, whereas those with a high risk of requiring insulin could also be remotely monitored, but have virtual appointments and monthly check-ins (**Figure 1.3**).

The insulin risk-prediction model (Chapter 6) indicated that there was feasibility in prediction insulin-need at the point of GDM diagnosis using routinely collected data. The models achieved moderate discrimination; they were sensitive but with low precision. From a clinical perspective, having high

sensitivity would be more acceptable in a triage context, where failing to successfully identify which women may need insulin has a greater risk that over triaging and enhancing monitoring. On the other hand, over-escalation into the enhanced monitoring stream could increase HCPs workloads and potentially women with GDM's anxiety too. Therefore, it is important to consider the model as a tool to support HCPs clinical judgment and not as a final decision on a women's care pathway.

The principal innovation of this thesis is together the insulin-risk prediction model within a mHealth tool for GDM care and management. The insulin-risk prediction model identifies women for different intensities of monitoring and contact. MyGDM then provide a platform for remote monitoring. This could then improve clinical efficiency, reduce unnecessary appointments, escalate care quicker and improve postpartum monitoring and appointment attendance. However, such a change in clinical workflow should be appropriately integrated. Risk prediction models raise concerns, reinforcing bias and over-reliance can inadvertently obscure accountability of outcomes.

Importantly, as MyGDM would sit within MyWay Digital Health and link directly into SCI-Diabetes, this would reduce implementation barriers, improve sustainability and allow scalability across Scotland. Integration within existing structures moves the concept beyond a standalone digital tool into a potential system-level intervention.

8.4 STRENGTHS AND LIMITATIONS

A scoping review was selected to synthesise evidence from the literature because I wanted to better understand the research that had already been established within this area, alongside identifying any research gaps, rather than providing a collated and succinct body of work (Munn et al., 2018). The results of the scoping review then guided the variable selection for data analysis (Chapters 4, 5) and the building of the models (Chapter 6).

To ensure that the results were transparent and reproducible, the analysis and results were reported following the PRISMA-ScR (Tricco et al., 2018). The quality of the predictive models was assessed using PROBAST (Wolff et al., 2019), which assesses the risk of bias and applicability of models within a clinical setting. Quality and risk of bias assessment are not traditionally completed for scoping reviews; however, I chose to, as model quality is important. Using PROBAST, I was able to gauge a more thorough understanding of the published models' ability to predict and their applicability in a clinical setting.

Due to the nature of a scoping review, the results found have breadth but lack depth, particularly as the results of the models were pooled together without grouping by methodological approach, as in a systematic review. Hence, the results give an overview of the field rather than provide specific evidence for a distinct question. Without a distinct question, the subjectivity of the results increases, contributing to bias in the selection of the studies. This was evident when screening literature that used logistic regression. Logistic regression can be a statistical (identifying associations) or machine learning method (producing predictions); therefore, a judgment was made, if the article reported predominantly statistical metrics, such as odds ratio or relative risk, then it was excluded.

Nevertheless, due to the heterogeneity and multidisciplinary nature of the literature within this area, a comprehensive and extensive search was completed through a wide search strategy and searching both medical and machine learning databases. This extended the volume of captured literature, making the review exhaustive.

The data used in this PhD were secondary use of clinical care data; therefore, it was vital to clean the data to remove any errors and noise (Sandfeld, 2024). I assumed that there was no bias involved during data input by HCPs. However, there are biases in the use of EHR data, which may arise from certain patient populations' access to care (Chen et al., 2021b). The bias in this data cleaning is likely to be negligible due to healthcare being free and accessible to all in

Glasgow. That being said, there will be groups that are less engaged or have distrust in healthcare services; consequently, this may lead to them having high volumes of missing data. To be inclusive, I tried to keep as much of the missing data in as possible. There may be a common theme within the groups with missing data, of which I am unaware and have not adjusted for. However, comparing with national statistics from PHS showed that the cohort had representative maternal characteristics of the Greater Glasgow and Clyde maternal population.

I was surprised by the inconsistency in GDM recording. Due to this, I had to estimate who had GDM by applying SIGN guidelines and through logic. Furthermore, the final treatment for GDM was not recorded, and I had to assume that if there was a record of metformin or insulin being dispensed during the pregnancy, it had been taken to manage GDM. These assumptions and estimations inherently will introduce bias, which will have influenced the descriptive and statistical analysis and modelling.

As a woman can be pregnant multiple times, I chose to remove any second pregnancies within the cohort dates to ensure that the data was independent for statistical analysis. Accordingly, subsequent pregnancies are also associated with an increased risk of GDM (Buchanan et al., 2012) and therefore removing the second pregnancy episodes reduced the noise and complexity of the data for the machine learning modelling.

By selecting pregnancy episodes that had care after the pandemic, I reduced the data to just over a third of what was initially available. However, it is important to prioritise data quality over quantity. From discussions with my clinical supervisors and exploratory analysis, care was systematically different during the pandemic. For this reason, it was appropriate to select a stable period of care after the pandemic. I selected booking dates after the second quarter of 2022, as this was when most, if not all, restrictions had been lifted.

The standard descriptive and statistical analysis tests were used, making the results reproducible. In addition to the p-value, the standard deviation was also calculated, giving the spread on the data. To select the appropriate test, the assumptions of the tests needed to be established. As the sample size of those with GDM was much smaller than those without, a visual inspection of Q-Q plots was done to understand the normality of the data, which is limited as it is subjective. There was a large difference in the sample size of the GDM and non-GDM complicated pregnancies, which affected the p-values' ability to detect actual differences within the data; hence, there were some instances of statistically significant but little difference after inspecting the median, mean and interquartile range of the boxplots.

There are many different options to choose from and parts to consider when modelling, and I tried to ensure that at each step, the most appropriate option was chosen with consideration of the problem, and for the data I had available. Given the volume of data, it was not possible to develop a multi-class model (identified as a research gap in Chapter 3, scoping review) for predicting treatment as lifestyle, metformin, or insulin. Nevertheless, I wanted to make a model that was appropriate for clinical use; therefore, I chose explainable algorithms, logistic regression and CART (Neha Margret et al., 2024). I then applied a standard modelling approach, involving a holdout train-test split and balancing the data with SMOTE.

Feature selection was customised for each algorithm. Although univariate feature selection is robust, it overlooks variable interactions and multicollinearity, meaning important features might have been excluded too early.

For the class imbalance, I chose to generate synthetic examples using SMOTE, rather than just resampling the data using other methods, which could discard relevant information (Fernández et al., 2018). Since SMOTE generates synthetic samples, it may lead to overfitting or fail to fully capture the data's complexity. To prevent data leakage, SMOTE was applied during training within each fold of the 10-fold cross-validation.

The most important results when modelling are the validation results, which show the model's ability to generalise on unseen data and give an understanding of how it could perform if implemented. Appropriate metrics were selected to enable comparison of the models with each other and with other models in the literature. To validate the model, I chose holdout validation, which is a very common method to evaluate a model (Maleki et al., 2020). Due to my small sample size and class imbalance, it was appropriate to do a 70:30 split (Maleki et al., 2020; Gupta and Sedamkar, 2020). Furthermore, to provide an unbiased evaluation during training, I used 10-fold cross-validation, which was stratified to account for the class imbalance (Maleki et al., 2020).

The validation test set was small (n=299, insulin: 23, non-insulin: 279), so the results are in the preliminary stages of developing a usable model. External validation of the models was intended on a cohort in Edinburgh, but despite obtaining ethical approval to access the data in 2023 (**Appendix K**), the data were not available in time to conduct this analysis. Additionally, I had hoped to temporarily validate the models on data from Greater Glasgow and Clyde that had later booking data, but this, too, did not become available in time for this analysis.

Involving users in early design stages increases the likelihood of developing a usable and wanted product (Still and Crane, 2017), hence, I chose a user-centred design where the end users are involved in the development (Chamberlain et al., 2022). Before designing, I needed to understand the views and needs of the end user, which would allow for the finalised digital tool to improve the user experience and usage (Duan et al., 2022). The qualitative design and method used in this chapter allowed for an in-depth understanding of what end-users would want from a digital tool in GDM care and management. The evaluation results were positive and have provided an understanding of opinions on MyGDM from a sample of HCPs and women with GDM. This has set the foundation for its further development.

GDM care is multidisciplinary; to gain concise information about GDM care, key HCPs were targeted during the interviews and feedback sessions. These included consultants, dietitians and diabetes specialist nurses or midwives who have vital roles through regular contact, education, prescribing and monitoring within GDM care.

As I have no medical background or experience with GDM, I needed to understand the environment in which I was designing (Still and Crane, 2017). Hence, I observed four antenatal diabetes clinics in three hospitals in Edinburgh and Glasgow. After the site visits, I wrote notes on the clinics, such as how the clinics were run, the emotional reactions of women with GDM, how HCPs interacted with women with GDM and where potential issues lay. After all the site visits were completed and I reflected on my observations, I considered the differences between the sites, which elements could be improved, and where a digital tool or CDSS could fit.

To explore the end-users' needs, thematic analysis of semi-structured interviews was conducted. Semi-structured interviews allow for an in-depth exploration of the research topics and a greater understanding of our users' wants (Gubrium et al., 2012). Given the time constraints of a PhD, this provided a rich source of information from a small sample.

Recruitment did have challenges; it was particularly difficult to recruit women with GDM during the interview stage. Reducing the interviews from approximately 30 minutes to 10 minutes and focusing on current care and features of a potential digital tool allowed for quick interviews between or after appointments. The small sample size is typical for qualitative studies (Braun and Clarke, 2006). Additionally, we reached data saturation where no new themes were emerging. For HCPs we were able to recruit representative roles within GDM care, however, this could be improved by having a minimum number for each role and equal numbers from our study locations. Recruitment of women with GDM was particularly difficult as I did not get a lot of interest, and there was a degree of trial-and-error in recruitment. Additionally, recruitment was just after the pandemic, as services were

returning to ‘normal’. The difficulties in recruiting women with GDM could be due to numerous reasons, such as a lack of time or incentive to take part. Furthermore, though the women with GDM were somewhat ethnically diverse, most had other similar characteristics; we were also unable to recruit any women with GDM from Glasgow, which could have further consolidated our findings.

User-centred design is an inherently iterative process (Still and Crane, 2017). Therefore, to gain insight into HCPs' opinions on the MyGDM prototype group a feedback session were run. The group setting was chosen so that ideas from different perspectives could be explored and argued within the group (Terry and Hayfield, 2021; Still and Crane, 2017). As the participants in the groups already knew each other before the session, this increased the ease and flow of conversation.

To evaluate the MyGDM prototype with women with GDM, I chose an evaluation method that could be completed quickly between or after appointments. In addition, I wanted to have consistency in what women with GDM interacted with. Therefore, a short video demonstrating the prototype after the iterative changes from the feedback session, followed by a short questionnaire, was used to evaluate the MyGDM prototype with women with GDM. This gave quick and structured feedback from women with GDM.

8.5 FUTURE WORK

Outlined in this thesis is the foundation for a digital tool and alternative approach to GDM care and management, which needs further development.

The predictive models presented are in their infancy; they need to be expanded, further developed, and thoroughly tested before they can be considered for testing or implementing clinically. With the current data, gradient boosting models could be developed and compared with the logistic regression and the CART model. Or an alternative data splitting approach, such as bootstrapping, could be applied. The models need to be externally

validated. There should soon be available to the research team data of pregnancy episodes that have delivery data after 31st December 2023 from Greater Glasgow and Clyde; therefore, a temporal validation could be conducted. Furthermore, there should also be data available from clinics in Edinburgh, and the models could also be geographically validated. If the model had performed sufficiently, then ethical and regulatory applications would be needed before it could be trialled in a pilot study. A potential pilot study could run the model on a random selection of participants in addition to their standard care. This could safely test out the models and provide insight into how it integrates into clinics and where improvements are needed. Furthermore, a qualitative study on HCPs' attitudes to the introduction of the insulin risk-stratifying model could be conducted to identify key barriers for adaptation. The aim would be to have the risk-stratifying model fully tested alongside the MyGDM digital tool and integrated into SCI-Diabetes in a multicentre RCT.

The MyGDM app and clinical dashboard are being further developed and tested by Prof. R. Reynolds' grant from the Chief Scientific Office (CSO) in collaboration with Dr D. Wake at MyWay Digital Health. Further work includes qualitative studies with HCPs using the dashboard with dummy cases. Issues with the dashboard will be identified either through talk-aloud methods, in which they are using the dashboard and explaining what they are doing and/or through interviews afterwards. The patient-facing app is well established in the literature, and therefore, it would not be necessary to do a similar study. Once each component of the digital tool has been finalised, a feasibility study or RCT should be implemented. This would assess the use of MyGDM in a clinical setting using a mixed-methods approach, quantifying the maternal and neonatal outcomes and evaluating qualitatively through interviews and surveys with women with GDM and HCPs on their experience of using it. Additionally, to improve upon current literature, there should be a follow-up within the post-partum period to assess the post-partum OGTT results and their attendance, alongside the long-term outcome of both mother and child.

Improvement in the records of GDM is also needed. Another work package of the CSO grant is currently underway to record GDM in SCI-Diabetes, which will improve the overall records of GDM and allow for annual audits. Once this is established, MyWay Digital Health can be integrated into the platform, either through a module or by linking to an external site, which could be an option for MyGDM to be integrated into clinical care. Improvement of GDM records and with the with MyGDM linkage to SCI-diabetes would allow for lifelong care for those that subsequently develop T2DM after GDM.

8.6 CONCLUSION

This thesis has presented the groundwork for an alternative care pathway for GDM, through risk-stratifying care and digital tools, depicted in **Figure 1.3**. I have developed a machine learning model that predicts the risk of requiring insulin therapy at the time of GDM diagnosis. This could be used to streamline care and identify low-risk women with GDM who could be managed remotely. Furthermore, I have designed MyGDM, a clinical dashboard and patient-facing app that was well received with end-users, HCPs and women with GDM, and uniquely would be integrated into existing healthcare systems.

There is work underway to further develop MyGDM, which can revolutionise GDM care in Scotland.

REFERENCES

6.1: Drugs used in diabetes [Online]. British National Formulary. Available: https://openprescribing.net/bnf/0601/ [Accessed 19 Febuary 2025].

Covidence is a web-based collaboration software platform that streamlines the production of systematic and other literature reviews. [Online]. Available: https://www.covidence.org/ [Accessed 2023].

2021. *International Diabetes Federation, IDF Diabetes Atlas Tenth Edition 2021* [Online]. 2021. [Accessed 05 July 2022].

ABDALRAHMAN MOHAMMAD ALI, M. O., ABDELGADIR ELHABEEB, S. M., ABDALLA ELSHEIKH, N. E., ABDALLA MOHAMMED, F. S., MAHMOUD ALI, S. H., IBRAHIM ABDELHALIM, A. A. & ALTOM, D. S. 2025. Advancing Obstetric Care Through Artificial Intelligence-Enhanced Clinical Decision Support Systems: A Systematic Review. *Cureus,* 17(3)**,** e80514.

ABEJIRINDE, I.-O. O., DOUWES, R., BARDAJÍ, A., ABUGNABA-ABANGA, R., ZWEEKHORST, M., VAN ROOSMALEN, J. & DE BROUWERE, V. 2018. Pregnant women's experiences with an integrated diagnostic and decision support device for antenatal care in Ghana. *BMC Pregnancy and Childbirth,* 18.

AINSCOUGH, K. M., O'BRIEN, E. C., LINDSAY, K. L., KENNELLY, M. A., O'SULLIVAN, E. J., O'BRIEN, O. A., MCCARTHY, M., DE VITO, G. & MCAULIFFE, F. M. 2020. Nutrition, Behavior Change and Physical Activity Outcomes From the PEARS RCT—An mHealth-Supported, Lifestyle Intervention Among Pregnant Women With Overweight and Obesity. *Frontiers in Endocrinology,* 10.

AKALPLER, Ö. & BAĞRIAÇIK, E. 2023. Education programs for gestational diabetes mellitus: A systematic review. *Human Nutrition & Metabolism,* 33**,** 200195.

ALBERT, L., CAPEL, I., GARCÍA-SÁEZ, G., MARTÍN-REDONDO, P., HERNANDO, M. E. & RIGLA, M. 2020. Managing gestational diabetes mellitus using a smartphone application with artificial intelligence (SineDie) during the COVID-19 pandemic: Much more than just telemedicine. *Diabetes Research and Clinical Practice,* 169**,** 108396.

ALI, A., SHASTRY, S., NITHIYANANTHAN, R., ALI, A. & GANAPATHY, R. 2018. Gestational diabetes–Predictors of response to treatment and obstetric outcome. *European Journal of Obstetrics & Gynecology and Reproductive Biology,* 220**,** 57-60.

ALQUDAH, A., MCMULLAN, P., TODD, A., O'DOHERTY, C., MCVEY, A., MCCONNELL, M., O'DONOGHUE, J., GALLAGHER, J., WATSON, C. J. & MCCLEMENTS, L. 2019. Service evaluation of diabetes management during pregnancy in a regional maternity hospital: potential scope for increased self-management and remote patient monitoring through mHealth solutions. *BMC Health Services Research,* 19**,** 662.

ALVAREZ-SILVARES, E., BERMÚDEZ-GONZÁLEZ, M., VILOUTA-ROMERO, M., GARCÍA-LAVANDEIRA, S. & SEOANE-PILLADO, T. 2022. Prediction of insulin therapy in women with gestational diabetes: a systematic review and meta-analysis of observational studies. *Journal of Perinatal Medicine,* 50**,** 608-619.

AMEYAW, E. K. & BAATIEMA, L. 2025. Mobile Health (mHealth) Across Maternity Care Continuum. Springer Nature Singapore.

BARNES, R. A., WONG, T., ROSS, G. P., JALALUDIN, B. B., WONG, V. W., SMART, C. E., COLLINS, C. E., MACDONALD-WICKS, L. & FLACK, J. R. 2016. A novel validated model for the prediction of insulin therapy initiation and adverse

perinatal outcomes in women with gestational diabetes mellitus. *Diabetologia,* 59**,** 2331-2338.

BAUER, E. M. 2024. Chapter 2 - Using Digital Health Tools in Medical Practice. *In:* KLONOFF, D. C., KERR, D. & ESPINOZA, J. C. (eds.) *Diabetes Digital Health, Telehealth, and Artificial Intelligence*. Academic Press.

BELSTI, Y., MORAN, L., HANDISO, D. W., VERSACE, V., GOLDSTEIN, R., MOUSA, A., TEEDE, H. & ENTICOTT, J. 2023. Models Predicting Postpartum Glucose Intolerance Among Women with a History of Gestational Diabetes Mellitus: a Systematic Review. *Current Diabetes Reports,* 23**,** 231-243.

BENHAM, J. L., GINGRAS, V., MCLENNAN, N.-M., MOST, J., YAMAMOTO, J. M., AIKEN, C. E., OZANNE, S. E. & REYNOLDS, R. 2023a. Precision Gestational Diabetes Treatment: Systematic review and Meta-analyses. Cold Spring Harbor Laboratory.

BENHAM, J. L., GINGRAS, V., MCLENNAN, N.-M., MOST, J., YAMAMOTO, J. M., AIKEN, C. E., OZANNE, S. E., REYNOLDS, R. M., TOBIAS, D. K., MERINO, J., AHMAD, A., AIKEN, C., BODHINI, D., CLARK, A. L., COLCLOUGH, K., CORCOY, R., CROMER, S. J., DUAN, D., FELTON, J. L., FRANCIS, E. C., GILLARD, P., GAILLARD, R., HAIDER, E., HUGHES, A., IKLE, J. M., JACOBSEN, L. M., KAHKOSKA, A. R., KETTUNEN, J. L. T., KREIENKAMP, R. J., LIM, L.-L., MÄNNISTÖ, J. M. E., MASSEY, R., MCLENNAN, N.-M., MILLER, R. G., MORIERI, M. L., NAYLOR, R. N., OZKAN, B., PATEL, K. A., PILLA, S. J., PRYSTUPA, K., RAGHAVAN, S., ROONEY, M. R., SCHÖN, M., SEMNANI-AZAD, Z., SEVILLA-GONZALEZ, M., SVALASTOGA, P., TAKELE, W. W., TAM, C. H.-T., THUESEN, A. C. B., TOSUR, M., WALLACE, A. S., WANG, C. C., WONG, J. J., YOUNG, K., AMOUYAL, C., ANDERSEN, M. K., BONHAM, M. P., CHEN, M., CHENG, F., CHIKOWORE, T., CHIVERS, S. C., CLEMMENSEN, C., DABELEA, D., DAWED, A. Y., DEUTSCH, A. J., DICKENS, L. T., DIMEGLIO, L. A., DUDENHÖFFER-PFEIFER, M., EVANS-MOLINA, C., FERNÁNDEZ-BALSELLS, M. M., FITIPALDI, H., FITZPATRICK, S. L., GITELMAN, S. E., GOODARZI, M. O., GRIEGER, J. A., GUASCH-FERRÉ, M., HABIBI, N., HANSEN, T., HUANG, C., HARRIS-KAWANO, A., ISMAIL, H. M., HOAG, B., JOHNSON, R. K., JONES, A. G., KOIVULA, R. W., LEONG, A., LEUNG, G. K. W., LIBMAN, I. M., LIU, K., LONG, S. A., LOWE, W. L., MORTON, R. W., MOTALA, A. A., ONENGUT-GUMUSCU, S., PANKOW, J. S., PATHIRANA, M., PAZMINO, S., PEREZ, D., PETRIE, J. R., POWE, C. E., et al. 2023b. Precision gestational diabetes treatment: a systematic review and meta-analyses. *Communications Medicine,* 3.

BERTINI, A., GÁRATE, B., PARDO, F., PELICAND, J., SOBREVIA, L., TORRES, R., CHABERT, S. & SALAS, R. 2022. Impact of Remote Monitoring Technologies for Assisting Patients With Gestational Diabetes Mellitus: A Systematic Review. *Frontiers in Bioengineering and Biotechnology,* 10.

BIRATI, Y., YEFET, E., PERLITZ, Y., SHEHADEH, N. & SPITZER, S. 2022. Cultural and Digital Health Literacy Appropriateness of App- and Web-Based Systems Designed for Pregnant Women With Gestational Diabetes Mellitus: Scoping Review. *J Med Internet Res,* 24**,** e37844.

BLAIR, R. A., HORN, C. E., DIAS, J. M., MCDONNELL, M. E. & SEELY, E. W. 2022. Development and Usability of a Text Messaging Program for Women With Gestational Diabetes: Mixed Methods Study. *JMIR Human Factors,* 9**,** e32815.

BOGDANET, D., EGAN, A., REDDIN, C., KIRWAN, B., CARMODY, L. & DUNNE, F. 2018. ATLANTIC DIP: Despite insulin therapy in women with IADPSG diagnosed GDM, desired pregnancy outcomes are still not achieved. What are we missing? *Diabetes research and clinical practice.,* 136**,** 116-123.

BONACCORSO, G. 2018. Machine Learning Algorithms: Popular algorithms for data science and machine learning. Birmingham, United Kingdom: Packt Publishing Limited.

BORGEN, I., GARNWEIDNER-HOLME, L. M., JACOBSEN, A. F., BJERKAN, K., FAYYAD, S., JORANGER, P., LILLEENGEN, A. M., MOSDØL, A., NOLL, J., SMÅSTUEN, M. C., TERRAGNI, L., TORHEIM, L. E. & LUKASSE, M. 2017. Smartphone application for women with gestational diabetes mellitus: a study protocol for a multicentre randomised controlled trial. *BMJ Open,* 7**,** e013117.

BORGEN, I., SMÅSTUEN, M. C., JACOBSEN, A. F., GARNWEIDNER-HOLME, L. M., FAYYAD, S., NOLL, J. & LUKASSE, M. 2019. Effect of the Pregnant+ smartphone application in women with gestational diabetes mellitus: a randomised controlled trial in Norway. *BMJ Open,* 9**,** e030884.

BRAUN, V. & CLARKE, V. 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology,* 3**,** 77-101.

BRINK, H., RICHARDS, J. W. & FETHEROLF, M. 2017. *Real-world machine learning,* Shelter Island, New York, Manning Publications.

BROMURI, S., PURICEL, S., SCHUMANN, R., KRAMPF, J., RUIZ, J. & SCHUMACHER, M. 2016. An expert Personal Health System to monitor patients affected by Gestational Diabetes Mellitus: A feasibility study. *Journal of Ambient Intelligence and Smart Environments,* 8**,** 219-237.

BRYSON, C. L. 2003. Association between Gestational Diabetes and Pregnancy-induced Hypertension. *American journal of epidemiology.,* 158**,** 1148-1153.

BUCHANAN, T. A., XIANG, A. H. & PAGE, K. A. 2012. Gestational diabetes mellitus: risks and management during and after pregnancy. *Nature Reviews Endocrinology,* 8**,** 639-649.

BURNIE, R., GOLOB, E. & CLARKE, S. 2022. Pregnancy in underweight women: implications, management and outcomes. *The Obstetrician & Gynaecologist,* 24**,** 50-57.

BUTTEN, K., LAURIE, J. G., VARNFIELD, M. & WILKINSON, S. A. 2024. Gestational diabetes mellitus care re-imagined: women's experiences of a major model of care change at a large metropolitan hospital. *Practical Diabetes,* 41**,** 22-30.

CABALLERO-RUIZ, E., GARCÍA-SÁEZ, G., RIGLA, M., VILLAPLANA, M., PONS, B. & HERNANDO, M. E. 2016. Automatic classification of glycaemia measurements to enhance data interpretation in an expert system for gestational diabetes. *Expert Systems with Applications,* 63**,** 386-396.

CABALLERO-RUIZ, E., GARCÍA-SÁEZ, G., RIGLA, M., VILLAPLANA, M., PONS, B. & HERNANDO, M. E. 2017. A web-based clinical decision support system for gestational diabetes: Automatic diet prescription and detection of insulin needs. *International Journal of Medical Informatics,* 102**,** 35-49.

CAO, W., CAO, X. & SUTHERLAND, A. D. 2025. Planning for the Unexpected and Unintended Effects of mHealth Interventions: Systematic Review. *Journal of Medical Internet Research,* 27**,** e68909-e68909.

CARR, D. B., UTZSCHNEIDER, K. M., HULL, R. L., TONG, J., WALLACE, T. M., KODAMA, K., SHOFER, J. B., HECKBERT, S. R., BOYKO, E. J., FUJIMOTO, W. Y. & KAHN, S. E. 2006. Gestational diabetes mellitus increases the risk of cardiovascular disease in women with a family history of type 2 diabetes. *Diabetes Care,* 29**,** 2078-83.

CHAMBERLAIN, S., DUTT, P., MITRA, R., GODFREY, A., LEFEVRE, A. E., SCOTT, K., KATIYAR, S., MENDIRATTA, J. & CHATURVEDI, S. 2022. Lessons learnt from applying a human-centred design process to develop one of the largest mobile health communication programmes in the world. *BMJ Innovations,* 8**,** 240-246.

CHEN, I. Y., PIERSON, E., ROSE, S., JOSHI, S., FERRYMAN, K. & GHASSEMI, M. 2021a. Ethical Machine Learning in Healthcare. *Annual Review of Biomedical Data Science,* 4**,** 123-144.

CHEN, Q., ZHANG, B., YANG, J., MO, X., ZHANG, L., LI, M., CHEN, Z., FANG, J., WANG, F., HUANG, W., FAN, R. & ZHANG, S. 2021b. Predicting Intensive Care Unit Length of Stay After Acute Type A Aortic Dissection Surgery Using Machine Learning. *Frontiers in Cardiovascular Medicine,* 8**,** 675431.

CHOWDHURY, M. Z. I. & TURIN, T. C. 2020. Variable selection strategies and its importance in clinical prediction modelling. *Family Medicine and Community Health,* 8**,** e000262.

COLLIER, A., ABRAHAM, E. C., ARMSTRONG, J., GODWIN, J., MONTEATH, K. & LINDSAY, R. 2017. Reported prevalence of gestational diabetes in Scotland: The relationship with obesity, age, socioeconomic status, smoking and macrosomia, and how many are we missing? *Journal of Diabetes Investigation,* 8**,** 161-167.

COLLINS, G. S., MOONS, K. G. M., DHIMAN, P., RILEY, R. D., BEAM, A. L., VAN CALSTER, B., GHASSEMI, M., LIU, X., REITSMA, J. B., VAN SMEDEN, M., BOULESTEIX, A.-L., CAMARADOU, J. C., CELI, L. A., DENAXAS, S., DENNISTON, A. K., GLOCKER, B., GOLUB, R. M., HARVEY, H., HEINZE, G., HOFFMAN, M. M., KENGNE, A. P., LAM, E., LEE, N., LODER, E. W., MAIER-HEIN, L., MATEEN, B. A., MCCRADDEN, M. D., OAKDEN-RAYNER, L., ORDISH, J., PARNELL, R., ROSE, S., SINGH, K., WYNANTS, L. & LOGULLO, P. 2024. TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. *BMJ,* 385**,** e078378.

CONWAY, N. T., ALLARDICE, B., WAKE, D. J. & CUNNINGHAM, S. G. 2019. User Experiences of an Electronic Personal Health Record for Diabetes. *Journal of Diabetes Science and Technology,* 13**,** 744-750.

COORAY, S. D., WIJEYARATNE, L. A., SOLDATOS, G., ALLOTEY, J., BOYLE, J. A. & TEEDE, H. J. 2020. The Unrealised Potential for Predicting Pregnancy Complications in Women with Gestational Diabetes: A Systematic Review and Critical Appraisal. *International Journal of Environmental Research and Public Health,* 17**,** 3048.

COTTER, A. P., DURANT, N., AGNE, A. A. & CHERRINGTON, A. L. 2014. Internet interventions to support lifestyle modification for diabetes management: A systematic review of the evidence. *Journal of Diabetes and its Complications,* 28**,** 243-251.

CREDIT, C. 2022. *CRediT, Contributor Roles Taxonomy* [Online]. American National Standard Developed, National Information Standards Organization. Available: https://www.niso.org/publications/z39104-2022-credit [Accessed 08/05/2025 2025].

CRESSWELL, K., CALLAGHAN, M., KHAN, S., SHEIKH, Z., MOZAFFAR, H. & SHEIKH, A. 2020. Investigating the use of data-driven artificial intelligence in computerised decision support systems for health and social care: A systematic review. *Health Informatics Journal,* 26**,** 2138-2147.

CUNNINGHAM, S. G., WAKE, D. J., WALLER, A., MORRIS, A. D. & WALKER, J. 2013. My Diabetes My Way: an electronic personal health record for diabetes. *The British Journal of Diabetes & Vascular Disease,* 13**,** 143-149.

DALEY, B. J., NI'MAN, M., NEVES, M. R., HUDA, M. S. B., MARSH, W., FENTON, N. E., HITMAN, G. A. & MCLACHLAN, S. 2021. mHealth Apps for Gestational Diabetes Mellitus that provide Clinical Decision Support or Artificial Intelligence: A Scoping Review. *Diabetic Medicine,* 39**,** e14735.

DALFRÀ, M. G., NICOLUCCI, A. & LAPOLLA, A. 2009. The effect of telemedicine on outcome and quality of life in pregnant women with diabetes. *Journal of Telemedicine and Telecare,* 15**,** 238-242.

DALTON, J. A., RODGER, D., WILMORE, M., HUMPHREYS, S., SKUSE, A., ROBERTS, C. T. & CLIFTON, V. L. 2018. The Health-e Babies App for antenatal education: Feasibility for socially disadvantaged women. *PLOS ONE,* 13**,** e0194337.

DALY, B., TOULIS, K. A., THOMAS, N., GOKHALE, K., MARTIN, J., WEBBER, J., KEERTHY, D., JOLLY, K., SARAVANAN, P. & NIRANTHARAKUMAR, K. 2018. Increased risk of ischemic heart disease, hypertension, and type 2 diabetes in women with previous gestational diabetes mellitus, a target group in general practice for preventive interventions: A population-based cohort study. *PLOS Medicine,* 15**,** e1002488.

DANKWA-MULLAN, I., RIVO, M., SEPULVEDA, M., PARK, Y., SNOWDON, J. & RHEE, K. 2019. Transforming Diabetes Care Through Artificial Intelligence: The Future Is Here. *Population Health Management,* 22**,** 229-242.

DAVIDSEN, E., MAINDAL, H. T., ROD, M. H., OLESEN, K., BYRNE, M., DAMM, P. & NIELSEN, K. K. 2022. The stigma associated with gestational diabetes mellitus: A scoping review. *eClinicalMedicine,* 52**,** 101614.

DAVIDSEN, E., TERKILDSEN MAINDAL, H., BYRNE, M., KELSTRUP, L., OVESEN, P., DAMM, P. & KRAGELUND NIELSEN, K. 2023. A qualitative investigation into the perceptions and experiences of the stigma attached to gestational diabetes mellitus among women in Denmark. *Diabetes Research and Clinical Practice,* 203**,** 110858.

DE SOUZA, P. C., DA SILVA, A. G. A., SANTOS, C. M. D. A. M., SANTIAGO, L. A. C. D. C., ARAÚJO, M. E. D. O., DE LIMA, I. L. B. & RIBEIRO, K. D. D. S. 2024. Comparison of perinatal outcomes of women with gestational diabetes mellitus according to type of treatment for glycemic control. *Jornal de Pediatria,* 101**,** 179-186.

DEVORE, J. L. & BERK, K. N. 2012. *Modern mathematical statistics with applications,* New York ;, Springer.

DEVORE, J. L., BERK, K. N. & CARLTON, M. A. 2021. *Modern Mathematical Statistics with Applications,* Cham, Springer International Publishing.

DOUALI, N., DOLLON, J. & JAULENT, M.-C. Personalized prediction of gestational Diabetes using a clinical decision support system. IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), 2015 2015 Istanbul, Turkey,. IEEE, 1-5.

DRAFFIN, C. R., ALDERDICE, F. A., MCCANCE, D. R., MARESH, M., HARPER, M. R., MCSORLEY, O. & HOLMES, V. A. 2016. Exploring the needs, concerns and knowledge of women diagnosed with gestational diabetes: A qualitative study. *Midwifery,* 40**,** 141-147.

DU, Y., MCNESTRY, C., WEI, L., ANTONIADI, A. M., MCAULIFFE, F. M. & MOONEY, C. 2023. Machine learning-based clinical decision support systems for pregnancy care: A systematic review. *International Journal of Medical Informatics,* 173**,** 105040.

DU, Y., RAFFERTY, A. R., MCAULIFFE, F. M., WEI, L. & MOONEY, C. 2022. An explainable machine learning-based clinical decision support system for prediction of gestational diabetes mellitus. *Scientific Reports,* 12**,** 1170.

DUAN, B., LIU, Z., LIU, W. & GOU, B. 2022. Assessing the Views and Needs of People at High Risk of Gestational Diabetes Mellitus for the Development of Mobile Health Apps: Descriptive Qualitative Study. *JMIR Formative Research,* 6**,** e36392.

DUAN, B., ZHOU, Z., LIU, M., LIU, Z., ZHANG, Q., LIU, L., MA, C., GOU, B. & LIU, W. 2024. Development and acceptability of a gestational diabetes mellitus prevention system (Better pregnancy) based on a user-centered approach: A clinical feasibility study. *Digital health.,* 10**,** 1-14.

DUCARME, G., DESROYS DU ROURE, F., GRANGE, J., VITAL, M., LE THUAUT, A. & CRESPIN-DELCOURT, I. 2019. Predictive factors of subsequent insulin requirement for glycemic control during pregnancy at diagnosis of gestational diabetes mellitus. *International Journal of Gynecology and Obstetrics,* 144**,** 265-270.

EDWARDS, K. J., MASLIN, K., ANDRADE, J., JONES, R. B. & SHAWE, J. 2022. Mobile health as a primary mode of intervention for women at risk of, or diagnosed with, gestational diabetes mellitus: a scoping review. *JBI Evidence Synthesis,* 20(9)**,** 2195-2243.

EL-GAYAR, O., TIMSINA, P., NAWAR, N. & EID, W. 2013. A systematic review of IT for diabetes self-management: Are we there yet? *International Journal of Medical Informatics,* 82**,** 637-652.

ELEFTHERIADES, M., CHATZAKIS, C., PAPACHATZOPOULOU, E., PAPADOPOULOS, V., LAMBRINOUDAKI, I., DINAS, K., CHROUSOS, G. & SOTIRIADIS, A. 2021. Prediction of insulin treatment in women with gestational diabetes mellitus. *Nutrition and Diabetes,* 11.

FARAO, J., MALILA, B., CONRAD, N., MUTSVANGWA, T., RANGAKA, M. X. & DOUGLAS, T. S. 2020. A user-centred design framework for mHealth. *PLOS ONE,* 15**,** e0237910.

FARRAR, D., SIMMONDS, M., GRIFFIN, S., DUARTE, A., LAWLOR, D. A., SCULPHER, M., FAIRLEY, L., GOLDER, S., TUFFNELL, D., BLAND, M., DUNNE, F., WHITELAW, D., WRIGHT, J. & SHELDON, T. A. 2016. The identification and treatment of women with hyperglycaemia in pregnancy: an analysis of individual participant data, systematic reviews, meta-analyses and an economic evaluation. *Health Technol Assess,* 20**,** 1-348.

FEGHALI, M. N., ABEBE, K. Z., COMER, D. M., CARITIS, S., CATOV, J. M. & SCIFRES, C. M. 2019. Response to Medical Nutritional Therapy and Need for Pharmacological Therapy in Women with Gestational Diabetes. *Am J Perinatol,* 36**,** 1250-1255.

FERNÁNDEZ, A., GARCÍA, S., GALAR, M., PRATI, R. C., KRAWCZYK, B. & HERRERA, F. 2018. Data Level Preprocessing Methods. *In:* FERNÁNDEZ, A., GARCÍA, S., GALAR, M., PRATI, R. C., KRAWCZYK, B. & HERRERA, F. (eds.) *Learning from Imbalanced Data Sets.* Cham: Springer International Publishing.

FERRARA, A. 2007. Increasing Prevalence of Gestational Diabetes Mellitus. *Diabetes Care,* 30**,** S141-S146.

FORD, H. L., CHAMPION, I., WAN, A., REDDY, M., MOL, B. W. & ROLNIK, D. L. 2022. Predictors for insulin use in gestational diabetes mellitus. *European Journal of Obstetrics & Gynecology and Reproductive Biology,* 272**,** 177-181.

FRASER, E. E., OGDEN, K. J., RADFORD, A., INGRAM, E. R., CAMPBELL, J. E., DENNIS, A. & CORBOULD, A. M. 2023. Exploring the psychological wellbeing of women with gestational diabetes mellitus (GDM): increased risk of anxiety in women requiring insulin. A Prospective Longitudinal Observational Pilot Study. *Health Psychology and Behavioral Medicine,* 11.

FREGOSO-APARICIO, L., NOGUEZ, J., MONTESINOS, L. & GARCÍA-GARCÍA, J. A. 2021. Machine learning and deep learning predictive models for type 2 diabetes: a systematic review. *Diabetology & Metabolic Syndrome,* 13**,** 148.

FUNG, N. L. S., WIDYA, I., BROENS, T., LARBURU, N., BULTS, R., SHALOM, E., JONES, V. & HERMENS, H. 2014. Application of a Conceptual Framework for the

Modelling and Execution of Clinical Guidelines as Networks of Concurrent Processes. *Procedia Computer Science,* 35**,** 671-680.
GARCÍA-SÁEZ, G., RIGLA, M., MARTÍNEZ-SARRIEGUI, I., SHALOM, E., PELEG, M., BROENS, T., PONS, B., CABALLERO-RUÍZ, E., GÓMEZ, E. J. & HERNANDO, M. E. 2014. Patient-oriented Computerized Clinical Guidelines for Mobile Decision Support in Gestational Diabetes. *Journal of Diabetes Science and Technology,* 8**,** 238-246.
GARG, N., ARUNAN, S. K., ARORA, S. & KAUR, K. 2022. Application of Mobile Technology for Disease and Treatment Monitoring of Gestational Diabetes Mellitus Among Pregnant Women: A Systematic Review. *Journal of Diabetes Science and Technology,* 16**,** 491-497.
GARNWEIDNER-HOLME, L., BORGEN, I., GARITANO, I., NOLL, J. & LUKASSE, M. 2015. Designing and Developing a Mobile Smartphone Application for Women with Gestational Diabetes Mellitus Followed-Up at Diabetes Outpatient Clinics in Norway. *Healthcare,* 3**,** 310-323.
GARNWEIDNER-HOLME, L., HENRIKSEN, L., TORHEIM, L. E. & LUKASSE, M. 2020. Effect of the Pregnant+ Smartphone App on the Dietary Behavior of Women With Gestational Diabetes Mellitus: Secondary Analysis of a Randomized Controlled Trial. *JMIR mHealth and uHealth,* 8**,** e18614.
GARNWEIDNER-HOLME, L., HOEL ANDERSEN, T., SANDO, M. W., NOLL, J. & LUKASSE, M. 2018. Health Care Professionals' Attitudes Toward, and Experiences of Using, a Culture-Sensitive Smartphone App for Women with Gestational Diabetes Mellitus: Qualitative Study. *JMIR mHealth and uHealth,* 6**,** e123.
GELMAN, A. & HILL, J. 2006. Missing-data imputation. *In:* GELMAN, A. & HILL, J. (eds.) *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge: Cambridge University Press.
GENTILI, A., FAILLA, G., MELNYK, A., PULEO, V., TANNA, G. L. D., RICCIARDI, W. & CASCINI, F. 2022. The cost-effectiveness of digital health interventions: A systematic review of the literature. *Frontiers in public health,* 10**,** 787135-787135.
GHADERI, M., FARAHANI, M. A., BALOUCHI, A. & HAGHANI, H. 2022. The effect of smartphone-based education on self-efficacy of women with gestational diabetes: A controlled clinical trial. *Romanian Journal of Military Medicine,* 125**,** 330-337.
GIAXI, P., VIVILAKI, V., SARELLA, A., HARIZOPOULOU, V. & GOUROUNTI, K. 2025. Artificial Intelligence and Machine Learning: An Updated Systematic Review of Their Role in Obstetrics and Midwifery. *Cureus,* 17**,** e80394.
GIVEN, J. E., BUNTING, B. P., O'KANE, M. J., DUNNE, F. & COATES, V. E. 2015. Tele-Mum: A Feasibility Study for a Randomized Controlled Trial Exploring the Potential for Telemedicine in the Diabetes Care of Those with Gestational Diabetes. *Diabetes technology & therapeutics.,* 17**,** 88-888.
GOMES MANTOVANI, R., HORVÁTH, T., ROSSI, A. L. D., CERRI, R., BARBON JUNIOR, S., VANSCHOREN, J. & CARVALHO, A. C. P. L. F. D. 2024. Better trees: an empirical study on hyperparameter tuning of classification decision tree induction algorithms. *Data mining and knowledge discovery.,* 38**,** 1364-1416.
GOV.UK. 2021. *Antenatal Screening Programme Gestational diabetes* [Online]. Goverment. Available: https://view-health-screening-recommendations.service.gov.uk/gestational-diabetes/ [Accessed 14 April 2025].
GOYAL, P. & MALVIYA, R. 2023. Challenges and opportunities of big data analytics in healthcare. *Health Care Science,* 2**,** 328-338.

GUBRIUM, J. F., HOLSTEIN, J. A., MARVASTI, A. B. & MCKINNEY, K. D. 2012. *The SAGE handbook of interview research: The complexity of the craft*, Sage Publications.
GUO, P., CHEN, D., XU, P., WANG, X., ZHANG, W., MAO, M., ZHENG, Q., JIN, Y. & FENG, S. 2023. Web-Based Interventions for Pregnant Women With Gestational Diabetes Mellitus: Systematic Review and Meta-analysis. *Journal of Medical Internet Research,* 25**,** e36922.
GUPTA, S. & SEDAMKAR, R. R. 2020. Machine Learning for Healthcare: Introduction. *Learning and Analytics in Intelligent Systems.* Springer International Publishing.
HALDORAI, A., R, B. L., MURUGAN, S. & BALAKRISHNAN, M. 2024. Risk Prediction of Maternal Health by Model Analysis Using Artificial Intelligence. *In:* HALDORAI, A., R, B. L., MURUGAN, S. & BALAKRISHNAN, M. (eds.) *Artificial Intelligence for Sustainable Development.* Cham: Springer Nature Switzerland.
HANDELMAN, G. S., KOK, H. K., CHANDRA, R. V., RAZAVI, A. H., LEE, M. J. & ASADI, H. 2018. eDoctor: machine learning and the future of medicine. *Journal of Internal Medicine,* 284**,** 603-619.
HARPER, L. M., GLOVER, A. V., BIGGIO, J. R. & TITA, A. 2016. Predicting failure of glyburide therapy in gestational diabetes. *Journal of Perinatology,* 36**,** 347-351.
HARRISON, T. N., SACKS, D. A., PARRY, C., MACIAS, M., LING GRANT, D. S. & LAWRENCE, J. M. 2017. Acceptability of Virtual Prenatal Visits for Women with Gestational Diabetes. *Womens Health Issues,* 27**,** 351-355.
HASAN, M. E., ISLAM, M. R., CHEN, D., SANIN, C. & XU, G. 2023. Applications of Artificial Intelligence for Health Informatics: A Systematic Review. *Journal of Artificial Intelligence for Medical Sciences,* 0**,** 1-28.
HASHIMOTO, K. & KOGA, M. 2015. Indicators of glycemic control in patients with gestational diabetes mellitus and pregnant women with diabetes mellitus. *World Journal of Diabetes,* 6**,** 1045.
HIRST, J., MACKILLOP, L., LOERUP, L., FARMER, A., KEVAT, D., BARTLETT, K., GIBSON, O., KENWORTHY, Y., LEVY, J. & TARASSENKO, L. 2015a. GDm-health: development of a real-time smartphone solution for the management of women with gestational diabetes mellitus (GDM): C27. 01. *BJOG: An International Journal of Obstetrics and Gynaecology,* 122**,** 404.
HIRST, J. E., MACKILLOP, L., LOERUP, L., KEVAT, D. A., BARTLETT, K., GIBSON, O., KENWORTHY, Y., LEVY, J. C., TARASSENKO, L. & FARMER, A. 2015b. Acceptability and User Satisfaction of a Smartphone-Based, Interactive Blood Glucose Management System in Women With Gestational Diabetes Mellitus. *Journal of Diabetes Science and Technology,* 9**,** 111-115.
HIVERT, M.-F., BACKMAN, H., BENHALIMA, K., CATALANO, P., DESOYE, G., IMMANUEL, J., MCKINLAY, C. J. D., MEEK, C. L., NOLAN, C. J., RAM, U., SWEETING, A., SIMMONS, D. & JAWERBAUM, A. 2024. Pathophysiology from preconception, during pregnancy, and beyond. *The Lancet,* 40**,** 158-174.
HOSMER, D. W., LEMESHOW, S. & STURDIVANT, R. X. 2013. *Applied logistic regression,* Hoboken, New Jersey, Wiley.
HU, Y., SUN, Y., ZHANG, F. & GUO, B. SmartCarb: an intelligent mobile system to assist diet control for gestational diabetes patients using deep learning neural networks. International Conference of Health Informatics and Medical Systems, 2018.
HUXLEY, K., ATKINSON, P., DELAMONT, S., CERNAT, A., SAKSHAUG, J. W. & WILLIAMS, R. A. 2020. *Data cleaning,* London, SAGE Publications Ltd.

HYAMS, E., BRACKENRIDGE, A. & WHITE, S. L. 2024. Impact of the use of a smartphone-based blood glucose management system within a maternal diabetes service. *Practical Diabetes,* 41**,** 16-23.
ISLAM, M. N., MUSTAFINA, S. N., MAHMUD, T. & KHAN, N. I. 2022. Machine learning to predict pregnancy outcomes: a systematic review, synthesizing framework and future research agenda. *BMC Pregnancy and Childbirth,* 22**,** 348.
ISTEPANIAN, R. S. H. 2022. Mobile Health (m-Health) in Retrospect: The Known Unknowns. *International Journal of Environmental Research and Public Health,* 19**,** 3747.
JOHNS, E. C., DENISON, F. C., NORMAN, J. E. & REYNOLDS, R. M. 2018. Gestational Diabetes Mellitus: Mechanisms, Treatment, and Complications. *Trends in Endocrinology & Metabolism,* 29**,** 743-754.
JORDAN, M. I. & MITCHELL, T. M. 2015. Machine learning: Trends, perspectives, and prospects. *Science,* 349**,** 255-260.
KAVAKIOTIS, I., TSAVE, O., SALIFOGLOU, A., MAGLAVERAS, N., VLAHAVAS, I. & CHOUVARDA, I. 2017. Machine Learning and Data Mining Methods in Diabetes Research. *Computational and Structural Biotechnology Journal,* 15**,** 104-116.
KC, K., SHAKYA, S. & ZHANG, H. 2015. Gestational Diabetes Mellitus and Macrosomia: A Literature Review. *Annals of Nutrition and Metabolism,* 66**,** 14-20.
KENNELLY, M. A., AINSCOUGH, K., LINDSAY, K., GIBNEY, E., MC CARTHY, M. & MCAULIFFE, F. M. 2016. Pregnancy, exercise and nutrition research study with smart phone app support (Pears): Study protocol of a randomized controlled trial. *Contemp Clin Trials,* 46**,** 92-99.
KHALIL, C. 2019. Understanding the Adoption and Diffusion of a Telemonitoring Solution in Gestational Diabetes Mellitus: Qualitative Study. *JMIR Diabetes,* 4**,** e13661.
KHIN, M. O., GATES, S. & SARAVANAN, P. 2018. Predictors of metformin failure in gestational diabetes mellitus (GDM). *Diabetes & Metabolic Syndrome: Clinical Research & Reviews,* 12**,** 405-410.
KIRKWOOD, J. R., DICKSON, J., MANATAKI, A., LINDSAY, R. S., WAKE, D. J. & REYNOLDS, R. M. 2025a. User Needs Assessment for Gestational Diabetes Care and Management Digital Tool: A Qualitative Study. *Intelligent Health Systems – From Technology to Data and Knowledge.* Medical Informatics Europe.
KIRKWOOD, J. R., DICKSON, J., STEVENS, M., MANATAKI, A., LINDSAY, R. S., WAKE, D. J. & REYNOLDS, R. M. 2024. The User-Centered Design of a Clinical Dashboard and Patient-Facing App for Gestational Diabetes. *J Diabetes Sci Technol,* 0(0).
KIRKWOOD, J. R., GALLOWAY, N., LINDSAY, R. S., MANATAKI, A., WAKE, D. J. & REYNOLDS, R. M. 2025b. The use of machine learning to predict pharmacological therapy in gestational diabetes: A scoping review. *Diabetic Medicine*.
KLONOFF, D. C. & TRUE, M. W. 2009. The Missing Element of Telemedicine for Diabetes: Decision Support Software. *Journal of Diabetes Science and Technology,* 3**,** 996-1001.
KRAMER, C. K., CAMPBELL, S. & RETNAKARAN, R. 2019. Gestational diabetes and the risk of cardiovascular disease in women: a systematic review and meta-analysis. *Diabetologia,* 62**,** 905-914.
KUSYANTI, T., WIRAKUSUMAH, F. F., RINAWAN, F. R., MUHITH, A., PURBASARI, A., MAWARDI, F., PUSPITASARI, I. W., FAZA, A. & STELLATA, A. G. 2022.

Technology-Based (Mhealth) and Standard/Traditional Maternal Care for Pregnant Woman: A Systematic Literature Review. *Healthcare,* 10**,** 1287.
KYTO, M., HOTTA, S., NIINISTO, S., MARTTINEN, P., KORHONEN, T. E., MARKUSSEN, L. T., JACUCCI, G., SIEVANEN, H., VAHA-YPYA, H., KORHONEN, I., VIRTANEN, S., HEIHONEN, S. & KOIVUSALO, S. B. 2024. Periodic mobile application (eMOM) with self-tracking of glucose and lifestyle improves treatment of diet-controlled gestational diabetes without human guidance: a randomized controlled trial. *American Journal of Obstetrics and Gynecology*.
KYTÖ, M., HOTTA, S., NIINISTÖ, S., MARTTINEN, P., KORHONEN, T. E., MARKUSSEN, L. T., JACUCCI, G., SIEVÄNEN, H., VÄHÄ-YPYÄ, H., KORHONEN, I., VIRTANEN, S., HEINONEN, S. & KOIVUSALO, S. B. 2024. Periodic mobile application (eMOM) with self-tracking of glucose and lifestyle improves treatment of diet-controlled gestational diabetes without human guidance: a randomized controlled trial. *American journal of obstetrics and gynecology.,* 231**,** 541.e1-541.e16.
KYTÖ, M., KOIVUSALO, S., RUONALA, A., STRÖMBERG, L., TUOMONEN, H., HEINONEN, S. & JACUCCI, G. 2022a. Behavior Change App for Self-management of Gestational Diabetes: Design and Evaluation of Desirable Features. *JMIR Human Factors,* 9**,** e36987.
KYTÖ, M., KOIVUSALO, S., TUOMONEN, H., STRÖMBERG, L., RUONALA, A., MARTTINEN, P., HEINONEN, S. & JACUCCI, G. 2023. Supporting the Management of Gestational Diabetes Mellitus With Comprehensive Self-Tracking: Mixed Methods Study of Wearable Sensors. *JMIR Diabetes,* 8**,** e43979.
KYTÖ, M., MARKUSSEN, L. T., MARTTINEN, P., JACUCCI, G., NIINISTÖ, S., VIRTANEN, S. M., KORHONEN, T. E., SIEVÄNEN, H., VÄHÄ-YPYÄ, H., KORHONEN, I., HEINONEN, S. & KOIVUSALO, S. B. 2022b. Comprehensive self-tracking of blood glucose and lifestyle with a mobile application in the management of gestational diabetes: a study protocol for a randomised controlled trial (eMOM GDM study). *BMJ Open,* 12**,** e066292.
LAWRENCE, J. M., CONTRERAS, R., CHEN, W. & SACKS, D. A. 2008. Trends in the prevalence of preexisting diabetes and gestational diabetes mellitus among a racially/ethnically diverse population of pregnant women, 1999-2005. *Diabetes Care,* 31**,** 899-904.
LAZAREVIC, N., PIZZUTI, C., ROSIC, G., BŒHM, C., WILLIAMS, K. & CAILLAUD, C. 2023. A mixed-methods study exploring women's perceptions and recommendations for a pregnancy app with monitoring tools. *npj Digital Medicine,* 6**,** 50.
LIAO, L. D., FERRARA, A., GREENBERG, M. B., NGO, A. L., FENG, J., ZHANG, Z., BRADSHAW, P. T., HUBBARD, A. E. & ZHU, Y. 2022. Development and validation of prediction models for gestational diabetes treatment modality using supervised machine learning: a population-based cohort study. *BMC Medicine,* 20**,** 307.
LIEW, S. J., SOON, C. S., CHOOI, Y. C., TINT, M. T. & ERIKSSON, J. G. 2023. A holistic approach to preventing type 2 diabetes in Asian women with a history of gestational diabetes mellitus: a feasibility study and pilot randomized controlled trial. *Frontiers in Clinical Diabetes and Healthcare,* 4.
LIN, X., LIANG, C., LIU, J., LYU, T., GHUMMAN, N. & CAMPBELL, B. 2024. Artificial Intelligence–Augmented Clinical Decision Support Systems for Pregnancy Care: Systematic Review. *Journal of Medical Internet Research,* 26**,** e54737.
LU, H. Y., DING, X., HIRST, J. E., YANG, Y., YANG, J., MACKILLOP, L. & CLIFTON, D. 2023. Digital Health and Machine Learning Technologies for Blood Glucose

Monitoring and Management of Gestational Diabetes. *IEEE Reviews in Biomedical Engineering,* 17**,** 98-117.
MÄÄTTÄNEN, S., KOIVUSALO, S., YLINEN, H., HEINONEN, S. & KYTÖ, M. 2025. The Effect of a Mobile App (eMOM) on Self-Discovery and Psychological Factors in Persons With Gestational Diabetes: Mixed Methods Study. *JMIR mHealth and uHealth.,* 13**,** e60855.
MACKENZIE, S. C., SAINSBURY, C. A. R. & WAKE, D. J. 2024. Diabetes and artificial intelligence beyond the closed loop: a review of the landscape, promise and challenges. *Diabetologia,* 67**,** 223-235.
MACKILLOP, L., HIRST, J. E., BARTLETT, K. J., BIRKS, J. S., CLIFTON, L., FARMER, A. J., GIBSON, O., KENWORTHY, Y., LEVY, J. C., LOERUP, L., RIVERO-ARIAS, O., MING, W.-K., VELARDO, C. & TARASSENKO, L. 2018. Comparing the Efficacy of a Mobile Phone-Based Blood Glucose Management System With Standard Clinic Care in Women With Gestational Diabetes: Randomized Controlled Trial. *JMIR mHealth and uHealth,* 6**,** e71.
MACKILLOP, L., LOERUP, L., BARTLETT, K., FARMER, A., GIBSON, O. J., HIRST, J. E., KENWORTHY, Y., KEVAT, D. A., LEVY, J. C. & TARASSENKO, L. 2014. Development of a Real-Time Smartphone Solution for the Management of Women With or at High Risk of Gestational Diabetes. *Journal of diabetes science and technology,* 8**,** 1105-1114.
MACKILLOP, L. H., BARTLETT, K., BIRKS, J., FARMER, A. J., GIBSON, O. J., KEVAT, D. A., KENWORTHY, Y., LEVY, J. C., LOERUP, L., TARASSENKO, L., VELARDO, C. & HIRST, J. E. 2016. Trial protocol to compare the efficacy of a smartphone-based blood glucose management system with standard clinic care in the gestational diabetic population. *BMJ open* [Online], 6. Available: http://europepmc.org/abstract/MED/26988348 [Accessed 2016/03//].
MAGLIANO, D. J., BOYKO, E. J. & COMMITTEE, I. D. F. D. A. T. E. S. 2021. IDF Diabetes Atlas. *Idf diabetes atlas.* Brussels: International Diabetes Federation

MALEKI, F., MUTHUKRISHNAN, N., OVENS, K., REINHOLD, C. & FORGHANI, R. 2020. Machine Learning Algorithm Validation. *Neuroimaging clinics of North America.,* 30**,** 433-445.
MAYNE, I. K., LEVY, M. J., VAIDYA, B. & MCGOVERN, A. P. 2025. The glucose management of gestational diabetes in the UK: a national survey. *BMC Pregnancy and Childbirth,* 25.
MESHEL, S., SCHEJTER, E., HAREL, T., MASLOVITZ, S., GERMEZ, N., ELIMELECH, B., COHEN, B. & AZURI, J. 2016. Can we predict the need for pharmacological treatment according to demographic and clinical characteristics in gestational diabetes? *The Journal of Maternal-Fetal & Neonatal Medicine,* 29**,** 2062-2066.
MINSCHART, C., MAES, T., DE BLOCK, C., VAN POTTELBERGH, I., MYNGHEER, N., ABRAMS, P., VINCK, W., LEURIDAN, L., MATHIEU, C. & BILLEN, J. 2020. Mobile-based lifestyle intervention in women with glucose intolerance after gestational diabetes mellitus (MELINDA), a multicenter randomized controlled trial: Methodology and design. *Journal of Clinical Medicine,* 9**,** 2635.
MOHAMED, H., ISMAIL, A., SUTAN, R., RAHMAN, R. A. & JUVAL, K. 2025. A scoping review of digital technologies in antenatal care: recent progress and applications of digital technologies. *BMC Pregnancy and Childbirth,* 25.
MOREIRA, M. W. L., RODRIGUES, J. J. P. C., KUMAR, N., AL-MUHTADI, J. & KOROTAEV, V. 2018. Evolutionary radial basis function network for

gestational diabetes data analytics. *Journal of Computational Science,* 27**,** 410-417.
MUNN, Z., PETERS, M. D. J., STERN, C., TUFANARU, C., MCARTHUR, A. & AROMATARIS, E. 2018. Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. *BMC Medical Research Methodology,* 18.
MURPHY, H. R., RAYMAN, G., LEWIS, K., KELLY, S., JOHAL, B., DUFFIELD, K., FOWLER, D., CAMPBELL, P. J. & TEMPLE, R. C. 2008. Effectiveness of continuous glucose monitoring in pregnant women with diabetes: randomised clinical trial. *BMJ,* 337**,** a1680-a1680.
NAGRAJ, S., KENNEDY, S. H., JHA, V., NORTON, R., HINTON, L., BILLOT, L., RAJAN, E., ARORA, V., PRAVEEN, D. & HIRST, J. E. 2021. SMARThealth pregnancy: feasibility and acceptability of a complex intervention for high-risk pregnant women in Rural India: protocol for a pilot cluster randomised controlled trial. *Frontiers in Global Women's Health,* 2**,** 620759.
NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE (NICE). 2015, 25 February *Diabetes in pregnancy: management from preconception to the postnatal period,* [Online]. Available: https://www.nice.org.uk/guidance/ng3/resources/diabetes-in-pregnancy-management-from-preconception-to-the-postnatal-period-pdf-51038446021 [Accessed 22 October 2021].
NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE (NICE) 2020. National Institute for Health and Care Excellence: Guidelines. *Diabetes in pregnancy: management from preconception to the postnatal period.* London: National Institute for Health and Care Excellence (NICE)

NAZARPOUR, S., SIMBAR, M., KIANI, Z., KHALAJI, N., KHORRAMI KHARGH, M. & NAEIJI, Z. 2024. The relationship between quality of life and some mental problems in women with gestational diabetes mellitus (GDM): a cross-sectional study. *BMC Psychiatry,* 24**,** 511.
NEHA MARGRET, I., RAJAKUMAR, K., ARULALAN, K. V., MANIKANDAN, S. & VALENTINA 2024. Statistical Insights Into Machine Learning-Based Box Models for Pregnancy Care and Maternal Mortality Reduction: A Literature Survey. *IEEE Access,* 12**,** 68184-68207.
NISHIKAWA, T., ONO, K., HASHIMOTO, S., KINOSHITA, H., WATANABE, T., ARAKI, H., OTSU, K., SAKAMOTO, W., HARADA, M., TOYONAGA, T., KAWAKAMI, S., FUKUDA, J., HAGA, Y., KUKIDOME, D., TAKAHASHI, T. & ARAKI, E. 2018. One-hour oral glucose tolerance test plasma glucose at gestational diabetes diagnosis is a common predictor of the need for insulin therapy in pregnancy and postpartum impaired glucose tolerance. *Journal of Diabetes Investigation,* 9**,** 1370-1377.
ODENDAAL, W. A., ANSTEY WATKINS, J., LEON, N., GOUDGE, J., GRIFFITHS, F., TOMLINSON, M. & DANIELS, K. 2020. Health workers' perceptions and experiences of using mHealth technologies to deliver primary healthcare services: a qualitative evidence synthesis. *The Cochrane database of systematic reviews*.
PELEG, M., SHAHAR, Y., QUAGLINI, S., BROENS, T., BUDASU, R., FUNG, N., FUX, A., GARCÍA-SÁEZ, G., GOLDSTEIN, A., GONZÁLEZ-FERRER, A., HERMENS, H., HERNANDO, M. E., JONES, V., KLEBANOV, G., KLIMOV, D., KNOPPEL, D., LARBURU, N., MARCOS, C., MARTÍNEZ-SARRIEGUI, I., NAPOLITANO, C., PALLÀS, À., PALOMARES, A., PARIMBELLI, E., PONS, B., RIGLA, M., SACCHI,

L., SHALOM, E., SOFFER, P. & VAN SCHOOTEN, B. 2017a. Assessment of a personalized and distributed patient guidance system. *International Journal of Medical Informatics,* 101**,** 108-130.

PELEG, M., SHAHAR, Y., QUAGLINI, S., FUX, A., GARCÍA-SÁEZ, G., GOLDSTEIN, A., HERNANDO, M. E., KLIMOV, D., MARTÍNEZ-SARRIEGUI, I., NAPOLITANO, C., PARIMBELLI, E., RIGLA, M., SACCHI, L., SHALOM, E. & SOFFER, P. 2017b. MobiGuide: a personalized and patient-centric decision-support system and its evaluation in the atrial fibrillation and gestational diabetes domains. *User Modeling and User-Adapted Interaction,* 27**,** 159-213.

PERIYATHAMBI, N., PARKHI, D., GHEBREMICHAEL-WELDESELASSIE, Y., PATEL, V., SUKUMAR, N., SIDDHARTHAN, R., NARLIKAR, L. & SARAVANAN, P. 2022. Machine learning prediction of non-attendance to postpartum glucose screening and subsequent risk of type 2 diabetes following gestational diabetes. *PLOS ONE,* 17**,** e0264648.

PERTOT, T., MOLYNEAUX, L., TAN, K., ROSS, G. P., YUE, D. K. & WONG, J. 2011. Can Common Clinical Parameters Be Used to Identify Patients Who Will Need Insulin Treatment in Gestational Diabetes Mellitus? *Diabetes Care,* 34**,** 2214-2216.

PETERS, M. D. J., MARNIE, C., TRICCO, A. C., POLLOCK, D., MUNN, Z., ALEXANDER, L., MCINERNEY, P., GODFREY, C. M. & KHALIL, H. 2020. Updated methodological guidance for the conduct of scoping reviews. *JBI Evidence Synthesis,* 18**,** 2119-2126.

PINTAUDI, B., FRESA, R., DALFRÀ, M., DODESINI, A. R., VITACOLONNA, E., TUMMINIA, A., SCIACCA, L., LENCIONI, C., MARCONE, T., LUCISANO, G., NICOLUCCI, A., BONOMO, M. & NAPOLI, A. 2018. The risk stratification of adverse neonatal outcomes in women with gestational diabetes (STRONG) study. *Acta Diabetologica,* 55**,** 1261-1273.

PONCELET, C., BOUAMOUD, L., MICHEL, P. & CAMPINOS, C. 2024. Monitoring gestational diabetes mellitus patients with myDiabby Healthcare® smartphone application vs classical diary. Results from the non-inferiority TELESUR-GDM study. *Diabetes & Metabolism,* 50**,** 101502.

POULTER, S. E., MELONCELLI, N. & MACK, M. 2022. Use of a smartphone-based, interactive blood glucose management system in women with gestational diabetes mellitus: A pilot study. *Diabetes Research and Clinical Practice,* 185**,** 109224.

PUBLIC HEALTH SCOTLAND 2024a. Births in Scotland

PUBLIC HEALTH SCOTLAND. 2024b. *Births in Scotland* [Online]. Public Health Scotland. Available: https://publichealthscotland.scot/publications/births-in-scotland/births-in-scotland-year-ending-31-march-2024/ [Accessed 19 March 2025].

PUBLIC HEALTH SCOTLAND 2024c. Births in Scotland Technical Report.

PUSTOZEROV, E. & POPOVA, P. Mobile-based decision support system for gestational diabetes mellitus. Ural Symposium on Biomedical Engineering, Radioelectronics and Information Technology (USBEREIT), 2018 2018 Yekaterinburg, Russia. IEEE, 45-48.

PUSTOZEROV, E., POPOVA, P., TKACHUK, A., BOLOTKO, Y., YULDASHEV, Z. & GRINEVA, E. 2018. Development and Evaluation of a Mobile Personalized Blood Glucose Prediction System for Patients With Gestational Diabetes Mellitus. *JMIR mHealth and uHealth,* 6**,** e6.

RAMSPEK, C. L., JAGER, K. J., DEKKER, F. W., ZOCCALI, C. & VAN DIEPEN, M. 2021. External validation of prognostic models: what, why, how, when and where? *Clinical Kidney Journal,* 14**,** 49-58.

RASEKABA, T. M., FURLER, J., BLACKBERRY, I., TACEY, M., GRAY, K. & LIM, K. 2015. Telemedicine interventions for gestational diabetes mellitus: A systematic review and meta-analysis. *Diabetes Research and Clinical Practice,* 110**,** 1-9.

REBALA, G., RAVI, A. & CHURIWALA, S. 2019. *An Introduction to Machine Learning,* Cham, Springer International Publishing.

RESEARCH DATA SCOTLAND. 2024. *NHS Greater Glasgow and Clyde Demographics* [Online]. Available: https://www.researchdata.scot/metadata-catalogue/search-results/result/?id=834&originId=2 [Accessed 31 July 2025].

RHEE, S. Y. & RHEE, E. J. 2024. Chapter 1 - Trends in Digital Health for Diabetes. *In:* KLONOFF, D. C., KERR, D. & ESPINOZA, J. C. (eds.) *Diabetes Digital Health, Telehealth, and Artificial Intelligence*. Academic Press.

RIGLA, M., MARTÍNEZ-SARRIEGUI, I., GARCÍA-SÁEZ, G., PONS, B. & HERNANDO, M. E. 2018. Gestational Diabetes Management Using Smart Mobile Telemedicine. *Journal of Diabetes Science and Technology,* 12**,** 260-264.

RILEY, R. D., ENSOR, J., SNELL, K. I. E., HARRELL, F. E., MARTIN, G. P., REITSMA, J. B., MOONS, K. G. M., COLLINS, G. & VAN SMEDEN, M. 2020. Calculating the sample size required for developing a clinical prediction model. *BMJ,* 368**,** m441.

ROWAN, J. A. 2007. A Trial in Progress: Gestational Diabetes. *Diabetes Care,* 30**,** S214-S219.

ROWLAND, S. P., FITZGERALD, J. E., LUNGREN, M., LEE, E., HARNED, Z. & MCGREGOR, A. H. 2022. Digital health technology-specific risks for medical malpractice liability. *npj Digital Medicine,* 5.

RUOPP, M. D., PERKINS, N. J., WHITCOMB, B. W. & SCHISTERMAN, E. F. 2008. Youden Index and Optimal Cut-Point Estimated from Observations Affected by a Lower Limit of Detection. *Biometrical Journal,* 50**,** 419-430.

SAFIEE, L., ROUGH, D., GEORGE, P. & MUDENHA, R. 2023. Baseline Perceptions of Women With Gestational Diabetes Mellitus and Health Care Professionals About Digital Gestational Diabetes Mellitus Self-Management Health Care Technologies: Interview Study Among Patients and Health Care Professionals. *JMIR Human Factors,* 10**,** e51691.

SAFIEE, L., ROUGH, D., GEORGE, P. & MUDENHA, R. Engaging Women with Gestational Diabetes Mellitus in the Design of Self-Management Apps. DIS: Designing Interactive Systems, 2024 2024 Copenhagen Denmark. ACM.

SAFIEE, L., ROUGH, D. J. & WHITFORD, H. 2022. Barriers to and Facilitators of Using eHealth to Support Gestational Diabetes Mellitus Self-management: Systematic Literature Review of Perceptions of Health Care Professionals and Women With Gestational Diabetes Mellitus. *Journal of Medical Internet Research,* 24**,** e39689.

SAIF-UR-RAHMAN, K., ISLAM, M. S., ALABOSON, J., OLA, O., HASAN, I., ISLAM, N., MAINALI, S., MARTINA, T., SILENGA, E., MUYANGANA, M. & JOARDER, T. 2023. Artificial intelligence and digital health in improving primary health care service delivery in LMICs: A systematic review. *Journal of Evidence-Based Medicine,* 16**,** 303-320.

SAKAMOTO, J. L., CARANDANG, R. R., KHAREL, M., SHIBANUMA, A., YAROTSKAYA, E., BASARGINA, M. & JIMBA, M. 2022. Effects of mHealth on the psychosocial health of pregnant women and mothers: a systematic review. *BMJ Open,* 12**,** e056807.

SANDFELD, S. 2024. *Materials Data Science : Introduction to Data Mining, Machine Learning, and Data-Driven Predictions for Materials Science and Engineering / by Stefan Sandfeld,* Cham, Springer International Publishing.

SARMA, A. D. & DEVI, M. 2025. Artificial intelligence in diabetes management: transformative potential, challenges, and opportunities in healthcare. *Hormones : international journal of endocrinology and metabolism.,* 24(2)**,** 307-322.

SARNO, L., NEOLA, D., CARBONE, L., SACCONE, G., CARLEA, A., MICELI, M., IORIO, G. G., MAPPA, I., RIZZO, G., GIROLAMO, R. D., D'ANTONIO, F., GUIDA, M. & MARUOTTI, G. M. 2023. Use of artificial intelligence in obstetrics: not quite ready for prime time. *American Journal of Obstetrics & Gynecology MFM,* 5**,** 100792.

SCOTTISH CARE INFORMATION DIABETES COLLABORATION, S.-D. 2015. *SCI-Diabetes* [Online]. Available: https://www.sci-diabetes.scot.nhs.uk/ [Accessed 16 Jun 2022].

SCOTTISH INTERCOLLEGIATE GUIDELINES NETWORK, S. 2010 *Management of diabetes 2010, (SIGN publication no. 116)* [Online]. SIGN. Available: https://www.sign.ac.uk/assets/sign116.pdf [Accessed 25 November 2021].

SCOTTISH INTERCOLLEGIATE GUIDELINES NETWORK, S. 2024. *Management of diabetes in pregnancy 2024, (SIGN publication no. 171)* [Online]. SIGN. Available: http://www.sign.ac.uk [Accessed 03/03/2025 2025].

SHALOM, E., SHAHAR, Y., GOLDSTEIN, A., ARIEL, E., SHEINBERGER, M., FUNG, N., JONES, V. & VAN SCHOOTEN, B. 2015. Implementation of a Distributed Guideline-Based Decision Support Model Within a Patient-Guidance Framework. Springer International Publishing.

SHANMUGAVEL, A., SHAKYA, P. R., SHRESTHA, A., NEPAL, J., SHRESTHA, A., DANEAULT, J.-F. & RAWAL, S. 2024. Designing and Developing a Mobile App for Management and Treatment of Gestational Diabetes in Nepal: User-Centered Design Study. *JMIR Formative Research,* 8**,** e50823.

SHEN, J., CHEN, J., ZHENG, Z., ZHENG, J., LIU, Z., SONG, J., WONG, S. Y., WANG, X., HUANG, M., FANG, P.-H., JIANG, B., TSANG, W., HE, Z., LIU, T., AKINWUNMI, B., WANG, C. C., ZHANG, C. J. P., HUANG, J. & MING, W.-K. 2020. An Innovative Artificial Intelligence–Based App for the Diagnosis of Gestational Diabetes Mellitus (GDM-AI): Development Study. *Journal of Medical Internet Research,* 22**,** e21573.

SIMMONS, D., GUPTA, Y., HERNANDEZ, T. L., LEVITT, N., VAN POPPEL, M., YANG, X., ZAROWSKY, C., BACKMAN, H., FEGHALI, M. & NIELSEN, K. K. 2024. Call to action for a life course approach. *The lancet.,* 404**,** 193-214.

SKAR, J. B., GARNWEIDNER-HOLME, L. M., LUKASSE, M. & TERRAGNI, L. 2018. Women's experiences with using a smartphone app (the Pregnant+ app) to manage gestational diabetes mellitus in a randomised controlled trial. *Midwifery,* 58**,** 102-108.

SMYTH, S., CURTIN, E., TULLY, E., MOLPHY, Z. & BREATHNACH, F. 2022. Smartphone Apps for Surveillance of Gestational Diabetes: Scoping Review. *JMIR Diabetes,* 7**,** e38910.

SMYTH, S., MOLPHY, Z., TULLY, E., BOLAND, F. & BREATHNACH, F. M. 2025. Impact of a novel smartphone application for remote monitoring of gestational diabetes on glycaemic control and birth outcomes: a pilot observational study. *Scientific Reports,* 15**,** 23724.

SOUZA, A. C. R. L. A., COSTA, R. A., PAGANOTI, C. F., RODRIGUES, A. S., ZUGAIB, M., HADAR, E., HOD, M. & FRANCISCO, R. P. V. 2019. Can we stratify the risk for insulin need in women diagnosed early with gestational diabetes by fasting blood glucose? *The Journal of Maternal-Fetal & Neonatal Medicine,* 32**,** 2036-2041.

SPAULONCI, C. P., BERNARDES, L. S., TRINDADE, T. C., ZUGAIB, M. & FRANCISCO, R. P. V. 2013. Randomized trial of metformin vs insulin in the management of gestational diabetes. *American Journal of Obstetrics and Gynecology,* 209**,** 34.e1-34.e7.

SRIVASTAVA, Y., KHANNA, P. & KUMAR, S. Estimation of Gestational Diabetes Mellitus using Azure AI Services. Amity International Conference on Artificial Intelligence (AICAI), 2019 2019 Dubai. IEEE.

STILL, B. & CRANE, K. 2017. *Fundamentals of User-Centered Design*, CRC Press.

SWEETING, A., HANNAH, W., BACKMAN, H., CATALANO, P., FEGHALI, M., HERMAN, W. H., HIVERT, M.-F., IMMANUEL, J., MEEK, C., OPPERMANN, M. L., NOLAN, C. J., RAM, U., SCHMIDT, M. I., SIMMONS, D., CHIVESE, T. & BENHALIMA, K. 2024. Epidemiology and management of gestational diabetes. *The Lancet,* 404**,** 175-192.

TAMAGAWA, M., KASUGA, Y., SAISHO, Y., TANAKA, Y., HASEGAWA, K., OISHI, M., ENDO, T., SATO, Y., IKENOUE, S., TANAKA, M. & OCHIAI, D. 2021. Predictors of later insulin therapy for gestational diabetes diagnosed in early pregnancy. *Endocrine Journal,* 68**,** 1321-1328.

TANG, L., XU, S., LI, P. & LI, L. 2019. Predictors of Insulin Treatment During Pregnancy and Abnormal Postpartum Glucose Metabolism in Patients with Gestational Diabetes Mellitus. *Diabetes, Metabolic Syndrome and Obesity: Targets and Therapy,* 12**,** 2655-2665.

TARRY-ADKINS, J. L., AIKEN, C. E. & OZANNE, S. E. 2019. Neonatal, infant, and childhood growth following metformin versus insulin treatment for gestational diabetes: A systematic review and meta-analysis. *PLOS Medicine,* 16**,** e1002848.

TERRY, G. & HAYFIELD, N. 2021. *Essentials of thematic analysis / Gareth Terry, Nikki Hayfield,* Washington, DC, American Psychological Association.

TRIANTAFYLLIDIS, A. K. & TSANAS, A. 2019. Applications of Machine Learning in Real-Life Digital Health Interventions: Review of the Literature. *Journal of Medical Internet Research,* 21**,** e12286.

TRICCO, A. C., LILLIE, E., ZARIN, W., O'BRIEN, K. K., COLQUHOUN, H., LEVAC, D., MOHER, D., PETERS, M. D. J., HORSLEY, T., WEEKS, L., HEMPEL, S., AKL, E. A., CHANG, C., MCGOWAN, J., STEWART, L., HARTLING, L., ALDCROFT, A., WILSON, M. G., GARRITTY, C., LEWIN, S., GODFREY, C. M., MACDONALD, M. T., LANGLOIS, E. V., SOARES-WEISER, K., MORIARTY, J., CLIFFORD, T., TUNÇALP, Ö. & STRAUS, S. E. 2018. PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation. *Annals of Internal Medicine,* 169**,** 467-473.

VAN DEN HEUVEL, J. F., GROENHOF, T. K., VEERBEEK, J. H., VAN SOLINGE, W. W., LELY, A. T., FRANX, A. & BEKKER, M. N. 2018. eHealth as the Next-Generation Perinatal Care: An Overview of the Literature. *Journal of Medical Internet Research,* 20**,** e202.

VARNFIELD, M., REDD, C., STONEY, R. M., HIGGINS, L., SCOLARI, N., WARWICK, R., IEDEMA, J., RUNDLE, J. & DUTTON, W. 2021. M♡THer, an mHealth System to Support Women with Gestational Diabetes Mellitus: Feasibility and Acceptability Study. *Diabetes Technology & Therapeutics,* 23**,** 358-366.

VELARDO, C., CLIFTON, D., HAMBLIN, S., KHAN, R., TARASSENKO, L. & MACKILLOP, L. 2021. Toward a Multivariate Prediction Model of Pharmacological Treatment for Women With Gestational Diabetes Mellitus: Algorithm Development and Validation. *Journal of Medical Internet Research,* 23**,** e21435.

VENKATESH, K. K., BUSCHUR, E. O. & ZORK, N. M. 2024. Chapter 17 - Telehealth for Pregnant Individuals with Diabetes. *In:* KLONOFF, D. C., KERR, D. & ESPINOZA, J. C. (eds.) *Diabetes Digital Health, Telehealth, and Artificial Intelligence*. Academic Press.

VOLANSKI, W., PRADO, A. L. D., AL-LAHHAM, Y., TELEGINSKI, A., PEREIRA, F. S., ALBERTON, D., REGO, F. G. D. M., VALDAMERI, G. & PICHETH, G. 2019. d-GDM: A mobile diagnostic decision support system for gestational diabetes. *Archives of Endocrinology and Metabolism,* 65**,** 5.

VOUNZOULAKI, E., KHUNTI, K., ABNER, S. C., TAN, B. K., DAVIES, M. J. & GILLIES, C. L. 2020. Progression to type 2 diabetes in women with a known history of gestational diabetes: systematic review and meta-analysis. *BMJ,* 369**,** m1361.

VOUNZOULAKI, E., MIKSZA, J. K., ZACCARDI, F., TAN, B. K., DAVIES, M. J., KHUNTI, K. & GILLIES, C. L. 2024. Association of ethnicity and socioeconomic status with health outcomes in women with gestational diabetes: Clinical Practice Research Datalink cohort study. *Diabetes & Metabolic Syndrome: Clinical Research & Reviews,* 18**,** 103010.

WAKE, D. J., HE, J., CZESAK, A. M., MUGHAL, F. & CUNNINGHAM, S. G. 2016. MyDiabetesMyWay. *Journal of Diabetes Science and Technology,* 10**,** 1050-1058.

WANG, H., LI, N., CHIVESE, T., WERFALLI, M., SUN, H., YUEN, L., HOEGFELDT, C. A., ELISE POWE, C., IMMANUEL, J., KARURANGA, S., DIVAKAR, H., LEVITT, N., LI, C., SIMMONS, D. & YANG, X. 2022. IDF Diabetes Atlas: Estimation of Global and Regional Gestational Diabetes Mellitus Prevalence for 2021 by International Association of Diabetes in Pregnancy Study Group's Criteria. *Diabetes research and clinical practice.,* 183**,** 109050.

WANG, L., ZHANG, Z., WANG, D., CAO, W., ZHOU, X., ZHANG, P., LIU, J., FAN, X. & TIAN, F. 2023. Human-centered design and evaluation of AI-empowered clinical decision support systems: a systematic review. *Frontiers in Computer Science,* 5.

WATANABE, M., KATAYAMA, A., KAGAWA, H., OGAWA, D. & WADA, J. 2016. Risk Factors for the Requirement of Antenatal Insulin Treatment in Gestational Diabetes Mellitus. *Journal of Diabetes Research,* 2016**,** 1-6.

WESCHENFELDER, F., LOHSE, K., LEHMANN, T., SCHLEUßNER, E. & GROTEN, T. 2021. Predictors of Treatment Requirements in Women with Gestational Diabetes: A Retrospective Analysis. *Journal of Clinical Medicine,* 10**,** 4421.

WHITE, S. L., AYMAN, G., BAKHAI, C., HILLIER, T. A. & MAGEE, L. A. 2023. Screening and diagnosis of gestational diabetes. *BMJ,* 381**,** e071920.

WOLFF, R. F., MOONS, K. G. M., RILEY, R. D., WHITING, P. F., WESTWOOD, M., COLLINS, G. S., REITSMA, J. B., KLEIJNEN, J. & MALLETT, S. 2019. PROBAST: A Tool to Assess the Risk of Bias and Applicability of Prediction Model Studies. *Annals of Internal Medicine,* 170**,** 51.

WORLD HEALTH ORGANISATION (WHO). 2019. *Recommendations on digital interventions for health system strengthening* [Online]. Available: https://www.who.int/publications/i/item/9789241550505 [Accessed 16 April 2025].

WORLD HEALTH ORGANISATION (WHO) & INTERNATIONAL DIABETES FEDERATION (IDF). 2021, 13 April. *Diabetes key facts,* [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/diabetes [Accessed].

XIE, W., DAI, P., QIN, Y., WU, M., YANG, B. & YU, X. 2020. Effectiveness of telemedicine for pregnant women with gestational diabetes mellitus: an updated meta-analysis of 32 randomized controlled trials with trial sequential analysis. *BMC Pregnancy and Childbirth,* 20.

XU, J., CHEN, Y., ZHAO, J., WANG, J., CHEN, J., PAN, X., ZHANG, W., ZHENG, J., ZOU, Z., CHEN, X. & ZHANG, Y. 2024. Current status of electronic health literacy among pregnant women with gestational diabetes mellitus and their perceptions of online health information: a mixed-methods study. *BMC Pregnancy and Childbirth,* 24**,** 392.

YANG, P., LO, W., HE, Z.-L. & XIAO, X.-M. 2018. Medical nutrition treatment of women with gestational diabetes mellitus by a telemedicine system based on smartphones. *Journal of Obstetrics and Gynaecology Research,* 44**,** 1228-1234.

YE, W., LUO, C., HUANG, J., LI, C., LIU, Z. & LIU, F. 2022. Gestational diabetes mellitus and adverse pregnancy outcomes: systematic review and meta-analysis. *BMJ***,** e067946.

YEE, L. M., JACKSON, J., LEZIAK, K., NIZNIK, C. M., SABER, R., YEH, C. & SIMON, M. 2020. 750: SweetMama: Usability testing of a novel mobile application for diabetes education and support during pregnancy. *American Journal of Obstetrics & Gynecology,* 222**,** S474-S475.

YEE, L. M., LEZIAK, K., JACKSON, J., STROHBACH, A., SABER, R., NIZNIK, C. M. & SIMON, M. A. 2021. Patient and provider perspectives on a novel mobile health intervention for low-income pregnant women with gestational or type 2 diabetes mellitus. *Journal of diabetes science and technology,* 15**,** 1121-1133.

YERLIKAYA, G., FALCONE, V., STOPP, T., MITTLBÖCK, M., TURA, A., HUSSLEIN, P., EPPEL, W. & GÖBL, C. S. 2018. To Predict the Requirement of Pharmacotherapy by OGTT Glucose Levels in Women with GDM Classified by the IADPSG Criteria. *Journal of Diabetes Research,* 2018**,** 1-6.

ZACCARA, T. A., MIKAMI, F. C. F., PAGANOTI, C. F., FRANCISCO, R. P. V. & COSTA, R. A. 2023. Predicting insulin use among women with gestational diabetes diagnosed in oral glucose tolerance test. *BMC Pregnancy and Childbirth,* 23.

ZHANG, Y., SHAO, J., LI, F. & XU, X. 2016. Factors in Gestational Diabetes Mellitus Predicting the Needs for Insulin Therapy. *International Journal of Endocrinology,* 2016**,** 1-5.

ZHANG, Z., YANG, L., HAN, W., WU, Y., ZHANG, L., GAO, C., JIANG, K., LIU, Y. & WU, H. 2022. Machine Learning Prediction Models for Gestational Diabetes Mellitus: Meta-analysis. *Journal of Medical Internet Research,* 24**,** e26634.

ZHAO, M., YAO, Z., ZHANG, Y., MA, L., PANG, W., MA, S., XU, Y. & WEI, L. 2025. Predictive value of machine learning for the progression of gestational diabetes mellitus to type 2 diabetes: a systematic review and meta-analysis. *BMC Medical Informatics and Decision Making,* 25.

ZHU, Y. & ZHANG, C. 2016. Prevalence of Gestational Diabetes and Risk of Progression to Type 2 Diabetes: a Global Perspective. *Current Diabetes Reports,* 16.

ZORK, N. M. 2022. Telehealth for the Management of Diabetes in Pregnancy. *Current Diabetes Reports,* 22**,** 365–369.

APPENDIX

Appendix A *Update Daley et al. 2021 and Lu et al. 2023 reviews*

To give an overview of the current landscape of mHealth apps available for GDM for this thesis, I have conducted a literature review using the search strategies described in Daley et al.'s (Daley et al., 2021) scoping review, and Lu et al.'s (Lu et al., 2023) review of clinical trials of patient monitoring technologies in GDM.

Update of Daley et al. 2021 scoping review, 'mHealth Apps for Gestational Diabetes Mellitus that provide Clinical Decision Support or Artificial Intelligence: A Scoping Review'

Daley et al.'s (Daley et al., 2021) scoping review on GDM mHealth apps, which included AI or CDSS, was conducted between 2014 to 2019, searched PubMed, Medline, ScienceDirect, Scopus, Directory of Open Access Journals, Web of Science and Elsevier using the following search terms:

"(Gestational Diabetes) AND (Decision Support)",
"(Gestational Diabetes) AND (Artificial Intelligence)",
"(Gestational Diabetes) AND (Machine Learning)",
"(Gestational Diabetes) AND (Neural Networks)".

Papers, of which there were 18, were included if they

1) had a focus on GDM
2) presented an mHealth app
3) addressed at least one element of the GDM pathway
4) had either CDSS or AI.

Adapting Daley et al.'s search strategy I searched PubMed, Medline, ScienceDirect, Scopus for English only papers using the search term "("Gestational Diabetes" AND (Decision Support OR Artificial Intelligence OR Machine Learning OR Neural Networks)) AND ("smartphone" OR "mHealth" OR "mobile health" or "app")", between 01/01/2020-31/03/2025. This returned 558 papers for title and abstract screening, 540 were removed,

leaving 18 for full text review, and finally, two were extracted (Albert et al., 2020; Shen et al., 2020). The additional papers alongside the originally included paper in Daley et al.'s scoping review can be seen in **Table A.2**.

Table A.1 *mHealth apps that provide clinical decision support or artificial intelligence: Update of Daley et al. (2021) 'mHealth Apps for Gestational Diabetes Mellitus that provide Clinical Decision Support or Artificial Intelligence: A Scoping Review'.*

		Clinical focus					
Main author, year, reference	**App name**	**Diagnosis**	**Management**	**Ongoing support**	**Data collection between clinical visits**	**Artificial intelligence type**	**Data storage**
San Fung, 2014 (Fung et al., 2014)	Unnamed^		x	x		CIG[a]	D[b]
García-Sáez, 2014 (García-Sáez et al., 2014)	MobiGuide		x	x	x	CIG[a]	NM[c]
Douali, 2015 (Douali et al., 2015)	Unnamed	x				ML (D)[d]	NM[c]
Shalom, 2015 (Shalom et al., 2015)	MobiGuide		x	x		CIG[a]	D[b], S[e]
Caballero-Ruiz, 2016(Caballero-Ruiz et al., 2016)	SineDie		x		x	NM[c]	NM[c]
Bromuri, 2016 (Bromuri et al., 2016)	Unnamed		x	x	x	CIG[a]	S[e]

Caballero-Ruiz, 2017 (Caballero-Ruiz et al., 2017)	SineDie		x	x	x	ML (D)[d]	S[e]
Peleg (a), 2017 (Peleg et al., 2017a)	MobiGuide		x	x		CIG[a]	S[e]
Peleg (b), 2017 (Peleg et al., 2017b)	MobiGuide		x	x		CIG[a]	NM[c]
Abejirind, 2018 (Abejirinde et al., 2018)	Bliss4Midwives	x				NM[c]	D[b]
Mackillop, 2018 (Mackillop et al., 2018)	GDm-Health		x	x	x	NM[c]	S[e]
Moreira, 2018 (Moreira et al., 2018)	Unnamed	x				NN[f]	N[g]
Hu, 2018 (Hu et al., 2018)	SmartCarb		x	x		NN[f], ML(I)[h]	D[b]
Pustozerov (a), 2018 (Pustozerov et al., 2018)	Unnamed		x	x	x	NM[c]	D[b], S[e]
Pustozerov (b), 2018 (Pustozerov and Popova, 2018)	Unnamed		x	x	x	NM[c]	D[b], S[e]

Rigla, 2018 (Rigla et al., 2018)	MobiGuide		x	x	x	CIG[a]	D[b], S[e]
Srivastava, 2019 (Srivastava et al., 2019)	Unnamed	x				ML (D)[d]	N[g]
Volanski, 2019 (Volanski et al., 2019)	d-GDM	x				NM[c]	D[b]
***Albert, 2020 (Albert et al., 2020)**	SineDie		x	x	x	ML (D)[d]	NM[c]
***Shen, 2020 (Shen et al., 2020)**	GDM-AI / Care-2-You	x		x	x	ML (D)[d]	S[e]

*Table adapted from Daley et al. (2021) Tables 2 and 3. x indicates the clinical focus of app. ^Describes an app framework (not an app) called MADE, which is the foundation of the MobiGuide app.*Papers added from the updated original review by Daley et al. (2021). [a]CIG Computer interpretable guidelines, [b]D Device, [c]NM Not mentioned, [d]ML (D) Machine learning (data), [e]S Server, [f]NN Neural network, [g]N None, [h]ML(I) Machine learning (images).*

Update of Lu et al. 2023 review, 'Digital Health and Machine Learning Technologies for Blood Glucose Monitoring and Management of Gestational Diabetes', section titled: 'Emerging monitoring devices and management platforms for GDM: Devices and applications trialled for clinical observations and interventions'

Lu et al. gave an overview of digital health and machine learning technologies for blood glucose monitoring and management of GDM (Lu et al., 2023). Clinical trials of GDM devices and applications before 08/08/2021 were reviewed, using the keyword "gestational diabetes" on the National Institutes of Health (NIH) ClinicalTrials.gov database, including only completed trials. Lu et al. then summarised a selection of the 19 included clinical trials.

Applying the same search strategy, from 08/08/2021-31/03/2025, 35 studies were returned, of these seven studied digital health technologies for GDM patient monitoring. **Table A.2** provides the clinical trials studies included in Lu et al.'s (2023) review (Lu et al., 2023), along with the additional studies found through the updated search.

Table A.2 *Updated review of Lu et al. (2023) 'Digital Health and Machine Learning Technologies for Blood Glucose Monitoring and Management of Gestational Diabetes', Emerging Monitoring Devices and Management Platforms for Gestational Diabetes, Devices a nd Application*

Application/ Author, year, reference	**Study title**	**Location, phase (study size)**	**Outcome measures**	**Study type: primary purpose**
Treat-GDM (GDm-Health), 2013-2016 (Hirst et al., 2015a) 52 53	Trial of Remote Evaluation and Treatment of Gestational Diabetes Mellitus	UK, N/A (203)	Glycosylated haemoglobin; mean Blood glucose levels for fasting, pre-prandial, and postprandial readings; percentage of 'on target's blood glucose readings; effectiveness of monitoring; maternal outcomes; maternal weight gain; birthweight; birth injury; neonatal hypoglycaemia	Behavioural: Health services research, self home blood glucose monitoring
Pears, 2013-2016 (Kennelly et al., 2016)	Pregnancy, Exercise And Nutrition Research Study with app support	Ireland, N/A (750)	Effectiveness of a smartphone-assisted targeted healthy lifestyle intervention; Support antenatal management of overweight and the obese pregnant population in preventing GDM[a]	Interventional

Hola bebe, 2014-2017	Mobile Health App to Reduce Diabetes in Latina Women	USA, Phase 1 (18)	Self-efficacy for healthy eating, physical activity, and body weight.	Behavioural: Prevention
MobiGuide, 2014 (Peleg et al., 2017b; García-Sáez et al., 2014)	Group and Mobile Care for Gestational Diabetes	Spain, N/A (22)	Patient data is automatically or manually entered in the App; the decision support system generates feedback to patients and clinicians based on clinical guidelines.	Interventional
Pregnant+, 2015-2017 (Borgen et al., 2017)	A Mobile Smartphone Application to Promote a Healthy Diet and Physical Activity Among Pregnant Women With GDM	Norway, N/A (238)	OGTT[b]; dietary intake; motivation for eating healthy; physical activity; motivation for physical activity; depression; complications of pregnancy; mode of delivery and complications at birth for the mother; complications for the newborn.	Interventional: Prevention
Pregnant+, 2019-2022	Effects of Dietary and	China, N/A (2000)	Incidence of GDM[a], gestational hypertension, Caesarean birth and premature birth.	Behavioural: Intensive dietary, weight

	Weight Management on Pregnancy Outcomes in Mobile Medical Platform			management, standard care
Pregnant+, 2021-2022	Development and Testing of a Mobile Health Application for Management of Gestational Diabetes	Nepal, N/A (60)	Maternal blood glucose levels at 6 weeks postpartum; neonatal birth weight; induction of labour; Caesarean birth; self-monitoring adherence; usability of telemonitoring; app acceptability and usability	Behavioural: Supportive care.
University of Colorado Hospital, 2015-2021	User Testing and Feedback for a Mobile Health Program for Postpartum	USA, N/A (200)	Usability of the application, engagement with the application, navigability of the application, and acceptability of the application	Observational

	Women: A Pilot Study			
Peking Union Medical College Hospital, 2016-2020	The Effect of Mobile Medical Used for the Standardized Management of Gestational Diabetes	China, N/A (400)	Glycaemic qualification rate; pregnancy outcome.	Behavioural: Prevention, intensive dietary, weight management and standard care
iHealth, 2017-2018	Evaluating the Feasibility of Using mHealth to Improve Serum Glucose Logs	USA, N/A (8)	Glucose log completeness; patient satisfaction.	Interventional
FAB, 2017-2020	Fit After Baby: Increasing Postpartum Weight Loss in	USA, N/A (34)	Changes in weight loss, postpartum weight retention, waist circumference fasting insulin, fasting glucose, HbA1c[c], adherence to self-	Interventional: Prevention

	Women at Increased Risk for Cardiometabolic Disease		monitoring, Satisfaction, use of App, number of interactions with lifestyle coaches	
LIVING, 2017-2021	Lifestyle Intervention IN Gestational Diabetes	Bangladesh and India, N/A (1612)	Change of glycaemic category; changes in fasting blood glucose; change in body weight; change in waist circumference; change in systolic blood pressure; change in physical activity level; change in dietary habits.	Interventional: Prevention
SweetMama, 2017-2021 (Yee et al., 2020; Yee et al., 2021)	SweetMama: Testing of a Novel Technology for Diabetes Education and Support to Pregnant Women	USA, N/A (80)	Focus group and individual user feedback; feasibility; user interactivity data; usability testing; diabetes self-efficacy; patient activation; difference in HbA1c[c]	Interventional: Supportive care

GlucoseMama, 2018-2019	Group and Mobile Care for Gestational Diabetes	USA, N/A (22)	Reduction in the number of pharmacotherapy for treatment, neonates born large, and infants with neonatal hypoglycaemia; increased number of screened in the postpartum period for type 2 diabetes.	Interventional: supportive care
eMoMs, 2019-2022	Usability Study of the Sensors and eMOM GDM Application	Finland, N/A (34)	Acceptability and usability of sensors; acceptability of prototype app; technical functionality.	Observational
MELINDA, 2019-2022 (Minschart et al., 2020)	Mobile-based Lifestyle Intervention in Women With Glucose Intolerance After Gestational Diabetes	Belgium, N/A (236)	Metabolic syndrome; insulin resistance Matsuda; insulin resistance HOMA- IR[d] beta-cell function ISSI-2[e] index; beta-cell function insulinogenic index; beta-cell function HOMA-B[f]; weight loss	Behavioural: Prevention

GUIDES, 2021-2022	Gestational Diabetes in Uganda and India: Improving Screening and Self-management	India; Uganda, N/A (20000)	GDM[a] diagnosis; mean fasting blood glucose; adverse perinatal outcomes (composite measure); HbA1c[c]	Interventional
SMARThealth, 2019-2020 (Nagraj et al., 2021)	SMARThealth Pregnancy: Feasibility & Acceptability of a Complex Intervention for High- risk Pregnant Women in Rural India	India, N/A (258)	Recruitment rate, retention rate, number of home visits, number of women with GDM[a], postpartum follow-up; number of hypertensive disorders; number of severe anaemia; mean postpartum haemoglobin; mean postpartum; Systolic and diastolic blood pressure	Interventional: Prevention
Johns Hopkins Health System and	Pragmatic Randomised	USA, N/A (380)	Total gestational weight gain; number of participants who gained excess weight,	Interventional: Health services research

University, 2021-2024	Clinical Trial to Limit Weight Gain in Pregnancy and Prevent Obesity		incidence of GDM[a], postpartum weight retention, infant weight, proportion of low birth weight infants	
***MyDiabby Healthcare, 2013-2015, 2021 Poncelet (Poncelet et al., 2024)**	Monitoring gestational diabetes mellitus patients with myDiabby Healthcare® smartphone application vs a classical diary. Results from the non-inferiority TELESUR-GDM study	France, NA, (644)	Maternal, fetal and neonatal complications	Observational; Standard care; health service research.

***Women's Hospital, Louisiana, 2021-2023**	The CGM[g] in GDM Labor and Delivery Study (CGMSGDMLABOR)	USA, NA, (61)	Average percentage time in target glucose range, above range, and below range. Neonatal hypoglycaemia, postpartum 2hr 75g OGTT[b]	Interventional; Supportive Care
***Women's Hospital, Louisiana, 2021-2024**	Effectiveness of CGMS Vs. Self-monitoring Blood Glucose (SMBG) in Woman with Gestational Diabetes (STEADYSUGAR)	USA, NA, (128)	Average percentage time in the target glucose range, above range, and below range. Percentage gestational weight gain, HbA1c[c], large for gestational age newborns, neonatal hypoglycaemia. CGM[g]-satisfaction survey, Perceived benefit questionnaire.	Interventional; Diagnostic
***DiP GlucoMo, University Hospital of Bern, 2021-2024**	Utility of Real Time Continuous Glucose Monitoring in the	Switzerland, NA, (302)	Neonatal outcomes, diabetic medicine taken during pregnancy, percentage of glucose in range, HbA1c[c], mode of delivery, maternal characteristics	Interventional; Treatment

	Care of Gestational Diabetes Versus Standard Care in Pregnancy Outcomes (DiP GlucoMo)			
***Oura ring with Oura App, HealthifySG App, and Oura app, 2022-2023 (Liew et al., 2023)**	A holistic approach to preventing type 2 diabetes in Asian women with a history of gestational diabetes mellitus: a feasibility study and pilot	Singapore, NA, (61)	Change in lifestyle behaviour, OGTT[b], sleep quality, physical activity, postprandial glucose, quality of life, anxiety, depression, meals	Intervention; feasibility

	randomised controlled trial			
***Nurten Terkes, Mehmet Akif Ersoy University, 2021-2022**	The Effect of SMS Counselling for Gestational Diabetes on Self-Efficacy and Knowledge Levels	Turkey, NA, (95)	Diabetes knowledge, self-efficacy scale in GDM[a]	Interventional; Supportive Care
***Grenye O'Malley, Icahn School of Medicine at Mount Sinai, 2023-2024**	Postpartum Dysglycemia Screening With Continuous Glucose Monitoring	USA, NA (50)	Specificity, sensitivity, PPV[h], PLR[i], NPV[j], NLR[k], for Time in Range 70-180 mg/dL of <96% for abnormal OGTT is reported.	Interventional; Diagnostic

*Table adapted from Lu et al. (2023) Table 1. *Clinical trials added from the updated original review by Lu et al. (2023). [a]GDM Gestational diabetes mellites, [b]OGTT Oral glucose tolerance test, [c]HbA1c glycosylated haemoglobin, [d]HOMA-IR Homeostatic model assessment for insulin resistance, [e]ISSI-2 Insulin secretion- sensitivity index-2, [f]HOMA-B Homeostatic model assessment for beta cell function, [g]CGM Continuous glucose monitoring, [h]PPV Positive predictive value, [i]PLR Positive likelihood ratio, [j]NPV Negative predictive value, [k]NLR Negative likelihood ratio*

Appendix B *Supplementary Materials of The Use of Machine Learning Pharmacological Therapy in Gestational Diabetes: a Scoping Review [Published in the Journal of Diabetic Medicine]*

SUPPLEMENTARY MATERIAL 1: SEARCH TERMS

Embase

((machine.mp. or *supervised/) adj2 learning.mp.
OR
((artificial or computer or machine) adj2 intelligence).mp.
OR
(Regression or classification or ai or ml).mp
OR
exp artificial intelligence/
OR
exp machine learning/)
AND
(((insulin or metformin or glyburide) adj2 (failure or pregan*or need)).mp.
OR
(prediction adj3 (metformin or insulin or glyburide or drug or pharmacological or model or diet)).mp

OR
exp metformin/ad, dt, pd [Drug Administration, Drug Therapy, Pharmacology]
OR
exp insulin/ad, pd, th [Drug Administration, Pharmacology, Therapy]
OR
exp glibenclamide/ad, pd [Drug Administration, Pharmacology]
OR
exp diet therapy/
OR
medication therapy management/
OR
exp glucose blood level/
OR
exp glucose level/)
AND
(((gestation* or pregan* or maternal) adj3 (diabete* or hyperglyc?emia or glucose)).mp.
OR
(Fasting blood glucose* or fgd or bg or gdm).mp
OR
exp pregnancy diabetes mellitus/dm, dt, th [Disease Management, Drug Therapy, Therapy]))
AND
limit XX to yr="2007-2024"
AND

limit XX to (english language and embase and (article-in-press status or embase status))

IEEE Xplore

[manual] Filters:
2007 – 2024
Journals
((machine OR *supervised) NEAR/2 learning
OR
((artificial OR computer OR machine) NEAR/2 intelligence)
OR
Regression OR classification OR ai OR ml
OR
Artificial intelligence
OR
Exp machine learning)
AND
(prediction NEAR/3 (metformin OR insulin OR drug OR pharmacological OR model OR diet OR glyburide)
OR
((insulin OR metformin OR glyburide) NEAR/2 (failure OR pregan*OR need))
OR
exp Metformin/
OR
Insulin/
OR
Medication Therapy Management
OR
Exp Blood glucose drug therapy/ OR glycemic control)
AND
((gestation* OR pregan* OR maternal) NEAR/3 (diabete* OR hyperglyc?emia OR glucose)
OR
Fasting blood glucose* OR fgd OR bg OR gdm
OR
diabetes, gestational/ OR pregnancy in diabetics/)

Medline

((machine.mp. or *supervised/) adj2 learning.mp.
OR
((artificial or computer or machine) adj2 intelligence).mp.
OR
exp Artificial Intelligence/cl [Classification]
OR
exp Machine Learning/ or exp Algorithms/
OR
(Regression or classification or ai or ml).mp.
AND

(exp Drug Therapy/cl, mt [Classification, Methods]
OR
exp Glycemic Control/

OR
exp Metformin/ad, pd, tu [Administration & Dosage, Pharmacology, Therapeutic Use]
OR
exp Insulin/ad, pd, tu [Administration & Dosage, Pharmacology, Therapeutic Use]
OR
exp Diet Therapy/cl [Classification]
OR
exp Blood Glucose/ad, pd [Administration & Dosage, Pharmacology]
OR
exp Glyburide/ad, bl, pd [Administration & Dosage, Blood, Pharmacology]
OR
(prediction adj3 (metformin or insulin or glyburide or drug or pharmacological or model or diet)).mp
OR
((insulin or metformin or glyburide) adj2 (failure or pregan*or need)).mp
AND
(((gestation* or pregan* or maternal) adj3 (diabete* or hyperglyc?emia or glucose)).mp.
OR
(Fasting blood glucose* or fgd or bg or gdm).mp
OR
Pregnancy in Diabetics/ or exp Diabetes, Gestational/
OR
exp Pregnancy in Diabetics/ or exp Diabetes, Gestational/)
AND
((2007* or 2008* or 2009* or 2010* or 2011* or 2012* or 2013* or 2014* or 2015* or 2016* or 2017* or 2018* or 2019* or 2020* or 2021* or 2022* or 2023* or 2024*).ed.)

Web of Science

((((TS=((machine or supervised or unsupervised) NEAR/2 learning) OR TS=(((artificial or computer or machine) near/2 intelligence))) OR TS=(Regression or classification or ai or ml))))
AND
(TS=(prediction near/3 (metformin or insulin or glyburide or drug or pharmacological or model or diet)) or TS=(((insulin or glyburide or metformin) near (failure))) or TS=(((insulin or metformin or glyburide) near (pregan*))) or TS=(((insulin or metformin or glyburide) near (need))))
AND
(TS=((gestation* or pregan* or maternal) NEAR/3 (diabete* or hyperglyc?emia or glucose)) or TS=(Fasting blood glucose* or fgd or bg or gdm))

MANUAL
Year ranges 2007-2024

Language English
Exclude all conference abstracts
Include articles, review articles

SUPPLEMENTARY MATERIAL 2: SUMMARY TABLE OF STUDY CHARACTERISTICS

	Author	Country	Study design	Type of prediction study	GDM[a] diagnostic criteria	Data collection period (years)	Number of participants	Algorithm used in each model described in the paper	The performance of the model	Aims relevant to review	Outcomes relevant to review
Predicting pharmacological therapy	Feghali et al. 2019 (Feghali et al., 2019)	United States	Retrospective cohort study	Development and validation	Carpenter and Coustan criteria	3	1174	Logistic regression	AUROC[i] 0.71	To identify if the initial response to diet can predict the need for pharmacological therapy in women with GDM[a]	Glucose values from the first week of diet were the strongest predictor of needing pharmacological therapy.
								Logistic regression	AUROC[i] 0.83 Sensitivity 72.2%, Specificity 86.8%, PPV[j] 92.5%, NPV[k] 58%		
	Liao et al. 2022 (Liao et al., 2022)	United States	Population-based	Development and validation	Carpenter and	10	30474	CART[e]	AUROC[i] 0.613 (0.603-0.622, 95% CI[l])	To investigate whether clinical data at varied	Clinical data demonstrated reasonably high

	Author	Country	Study design	Type of prediction study	GDM[a] diagnostic criteria	Data collection period (years)	Number of participants	Algorithm used in each model described in the paper	The performance of the model	Aims relevant to review	Outcomes relevant to review
			cohort study		Coustan criteria			CART[e]	AUROC[i] 0.618 (0.609-0.628 95% CI[l])	stages of pregnancy can predict GDM[a] treatment modality.	predictability for GDM[a] treatment modality at the time of GDM[a] diagnosis and high predictability at 1-week post GDM[a] diagnosis.
								CART[e]	AUROC[i] 0.740 (0.732-0.748, 95% CI[l])		
								CART[e]	AUROC[i] 0.785 (0.777-0.792, 95% CI[l])		
								LASSO[f]	AUROC[i] 0.670 (0.663-0.676, 95% CI[l])		
								LASSO[f]	AUROC[i], 0.685 (0.678-0.691, 95% CI[l])		

	Author	Country	Study design	Type of prediction study	GDM[a] diagnostic criteria	Data collection period (years)	Number of participants	Algorithm used in each model described in the paper	The performance of the model	Aims relevant to review	Outcomes relevant to review
								LASSO[f]	AUROC[i] 0.785 (0.780-0.791, 95% CI[l])		
								LASSO[f]	AUROC[i] 0.849 (0.845-0.854, 95% CI[l])		
								Simple super learner[g]	AUROC[i] 0.673 (0.667-0.679, 95% CI[l])		
								Simple super learner[g]	AUROC[i] 0.688 (0.682-0.695, 95% CI[l])		
								Simple super learner[g]	AUROC[i] 0.790 (0.785-0.796, 95% CI[l])		

	Author	Country	Study design	Type of prediction study	GDM[a] diagnostic criteria	Data collection period (years)	Number of participants	Algorithm used in each model described in the paper	The performance of the model	Aims relevant to review	Outcomes relevant to review
								Simple super learner[g]	AUROC[i] 0.852 (0.848-0.857, 95% CI[l])		
								Complex super learner[h]	AUROC[i] 0.683 (0.676-0.689, 95% CI[l])		
								Complex super learner[h]	AUROC[i] 0.761 (0.756-0.767, 95% CI[l])		
								Complex super learner[h]	AUROC[i] 0.869 (0.865-0.873, 95% CI[l])		
								Complex super learner[h]	AUROC[i] 0.934 (0.931-0.936, 95% CI[l])		
								Logistic regression	AUROC[i] 0.632 (0.623-0.640, 95% CI[l])		

	Author	Country	Study design	Type of prediction study	GDM[a] diagnostic criteria	Data collection period (years)	Number of participants	Algorithm used in each model described in the paper	The performance of the model	Aims relevant to review	Outcomes relevant to review
								Logistic regression	AUROC[i] 0.648 (0.640-0.656, 95% CI[l])		
								Logistic regression	AUROC[i] 0.770 (0.764-0.775, 95% CI[l])		
								Logistic regression	AUROC[i] 0.825 (0.820-0.830, 95% CI[l])		
	Velardo et al. 2021 (Velardo et al., 2021)	United Kingdom	Retrospective cohort study	Development and validation	IADSPG[c] and national guidelines	5	1789	Logistic regression	Median AUROC[i] 0.8	To assess whether data collected through a mHealth system can be analysed to automatically evaluate the switch to pharmacological	Using real-time data collected via a mHealth system may further improve the timeliness of the intervention and potentially improve patient care

	Author	Country	Study design	Type of prediction study	GDM[a] diagnostic criteria	Data collection period (years)	Number of participants	Algorithm used in each model described in the paper	The performance of the model	Aims relevant to review	Outcomes relevant to review
										treatment from diet-based management of GDM.	
	Yerlikaya et al. 2018 (Yerlikaya et al., 2018)	Austria	Retrospective cohort study	Development	IADSPG[c]	2	203	Logistic regression	AUROC[i] 71.1	To assess the association between OGTT[m] glucose levels and the requirement of pharmacotherapy in GDM[a] patients	OGTT[m] glucose measures in addition to clinical risk factors showed promising properties for risk stratification in GDM[a] patients classified by the recently established IADPSG[c] criteria.
								Logistic regression	AUROC[i] 72.1		
								Logistic regression	AUROC[i] 77.5		
								Random forest	Performance not reported.		

	Author	Country	Study design	Type of prediction study	GDM[a] diagnostic criteria	Data collection period (years)	Number of participants	Algorithm used in each model described in the paper	The performance of the model	Aims relevant to review	Outcomes relevant to review
										classified by the IADPSG[c] criteria.	
Predicting insulin	Barnes et al. 2016 (Barnes et al., 2016)	Australia	Prospective cohort study	Development and validation	ADIPS[b]	23	3317	Logistic regression	AUROC[i] 0.710 (0.675-0.745, 95% CI[l])	To identify women with GDM who are more likely to require insulin therapy vs diet alone.	A validated model has been shown to predict therapy type and the likelihood of several adverse perinatal outcomes in women with GDM[a].
	Ducarme et al. 2019 (Ducarme et al., 2019)	France	Secondary analysis of a prospective observational study	Development and validation	IADSPG[c] and national guidelines	1	200	Logistic regression	HbA1c[o]: AUROC[i] 0.58 (0.48–0.67, 95% CI[l]), Sensitivity: 29.7%, Specificity: 87.7%, PPV[j] 59.4%. 1hr OGTT[m]:	To determine maternal and biological parameters at diagnosis of GDM[a] as predictors of antenatal insulin therapy for	HbA1c[o] at diagnosis of GDM[a] and elevated 1-hour OGTT[m] were independent predictors of insulin for glycaemic control.

	Author	Country	Study design	Type of prediction study	GDM[a] diagnostic criteria	Data collection period (years)	Number of participants	Algorithm used in each model described in the paper	The performance of the model	Aims relevant to review	Outcomes relevant to review
									AUROC[i] 0.62 (0.50–0.74, 95% CI[l]), Sensitivity: 29.7, Specificity: 66.7%, PPV[j] 44.0%.	glycaemic control.	
	Eleftheriades et al. 2021 (Eleftheriades et al., 2021)	Greece	Prospective cohort study	Development and validation	IADSPG[c]	8	775	CART[e]	AUROC[i] 0.75 (0.7-0.78, 95% CI[l])	To develop a predictive model for the necessity of insulin treatment in women with GDM[a].	Overweight women with an abnormal baseline blood glucose at OGTT[m] have a high likelihood of insulin treatment.
	Ford et al. 2022 (Ford et al., 2022)	Australia	Retrospective cohort study	Development and validation	ADIPS[b]	1	2048	Logistic regression	AUROC[i] 0.744 (0.720-0.767, 95% CI[l])	To identify predictors of insulin therapy in women	Reasonable prediction of the need for insulin use can be achieved with

	Author	Country	Study design	Type of prediction study	GDM[a] diagnostic criteria	Data collection period (years)	Number of participants	Algorithm used in each model described in the paper	The performance of the model	Aims relevant to review	Outcomes relevant to review
										diagnosed with GDM[a] once an OGTT[m] is performed during pregnancy.	information routinely collected in antenatal care.
	Harper et al. 2016 (Harper et al., 2016)	United States	Retrospective cohort study	Development and validation	Carpenter and Coustan criteria	6	360	Logistic regression	AUROC[i] 0.86, Sensitivity 39.7%, Specificity 97.5%	To develop a prediction model to identify women with GDM[a] who require insulin to achieve glycaemic control.	Women with GDM[a] who will require insulin can be identified at the initiation of pharmacological therapy.
								Logistic regression	AUROC[i] 0.87, Sensitivity 30.8%, Specificity 98.8%		

	Author	Country	Study design	Type of prediction study	GDM[a] diagnostic criteria	Data collection period (years)	Number of participants	Algorithm used in each model described in the paper	The performance of the model	Aims relevant to review	Outcomes relevant to review
	Khin et al. 2018 (Khin et al., 2018)	United Kingdom	Retrospective cohort study	Development	IADSPG[c]	3	228	Logistic regression	Specificity 64%, Sensitivity 87%, PPV[j] 74%, NPV[k] 70%,	To identifying the characteristics of these women with GDM[a] may help insulin in addition to metformin define optimal therapeutic strategy.	Women with higher fasting glucose levels have a higher chance of necessitating insulin in later pregnancies.
	Nishikawa et al. 2018 (Nishikawa et al., 2018)	Japan	Retrospective cohort study	Development	IADSPG[c]	1	529	Logistic regression	AUROC[i] 0.723, Sensitivity 67.3%, Specificity 67.3%	To identify factors predicting the need for insulin therapy in GDM[a] patients.	Antepartum 1-h glucose levels in a 75-g OGTT[m] were a predictor of the need for insulin therapy in pregnancy.
	Souza et al. 2019 (Souza et al., 2019)	Brazil	Retrospective	Development and validation	IADSPG[c]	3	408	Logistic regression	Specificity 90%, Sensitivity	To evaluate risk factors and propose a model	The need for insulin therapy in women with early diagnosis

	Author	Country	Study design	Type of prediction study	GDM[a] diagnostic criteria	Data collection period (years)	Number of participants	Algorithm used in each model described in the paper	The performance of the model	Aims relevant to review	Outcomes relevant to review
			cohort study						36%, PPV[j] 66%, NPV[k] 74%, Accuracy 74%	for the prediction of insulin requirement during the treatment of early-diagnosed GDM[a].	of GDM[a] can be predicted by a logistic regression model.
	Tamagawa et al. 2021 (Tamagawa et al., 2021)	Japan	Retrospective cohort study	Development	IADSPG[c] and national guidelines	9	388	Logistic regression	Pre-pregnancy BMI[n]: AUROC[i] 0.62, Sensitivity 63.3%, Specificity 57.8%, PPV[j] 54.4%, NPV[k] 73.8%. Fasting plasma glucose: AUROC[i] 0.52,	To elucidate factors that predict patients with GDM[a] diagnosed before 24 gestational weeks who require insulin therapy later during pregnancy.	Women with an early diagnosis of GDM[a], a pre-pregnancy BMI[n] ≥25 kg/m^2, and a family history of diabetes are more likely to require insulin therapy later during pregnancy.

	Author	Country	Study design	Type of prediction study	GDM[a] diagnostic criteria	Data collection period (years)	Number of participants	Algorithm used in each model described in the paper	The performance of the model	Aims relevant to review	Outcomes relevant to review
									Sensitivity 20.1%, Specificity 92.0%, PPV[j] 58.3%, NPV[k] 82.3%. 1hr plasma glucose: AUROC[i] 0.77, Sensitivity 71.2%, Specificity 74.7%, PPV[j] 61.1%, NPV[k] 84.3%. 2hr plasma glucose: AUROC[i] 0.75,		

	Author	Country	Study design	Type of prediction study	GDM[a] diagnostic criteria	Data collection period (years)	Number of participants	Algorithm used in each model described in the paper	The performance of the model	Aims relevant to review	Outcomes relevant to review
									Sensitivity 78.4%, Specificity 64.7%, PPV[j] 61.1%, NPV[k] 84.3%.		
	Tang et al. 2019 (Tang et al., 2019)	China	Retrospective cohort study	Development	IADSPG[c] and national guidelines	3	534	Logistic regression	Fasting plasma glucose 5.7 mmol/L: AUROC[i] 07.88, (0.704–0.872 CI[l]) Sensitivity 59.6%, Specificity 89.9%, 1hr plasma glucose, 11.4 mmol/L:	To investigate the potential predictors of insulin treatment during pregnancy and abnormal postpartum glucose metabolism in GDM[a].	Patients with fasting plasma glucose >5.7 mmol/L, 1 h plasma glucose >11.4 mmol/L, or HbA1c[o] >5.3% of GDM[a] diagnoses required insulin treatment. With fasting plasma glucose at GDM[a] diagnosis was the most important predictor.

	Author	Country	Study design	Type of prediction study	GDM[a] diagnostic criteria	Data collection period (years)	Number of participants	Algorithm used in each model described in the paper	The performance of the model	Aims relevant to review	Outcomes relevant to review
									AUROC[i] 0.642, (0.540-0.744 CI[i]) Sensitivity 34.0%, Specificity 94.4%, HbA1c[o] 5.3% AUROC[i] 0.683, (0.587-0.779 CI[i]) Sensitivity 59.6%, Specificity 70.8%,		

	Author	Country	Study design	Type of prediction study	GDM[a] diagnostic criteria	Data collection period (years)	Number of participants	Algorithm used in each model described in the paper	The performance of the model	Aims relevant to review	Outcomes relevant to review
	Watanabe et al. 2016 (Watanabe et al., 2016)	Japan	Retrospective cohort study	Development	IADSPG[c]	6	37	Logistic regression	1hr plasma glucose, 10.25 mmol/L: AUROC[i] 0.872, Sensitivity 100%, Specificity 77.8%. 2hr plasma glucose, 8.75 mmol/L: AUROC[i] 0.756, Sensitivity 70%, Specificity 70.4%. 75g OGTT[m] 1.5: AUROC[i] 0.783,	To investigate the clinical characteristics of patients with GDM[a] to identify risk factors for antenatal insulin treatment.	1- h plasma glucose levels in 75 g OGTT[m] are useful parameters in predicting the requirement for insulin in GDM[a].

	Author	Country	Study design	Type of prediction study	GDM[a] diagnostic criteria	Data collection period (years)	Number of participants	Algorithm used in each model described in the paper	The performance of the model	Aims relevant to review	Outcomes relevant to review
									Sensitivity 80%, Specificity 74.1%		
	Weschenfelder et al. 2021 (Weschenfelder et al., 2021)	Germany	Retrospective cohort study	Development	IADSPG[c] and national guidelines	5	454	Logistic regression	Fasting plasma glucose 5.5 mmol/L: AUROC[i] 0.643 (0.590-0.696, 95% CI[l]) Specificity 84.4%, Sensitivity 42.5%, PPV[j] 69.3%, NPV[k] 63.9%.	To find predictors of both the general insulin requirement as well as for the described treatment subgroups within women diagnosed with GDM[a] after 24	Significant cut-offs for insulin dependency were HbA1c[o] level of 5.4%, fasting plasma glucose of 5.5 mmol/L and 1 hr glucose of 10.6 mmol/L.

	Author	Country	Study design	Type of prediction study	GDM[a] diagnostic criteria	Data collection period (years)	Number of participants	Algorithm used in each model described in the paper	The performance of the model	Aims relevant to review	Outcomes relevant to review
									1hr glucose 10.6 mmol/L AUROC[i] 0.643 (0.582-0.686, 95% CI[l]) Specificity 76.7%, Sensitivity 45.3%, PPV[j] 55.9%, NPV[k] 68.3%. HbA1c[o], 5.4%: AUROC[i] 0.653 (0.603-0.675, 95% CI[l]) Specificity 84.4%, Sensitivity 42.5%,	weeks of gestation.	

	Author	Country	Study design	Type of prediction study	GDM[a] diagnostic criteria	Data collection period (years)	Number of participants	Algorithm used in each model described in the paper	The performance of the model	Aims relevant to review	Outcomes relevant to review
									PPV[j] 51.9%, NPV[k] 69.4%.		
								Logistic regression	1hr glucose, 10.6 mmol/L: AUROC[i] 0.63 (0.55-0.711, 95% CI[l]) Specificity 70.2%, Sensitivity 51.1%, PPV[j] 15.9%, NPV[k] 92.9%. HbA1c[o], 5.4%: AUROC[i] 0.6 (0.525-0.675, 95% CI[l])		

	Author	Country	Study design	Type of prediction study	GDM[a] diagnostic criteria	Data collection period (years)	Number of participants	Algorithm used in each model described in the paper	The performance of the model	Aims relevant to review	Outcomes relevant to review
									Specificity 60.1%, Sensitivity 57.8%, PPV[j] 13.8%, NPV[k] 92.8%.		
								Logistic regression	Fasting plasma glucose, 5.2 mmol/L: AUROC[i] 0.613 (0.543-0.682, 95% CI[l]) Specificity 48.9%, Sensitivity 70.5%, PPV[j] 17.6%, NPV[k] 91.4%.		

	Author	Country	Study design	Type of prediction study	GDM[a] diagnostic criteria	Data collection period (years)	Number of participants	Algorithm used in each model described in the paper	The performance of the model	Aims relevant to review	Outcomes relevant to review
								Logistic regression	Fasting plasma glucose, 5.6mmol/L: AUROC[i] 0.723 (0.653-0.793, CI[l]) Specificity 83.7%, Sensitivity 56.2%, PPV[j] 39.8%, NPV[k] 90.9%. 1hr glucose, 10.7 mmol/L: AUROC[i] 0.655 (0.583-0.727),95% CI[l])		

	Author	Country	Study design	Type of prediction study	GDM[a] diagnostic criteria	Data collection period (years)	Number of participants	Algorithm used in each model described in the paper	The performance of the model	Aims relevant to review	Outcomes relevant to review
									Specificity 73.5%, Sensitivity 49.3%, PPV[j] 26.3%, NPV[k] 88.3%. HbA1c[o], 5.4%: AUROC[i] 0.734 (0.672-0.796, 95% CI[l]) Specificity 63.8%, Sensitivity 69.9%, PPV[j] 27%, NPV[k] 91.7%. Abdominal circumference 69th percentile:		

	Author	Country	Study design	Type of prediction study	GDM[a] diagnostic criteria	Data collection period (years)	Number of participants	Algorithm used in each model described in the paper	The performance of the model	Aims relevant to review	Outcomes relevant to review
									AUROC[i] 0.662 (0.591-0.733, 95% CI[l]) Specificity 60.3%, Sensitivity 69%, PPV[j] 26%, NPV[k] 89.8%.		
	Zaccara et al. 2023 (Zaccara et al., 2023)	Brazil	Retrospective cohort study	Development	ADA[d]	8	869	Logistic regression	AUROC[i] 0.77 (0.72-0.81, 95% CI[l])	To identify risk factors associated with insulin need in women with GDM[a].	Regularly collected data from patients can calculate the risk of a woman with GDM[a] diagnosed in OGTT[m] needing insulin.

[a]*GDM Gestational diabetes mellitus*
[b]*ADIPS Australasian Diabetes in Pregnancy Society*
[c]*IADPSG International Association of the Diabetes and Pregnancy Study Group*
[d]*ADA American Diabetes Association*

[e]*CART Classification and regression tree*
[f]*LASSO Least absolute shrinkage and selection operator*
[g]*Simple super learner could have been included response-mean, least absolute shrinkage and selection operator regression, and Classification and regression tree*
[h]*Complex super learner could have been response-mean, least absolute shrinkage and selection operator regression, Classification and regression tree, random forest, and extreme gradient boosting*
[i]*AUROC Area under the receiver operating characteristics*
[j]*PPV Positive predictive value*
[k]*NPV Negative predictive value*
[l]*CI Confidence interval*
[m]*OGTT Oral Glucose tolerance test*
[n]*BMI Body mass index*
[o]*HbA1c glycohemoglobin, haemoglobin*

SUPPLEMENTARY MATERIAL 3: PERFORMANCE METRICS

Whole model algorithm performance, median (range)

	AUROC[a]	Sensitivity, %	Specificity, %	PPV[b]	NPV[c]
Any algorithm	0.75 (0.61-0.93)	59.6 (9.3-90.0)	81.2 (36.0-91.5)	67.5 (17.6-92.5)	70.0 (58.0-74.0)
Logistic regression	0.75 (0.63-0.87)	59.6 (9.3-90.0)	89.9 (36-99.4)	67.54 (58.0-74.0)	70.0 (58.0-74.0)
CART[d]	0.74 (0.61-0.79)	-	-	-	-
LASSO[e]	0.74 (0.67-0.84)	-	-	-	-
Simple super learner[f]	0.74 (0.67-0.85)	-	-	-	-
Complex super learner[g]	0.82 (0.64-0.93)	-	-	-	-

[a]*AUROC Area under the receiver operating curve*
[b]*PPV Positive predictive value*
[c]*NPV Negative predictive value*
[d]*CART Classification and regression trees*
[e]*LASSO Least absolute shrinkage and selection operator*
[f]*Simple super learner could have been included response-mean, least absolute shrinkage and selection operator regression, and Classification and regression tree*
[g]*Complex super learner could have been response-mean, least absolute shrinkage and selection operator regression, Classification and regression tree, random forest, and extreme gradient boosting*

Whole model algorithm performance predicting insulin, median (range)

	AUROC[a]	Sensitivity, %	Specificity, %	PPV[b]	NPV[c]
Any algorithm	0.74 (0.71 – 0.87)	54.3 (9.3 - 90/0)	90.65 (36.0 – 99.4)	70/0 917.6-87.6)	70.0 (69.9-74.0)
Logistic regression	0.76 (0.71-0.87)	54.3 (9.3-90.0)	90.7 (36.0-99.4)	70.0 (17.6 -87.6)	70 (69.9-74)
CART[d]	0.74*	-	-	-	-

[a]*AUROC Area under the receiver operating curve*
[b]*PPV Positive predictive value*
[c]*NPV Negative predictive value*
[d]*CART Classification and regression trees*
** only appeared in one model*

Whole model algorithm performance predicting pharmacological therapy, median (range)

	AUROC[a]	Sensitivity, %	Specificity, %	PPV[b]	NPV[c]
Any algorithm	0.75 (0.61 – 0.93)	72.2*	86.8*	92.5*	58*
Logistic regression	0.75 (0.63-0.83)	72.2*	86.8*	92.5*	58.0*
CART[d]	0.68 (0.61-0.79)	-	-	-	-
LASSO[e]	0.74 (0.67-0.85)	-	-	-	-
Simple super learner[f]	0.74 (0.67-0.85)	-	-	-	-
Complex super learner[g]	0.82 (0.64-0.94)	-	-	-	

[a]*AUROC Area under the receiver operating curve*
[b]*PPV Positive predictive value*
[c]*NPV Negative predictive value*
[d]*CART Classification and regression trees*
[e]*LASSO Least absolute shrinkage and selection operator*
[f]*Simple super learner could have been included response-mean, least absolute shrinkage and selection operator regression, and Classification and regression tree*
[g]*Complex super learner could have been response-mean, least absolute shrinkage and selection operator regression, Classification and regression tree, random forest, and extreme gradient boosting*
**only appeared in one model*

Individual model algorithm performance, median (range) [All predicting insulin and used logistic regression]

	AUROC[a]	Sensitivity, %	Specificity, %	PPV[b]	NPV[c]
HbA1c	0.63 (0.58-0.73)	62.4 (29.7 - 69.9)	62.0 (54.7-87.7)	39.5 (13.8 - 59.4)	69.4 (61.7-92.8)
Fasting 75g OGTT	0.64 (0.52-0.72)	42.5 (20.1-56.2)	84.4 (83.7-92.0)	58.3 (39.8-63.9)	82.3 (69.3 - 90.9)
1hr 75G OGTT	0.65 (0.62-0.88)	50.2 (29.7-100)	74.1 (66.7-77.8)	44.0 (15.9 - 61.1)	86.3 (68.3-92.9)
2hr 75g OGTT	0.75 (0.75 - 0.76)	74.2 (70. - 78.4)	67.6 (64.7-70.4)	55.3*	84.3 (84.3-84.3)
Number of abnormal values in 75g OGTT *	0.78	80.0	74.1	-	-
Pre-pregnancy BMI*	0.62	63.3	57.8	55.4	73.8
Abdominal circumference percentile*	0.66	60.3	67.2	26.0	89.3

[a]*AUROC Area under the receiver operating curve*
[b]*PPV Positive predictive value*
[c]*NPV Negative predictive value*
**only appeared in one model*

Whole model performance within class imbalance, median (range)

Predictive group, %	AUROC[a]
10-20	0.76 (0.72-0.87)
21-30	0.78 (0.71-0.86)
31-40	0.76 (0.61-0.93)
41-50	0.72 (0.71-0.78)
>51	0.77 (0.70-0.83)

[a]*AUROC Area under the receiver operating curve*

Whole model performance within GDM diagnostic criteria, median (range)

	AUROC[a]	Sensitivity, %	Specificity, %	PPV[b]	NPV[c]
ADA[d*]	0.77	-	-	-	-
ADIPS[e]	0.72 (0.71 – 0.74)	-	-	-	-
Carpenter and Coustan criteria	0.77 (0.61 – 0.93)	36.7 (30.8 – 98.8)	97.5 (86.6 – 92.5)	92.5*	58*
IADSPG[f]	0.74 (0.71 – 0.78)	67.3 (64.0 - 90.0)	67.3 (36.0 - 87.0)	70.0 (66.0 – 74.0)	72.0 (70.0 - 74.0)
IADSPG[f] and national guidelines*	0.80	-	-	-	-

[a]AUROC Area under the receiver operating curve
[b]PPV Positive predictive value
[c]NPV
[d]ADA American Diabetes Association
[e]ADIPS Australasian Diabetes in Pregnancy Society
[f]IADPSG International Association of the Diabetes and Pregnancy Study Group
* only appeared in one model

SUPPLEMENTARY MATERIAL 4: RISK OF BIAS AND APPLICABILITY ASSESSMENT

PROBAST (Prediction model Risk Of Bias ASsessment Tool) risk of bias and applicability assessment

	Author	Model	Risk of bias				Applicability			Overall	
		Model description	Participants	Predictors	Outcome	Analysis	Participants	Predictors	Outcome	Risk of bias	Applicability
Predicting pharmacological therapy	Feghali et al. 2019 (Feghali et al., 2019)	Without SMBG[a]	-	+	+	-	+	+	+	-	?
		With SMBG[a]	-	+	+	-	+	+	+	-	?
	Liao et al. 2022 (Liao et al., 2022)	CART[b] L1[c]	-	+	+	+	-	+	?	-	-
		CART[b] L1-2[d]	-	+	+	+	-	+	?	-	-
		CART[b] L1-3[e]	-	+	+	+	-	+	+	-	-
		CART[b] L1-4[f]	-	+	+	+	-	+	+	-	-
		LASSO[g] L1[c]	-	+	+	+	-	+	?	-	-
		LASSO[g] L1-2[d]	-	+	+	+	-	+	?	-	-
		LASSO[g] L1-3[e]	-	+	+	+	-	+	+	-	-
		LASSO[g] L1-4[f]	-	+	+	+	-	+	+	-	-
		SL[h] L1[c]	-	+	+	+	-	+	?	-	-
		SL[h] L1-2[d]	-	+	+	+	-	+	?	-	-
		SL[h] L1-3[e]	-	+	+	+	-	+	?	-	-
		SL[h] L1-4[f]	-	+	+	+	-	+	+	-	-

	Author	Model	Risk of bias				Applicability			Overall	
		Model description	Participants	Predictors	Outcome	Analysis	Participants	Predictors	Outcome	Risk of bias	Applicability
		CL[i] L1[c]	-	+	+	+	-	+	?	-	-
		CL[i] L1-2[d]	-	+	+	+	-	+	?	-	-
		CL[i] L1-3[e]	-	+	+	+	-	+	?	-	-
		CL[i] L1-4[f]	-	+	+	+	-	+	+	-	-
		LR[j] L1[c]	-	+	+	+	-	+	?	-	-
		LR[j] L1-2[d]	-	+	+	+	-	+	?	-	-
		LR[j] L1-3[e]	-	+	+	+	-	+	+	-	-
		LR[j] L1-4[f]	-	+	+	+	-	+	+	-	-
	Velardo et al. 2021 (Velardo et al., 2021)		+	-	+	+	-	-	+	-	-
	Yerlikaya et al. 2018 (Yerlikaya et al., 2018)	Using OGTT[k]	+	+	+	-	+	+	+	?	+
		Using clinical variables	+	+	+	-	+	+	+	?	+
		Using OGTT[k]and clinical variables	+	+	+	-	+	+	+	?	+
		Random forest	+	+	+	-	+	+	+	-	?

	Author	Model	Risk of bias				Applicability			Overall	
		Model description	Participants	Predictors	Outcome	Analysis	Participants	Predictors	Outcome	Risk of bias	Applicability
Predicting insulin	Barnes et al. 2016 (Barnes et al., 2016)		+	+	+	?	+	+	?	-	?
	Ducarme et al. 2019 (Ducarme et al., 2019)		-	+	+	?	-	+	+	-	-
	Eleftheriades et al. 2021 (Eleftheriades et al., 2021)		-	+	+	-	+	-	+	-	?
	Ford et al. 2022 (Ford et al., 2022)		+	+	+	+	+	+	+	+	?
	Harper et al. 2016 (Harper et al., 2016)	Including pre-diabetes	-	+	+	-	+	+	-	-	-
		Excluding pre-diabetes	-	+	+	?	+	+	-	-	-
	Khin et al. 2018		?	+	+	-	+	+	+	-	-

	Author	Model	Risk of bias				Applicability			Overall	
		Model description	Participants	Predictors	Outcome	Analysis	Participants	Predictors	Outcome	Risk of bias	Applicability
	(Khin et al., 2018)										
	Nishikawa et al. 2018 (Nishikawa et al., 2018)		+	+	+	?	+	+	+	-	-
	Souza et al. 2019 (Souza et al., 2019)		+	+	+	?	+	+	+	?	+
	Tamagawa et al. 2021 (Tamagawa et al., 2021)		-	+	+	-	+	+	+	-	+
	Tang et al. 2019 (Tang et al., 2019)		?	+	+	-	+	+	+	-	+
	Watanabe et al. 2016 (Watanabe et al., 2016)		?	+	+	-	+	+	+	-	?

	Author	Model	Risk of bias				Applicability			Overall	
		Model description	Participants	Predictors	Outcome	Analysis	Participants	Predictors	Outcome	Risk of bias	Applicability
	Weschenfelder et al. 2021 (Weschenfelder et al., 2021)	Predicting insulin	+	+	+	-	+	+	+	-	?
		Predicting Bolus insulin	+	+	+	-	+	+	+	-	?
		Predicting basal insulin	+	+	+	-	+	+	+	-	?
		Predicting multiple injections	+	+	+	-	+	+	+	-	?
	Zaccara et al. 2023 (Zaccara et al., 2023)		+	+	+	-	+	+	+	-	?

+: Low risk of bias/concern about applicability, -: high risk of bias/concern about applicability, ?: unclear risk of bias/concern about applicability

[a]*SMBG Self monitoring blood glucose*
[b]*CART Classification and Regression Tree*
[c]*L1 Data from 1-year preconception to last menstrual period*
[d]*L1-2 Data from 1-year preconception to last menstrual period until last menstrual period to before diagnosis of GDM*
[e]*L1-3 Data from 1-year preconception until the time of diagnosis of GDM*
[f]*L1-4 L1 Data from 1-year preconception to last menstrual period until 1 week after diagnosis of GDM (including self-monitoring blood glucose data)*
[g]*LASSO Least absolute shrinkage and selection operator*
[h]*SL Simple super learner could have been included response-mean, least absolute shrinkage and selection operator regression, and Classification and regression tree*

[i]*CL Complex super learner could have been response-mean, least absolute shrinkage and selection operator regression, Classification and regression tree, random forest, and extreme gradient boosting*
[j]*LR Logistic regression*
[k]*OGTT Oral glucose tolerance test*

SUPPLEMENTARY MATERIAL 5: SENSITIVITY ANALYSIS

Liao et al.(Liao et al., 2022), presented 20 studies, (45.5%) of the included models, therefore a sensitivity analysis was conducted.

Topic	Before (including Liao et al. (2022))	After sensitivity analysis (excluding Liao et al. (2022))
Number of included models	44	24
Prediction groups	Predicting insulin: 38.6%, 17/44 Predicting pharmacological therapy: 61.4%, 27/44	Predicting insulin: 70.8%, 17/24 Predicting pharmacological therapy: 29.2%, 7/24
Study period (median (range))	5 years (range 1-23 years)	5 years (range 1-23 years)
Number of participants (median (range))	1919 participants (range 37 - 30,474)	304 participants (range 37-2217)
Percentage of participants in control or predictive group (median (range))	Control group: 61.2% (range 30.2-89.2%) Predictive group: 38.8% (range 10.8-69.8%)	Control group: 65.1% (range 30.2-89.2) Predictive group: 35.0% (range 10.8-69.8%)
Algorithms used (percentage, frequency)	Logistic regression (59.1%, 26/44), CART[a] (11.4%, 5/44), LASSO[b] (9.1%, 4/44), Simple super learner (either response-mean, LASSO[b], and CART[a]) (9.1%, 4/44), Complex super learner (either response-mean, LASSO[b], CART[a], random forest, or extreme gradient boosting), (9.1%, 4/44)	Logistic regression (97.7%, 22/24), CART[a] (4.2%, 1/24), Random forest (4.2%, 1/24)
Frequently used variables overall (percentage, frequency)	History of GDM[c] (47.7%, 21/44), Gestational week at GDM[c] diagnosis (45.5%, 20/44), Pregestational BMI[d] (40.9%, 18/44), and Maternal age (38.6%, 17/44).	1hr 75g OGTT[e] (58.3% 14/24), Fasting 75g OGTT[e] (54.2%, 13/24), Maternal age (41.7%, 10/24) Gestational week at GDM[c] diagnosis (41.7%, 10/24)
Frequently used variables predicting pharmacological therapy (percentage, frequency)	History of GDM[c] (51.9%, 14/27), Gestational week at GDM[c] Diagnosis (51.9%, 14/27), Pregestational BMI[d] (48.1%, 13/27), and Maternal age (41.8%, 13/27)	Maternal age (85.7%, 6/7), Gestational week at GDM[c] diagnosis (57.1%, 4/7), Pregestational BMI[d] (57.1%, 4/7) and Parity (57.1%, 4/7)
Performance metrics used	AUROC[f] (95.5%, 42/44),	AUROC[f] (91.7%, 22/24),

(percentage, frequency)	Sensitivity and specificity (36.4%, 16/44), PPV[g] and NPV[h] (25.0%, 11/44)	Sensitivity and specificity: (66.7%, 16/24), PPV[g] and NPV[h] (45.8% ,11/24)
Overall AUROC[f] performance (median (range))	0.75 (range 0.61-0.93)	0.74 (range 0.70-0.87)
Logistic regression AUROC[f] performance (median (range))	0.76 (range 0.63-0.87)	0.76 (range 0.70-0.87)
Percentage of models validated, (frequency)	65.9% (29/44)	37.5% (9/24)
Percentage of overall risk of bias using PROBAST[i] (Wolff et al., 2019), (frequency)	High: 88.6%, (39/44), Unclear: 9.1%, (4/44), Low: 2.3%, (1/44)	High: 79.2%, (19/24) Unclear: 16.7%, (4/24), Low: 4.2%, (1/24)
Percentage of overall concern about applicability using PROBAST[i] (Wolff et al., 2019), (frequency)	High: 59.1%, (26/44), Unclear: 27.3% (12/44), Low: 13.6% (6/44)	High: 25.0%, (6/24), Unclear: 50.0%, (12/24), Low: 25.0%, (6/24)

[a]*CART Classification and regression trees*
[b]*LASSO Least absolute shrinkage and selection operator*
[c]*GDM Gestational diabetes*
[d]*BMI Body Mass Index*
[e]*OGTT Oral Glucose Tolerance Test*
[f]*AUROC Area under the receiver operating curve*
[g]*PPV Positive predictive value*
[h]*NPV Negative predictive value*
[i]*PROBAST Prediction model Risk Of Bias ASsessment Tool*

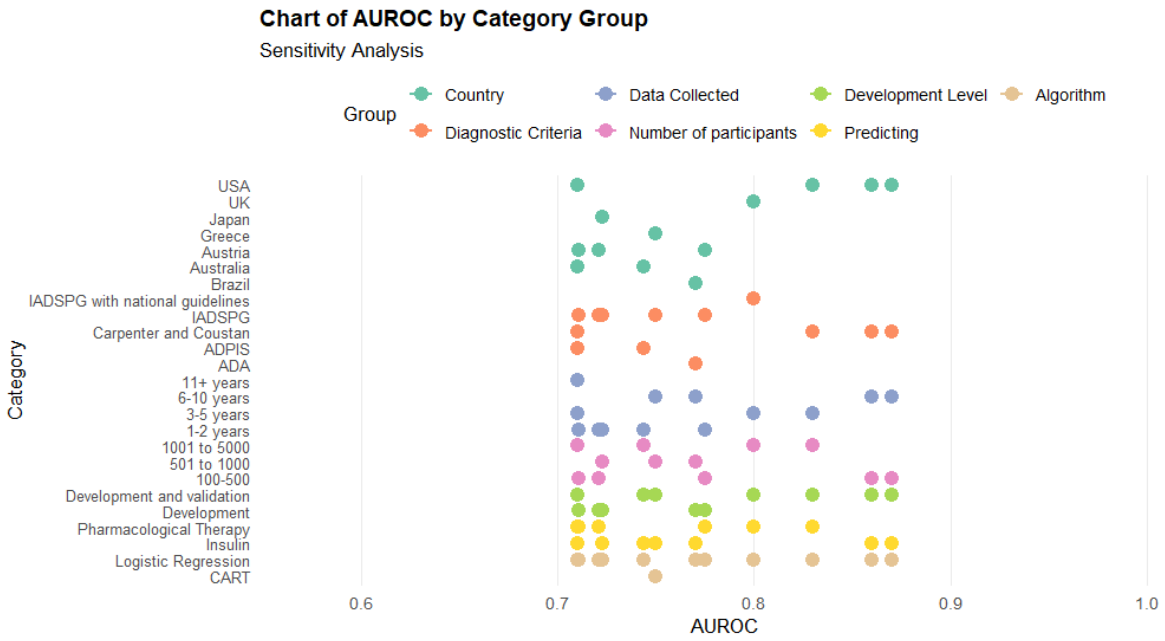


Figure *Sensitivity analysis results of AUROC for the whole model performance with the removal of Liao et al. (2022) removed, for categories grouped by country, GDM diagnosis criteria, length of data collection and number of participants within the study, development level of the model, the prediction of the model and the algorithm used.*

AUROC Area Under the Receiver Operating Characteristics, USA – United States of America, UK- United Kingdon, IADSPG – International Association of the Diabetes and Pregnancy Study Group, ADIPS - Australasian Diabetes in Pregnancy Society, ADA - American Diabetes Association, CART Classification and regression Tree, LASSO Least absolute shrinkage and selection operator, Simple super learner could have been included response-mean, least absolute shrinkage and selection operator regression, and classification and regression tree, Complex super learner could have been response-mean, least absolute shrinkage and selection operator regression, Classification and regression tree, random forest, and extreme gradient boosting.

Appendix C *The Use of Machine Learning to Predict Pharmacological Therapy in Gestational Diabetes: A Scoping Review Protocol [Registered on the Open Science Framework (OSF)]*

The use of machine learning to predict pharmacological therapy in gestational diabetes: a scoping review protocol

Authors

Jasmine R Kirkwood[1] Natasha Galloway[2] Robert S Lindsay[3], Areti Manataki[4], Deborah Wake[5], Rebecca M Reynolds[1]
1. Centre for Cardiovascular Science, Queen's Medical Research Institute, University of Edinburgh
2. Edinburgh Centre for Endocrinology and Diabetes, NHS Lothian Hospitals Trust
3. School of Cardiovascular and Metabolic Health, University of Glasgow
4. School of Computer Science, University of St Andrews
5. Usher Institute, University of Edinburgh

Abstract

Objective: This scoping review aims to comprehensively review the methods and outcomes in machine learning used to predict need for pharmacological therapy in gestational diabetes, whilst identifying the research gaps.

Introduction: Successful management of blood glucose during gestational diabetes reduces the risk of morbidities for both women and their child. Initial treatment is through diet and lifestyle changes, and where this treatment is not sufficient results in the need for pharmacological therapy. Early identification could allow for better therapeutic strategies, rapid decision making, and allocation of resources and reduce associated morbidities.

Inclusion criteria: Studies that have/are the following will be included in this scoping review: 1) studies predicting pharmacological therapy for gestational diabetes using machine learning; 2) studies written in English (authors not familiar with any other language); 3) studies published from 1st July 2007 onwards. Studies that have/are the following will be excluded from this scoping review: 1) studies predicting pharmacological therapy for gestational diabetes that do not use machine learning; 2) conference abstracts.

Methods: The scoping review will search electronic databased for published literature in English, including Embase, Medline, IEEE Xplore, and Web of Science from 1st July 2007 onwards using the following keyword, machine learning, prediction, insulin, metformin, and gestational diabetes. It will be conducted using the framework set out in the Joanna Briggs institute [1], whist following the Preferred reporting items for systematic reviews and meta-analyses – scoping review (PRISMA-ScR) checklist [2].

Introduction

Gestational diabetes is hyperglycaemia that is first onsets or recognised during pregnancy [3-5]. It has a high prevalence, affecting approximately 1-3% of births in the UK [6], and 13.4% of live births worldwide in 2019 [5]. This is expected to increase due to women having pregnancies at a later age [6] and the rise in obesity among women of reproductive age [6, 7].

Gestational diabetes causes short-term complications with the pregnancy and can also have long-term health complications for both mother and baby. During pregnancy there is an increased risk of having preeclampsia [5], a large for gestational age baby, which increases the potential of birth trauma, induction of labour, instrumental or Caesarean section births [4] and pre-term births [8]. The child has a greater risk of future health problems later in life such as obesity or diabetes [4, 5, 8]. After birth, women with gestational diabetes are ten times more likely to develop type 2 diabetes [9], and they are more likely to have gestational diabetes during subsequent pregnancies [8].

After diagnosis, the woman is educated about how to closely monitor her blood glucose using home blood glucose monitoring, make modifications to her diet and exercise and is invited to attend additional multi-disciplinary specialised clinics, typically fortnightly, to discuss the home blood glucose monitoring results and adjust the treatment plan accordingly [4, 10]. Some women can manage their gestational diabetes through diet and exercise alone, whereas others will have to take additional medication, in the form of an oral tablet metformin, and/or insulin injections [4].

From diagnosis to birth there are 9-13 weeks during which a woman with gestational diabetes must learn and understand the consequences of gestational diabetes whilst also managing the extra responsibility of monitoring blood glucose levels, changes in lifestyle and attending extra hospital appointments.

Achieving good glycaemic control prevents adverse pregnancy outcomes for both mother and child [11]; it is thought that identifying women who are less likely to achieve optimum glucose levels through diet-only method, reduces adverse pregnancy outcomes [12]. In addition, early identification could allow for better therapeutic strategies, rapid decision making and allocation of resources.

Alvarez-Silvares et al [13] presented a systematic review and meta-analysis to identify risk factors that may indicate the need for insulin therapy upon the diagnosis of GDM. The authors reviewed 18 observational studies from 1975 to 2021; these studies did not exclusively use machine learning. It was found that there was a strong relationship between having a BMI ≥ 30 kg/m^2, a family history of type 2 diabetes, prior history of GDM and higher glycosylated haemoglobin at diagnosis, with the need for insulin treatment in GDM.

Furthermore, a primarily search on PubMed, Web of Science, ScienceDirect, the Cochrane Database of Systematic Reviews, IEEE Xplore and JBI Evidence Synthesis found there has been no scoping reviews conducted on machine learning use in predicting pharmacological therapy in gestational diabetes. Therefore, there is an opportunity to conduct such a review and identify what methods and outcomes of the machine learning models have been used and how well they perform.

Review question

This scoping review aims to comprehensively review the methods and outcomes in machine learning used to predict the need of prescribing or the escalation of pharmacological therapy in gestational diabetes both at the start of or continuously throughout pregnancy, whilst identity the research gaps. The scoping review will facilitate the direction of future research. Furthermore, the results of the review could aid clinics to risk-stratify patients and identify

those that are a higher risk of need pharmacological interventions and therefore allocate resources appropriately.

The primary research question is:

1. What machine learning techniques, methods, models, and algorithms have been used in predicting need for pharmacological therapy in gestational diabetes?

The secondary research questions are:

1. What features were included in the machine learning models to the need for predict pharmacological therapy in gestational diabetes?
2. What performing metrics have been reported and how well have the machine learning models worked for predicting need for pharmacological therapy in gestational diabetes?
3. How were machine learning models for predicting need for pharmacological therapy in gestational diabetes validated?

Keywords

Gestational diabetes; machine learning; medication; pharmacological therapy, prediction

Eligibility criteria

Participants

Participants that will be included are women that have been diagnosed with gestational diabetes via any method described by the author.

Concept

Participants diagnosed with type 1 or type 2 diabetes prior to or during their pregnancy will be excluded. Only models that use common supervised and unsupervised machine learning techniques will be included in this study as we are interested in how machine learning has been used and how well it has performed. Therefore, models that only used statistical powers to make predictions will be excluded from this review.

Context

The review will be as broad as possible and will try to capture as much detail about the population and settings of the use of machine learning to predict need for pharmacological therapy in gestational diabetes.

Types of Sources

This scoping review will consider both experimental and quasi-experimental study designs including randomized controlled trials, non-randomized controlled trials, before and after studies and interrupted time-series studies. In addition, analytical observational studies including prospective and retrospective cohort studies and analytical cross-sectional studies will be considered for inclusion. This review will also consider descriptive observational study designs including and descriptive cross-sectional studies for inclusion.

Methods

The scoping review will be conducted using the framework set out in the Joanna Briggs institute [1], whist following the Preferred reporting items for systematic reviews and meta-analyses – scoping review (PRISMA-ScR) checklist [2].

Search strategy

The review will search electronic databases for published literature, including Embase, Medline, IEEE Xplore, and Web of Science, the search strategy for each database is in the Appendix. A three-step search strategy as detailed in Joanna Briggs institute [1] will be used to comprehensively review the literature. Firstly, an initial search on Embase and Medline will be conducted the text words in the titles and abstracts of the papers selected will be analysed. A second search using all keywords and index terms over all databased will be completed. Finally, the reference list of identified literature will be used to identify any further literature.

Study/Source of Evidence selection

The literature found will be uploaded to EndNote (https://endnote.com/) and imported in to Covidence (https://www.covidence.org/), and duplications will be removed. The first reviewer (JK) will review and identify the literature through the title and abstract first based on the inclusion criteria. Following this a full text screening will be conducted, where the inclusion criteria can be refined. The second reviewer (NG) will review 10% of the literature that has met the inclusion criteria. Disagreements on literature inclusion will be discussed and agreed upon if no agreement is made a third reviewer (DW) will be consulted.

The results of the literature search will be reported in full and present in a flow diagram in keeping with the PRISMA-ScR checklist [2].

Data Extraction

The first reviewer (JK) will extract the data, summarised in table 1, from the literature that has met the inclusion criteria. Initial pilot of three articles will be conducted which will be summarised on the following author, publication year, country of research institution, location of data extract, time range of data used, study type and aims. Study features that will be recorded are, diagnosis criteria used, what is being predicted, characteristics used in the models, and the characteristics that are reported either have a significant or no significant effect on medication need, what machine learning model was use, how it performed, and how it was validated. If after the pilot the research questions are not being met by the data abstracted, then amendments will be made to the data form; following which the full text reviewing process will begin. Any uncertainties will be resolved by consulting experts, (RR, RL, AM).

Table 1 Pilot data form for data abstraction

	Study 1	Study 2	etc
Author			
Publication year			
Country of research institute			
Location of participants/data			
Time range of data			

Gestational diabetes diagnosis guideline/criteria			
Study aim			
Study type			
Sample number			
Therapy predicted (e.g., insulin, metformin, both)			
Characteristic used in model			
Characteristic reported to be significant in effecting medication need			
Characteristic reported to not be significant in effecting medication need			
Machine learning model used			
Performance/validation metric used			
Performance/validation of model			
Clinical validation			
External validation			

Data Analysis and Presentation

The results of the data abstraction and charting will be analysis in relation to the research questions. Comparisons between the studies will be made regarding the machine learning models and characteristics used. Summarisation of the results and quantification will be done where possible will be presented in accordance with the PRIMSA-ScR checklist [2].

Acknowledgements

The authors would like to thank Marshal Dozier and Bohee Lee for assistance is developing the search terms.

Funding

The review will be conducted as part of a PhD that is funded by Medical Research Scotland (PHD-50224-2020).

Conflicts of interest

There are no conflicts of interest.

References

1. Peters, M.D.J., et al., *Updated methodological guidance for the conduct of scoping reviews.* JBI Evidence Synthesis, 2020. **18**(10).
2. Tricco, A.C., et al., *PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation.* Annals of Internal Medicine, 2018. **169**(7): p. 467-473.
3. Diabetes, W.H.O.W.I. and Federation (IDF). *Diabetes key facts,*. 2021, 13 April; Available from: https://www.who.int/news-room/fact-sheets/detail/diabetes.

4. National Institute for Health and Care Excellence (NICE). *Diabetes in pregnancy: management from preconception to the postnatal period,*. 2015, 25 February 2020, 16 December [cited 2021 22 October]; Available from: https://www.nice.org.uk/guidance/ng3/resources/diabetes-in-pregnancy-management-from-preconception-to-the-postnatal-period-pdf-51038446021.
5. *International Diabetes Federation, IDF Diabetes Atlas Tenth Edition 2021*. 2021 [cited 2022 05 July]; 10:[Available from: https://diabetesatlas.org/idfawp/resource-files/2021/07/IDF_Atlas_10th_Edition_2021.pdf.
6. Johns, E.C., et al., *Gestational Diabetes Mellitus: Mechanisms, Treatment, and Complications.* Trends in Endocrinology & Metabolism, 2018. **29**(11): p. 743-754.
7. Ferrara, A., *Increasing Prevalence of Gestational Diabetes Mellitus.* Diabetes Care, 2007. **30**(Supplement_2): p. S141-S146.
8. Buchanan, T.A., A.H. Xiang, and K.A. Page, *Gestational diabetes mellitus: risks and management during and after pregnancy.* Nature Reviews Endocrinology, 2012. **8**(11): p. 639-649.
9. Vounzoulaki, E., et al., *Progression to type 2 diabetes in women with a known history of gestational diabetes: systematic review and meta-analysis.* BMJ, 2020: p. m1361.
10. Mackillop, L., et al., *Development of a Real-Time Smartphone Solution for the Management of Women With or at High Risk of Gestational Diabetes.* Journal of diabetes science and technology, 2014. **8**(6): p. 1105-1114.
11. Hashimoto, K., *Indicators of glycemic control in patients with gestational diabetes mellitus and pregnant women with diabetes mellitus.* World Journal of Diabetes, 2015. **6**(8): p. 1045.
12. Meshel, S., et al., *Can we predict the need for pharmacological treatment according to demographic and clinical characteristics in gestational diabetes?* The Journal of Maternal-Fetal & Neonatal Medicine, 2016. **29**(13): p. 2062-2066.
13. Alvarez-Silvares, E., et al., *Prediction of insulin therapy in women with gestational diabetes: a systematic review and meta-analysis of observational studies.* Journal of Perinatal Medicine, 2022.

Appendix

Search strategy

Database	Keywords
Medline	((machine or *supervised) adj2 learning [TEXT] OR ((artificial or computer or machine) adj2 intelligence) [TEXT] OR ((logistic or linear) adj1 regression) [TEXT] OR Regression[TEXT] OR classification [TEXT] OR ai [TEXT] OR ml [TEXT] OR (decision adj1 (forest* OR tree*) [TEXT] OR Artificial intelligence [MESH] OR Exp machine learning [MESH] OR Exp algorithms [MESH] OR Exp sensitivity and specificity [MESH] OR Exp regression analysis [MESH]) AND (prediction adj3 (metformin OR insulin OR drug OR pharmacological OR model OR diet OR glibenclamide OR glyburide)

	[TEXT] OR ((insulin OR metformin OR glibenclamide OR glyburide) adj2 (failure OR pregan* OR need)) [TEXT] OR Medication or Drug or Pharmacolog* [TEXT] OR drug therapy/ [MESH] OR glycemic control [MESH] OR exp Metformin/[MESH] OR Insulin/[MESH] OR Glyburide/[MESH] OR exp Diet Therapy[MESH] OR Pharmacology/[MESH] OR Medication Therapy Management [MESH] OR Exp Blood glucose [MESH]) AND ((gestation* OR pregan* OR maternal) adj3 (diabete* OR hyperglyc?emia OR glucose) [TEXT] OR Fasting blood glucose* [TEXT] OR fgd [TEXT] OR bg [TEXT] OR gdm [TEXT] OR diabetes, gestational/ [MESH] OR pregnancy in diabetics/ [MESH]) AND ((2007* or 2008* or 2009* or 2010* or 2011* or 2012* or 2013* or 2014* or 2015* or 2016* or 2017* or 2018* or 2019* or 2020* or 2021* or 2022*).ed)
Embase	((machine or *supervised) adj2 learning [TEXT] OR ((artificial or computer or machine) adj2 intelligence) [TEXT] OR ((logistic or linear) adj1 regression) [TEXT] OR Regression[TEXT] OR classification [TEXT] OR ai [TEXT] OR ml [TEXT] OR (decision adj1 (forest* OR tree*) [TEXT] OR exp artificial intelligence/ OR exp machine learning/ OR exp algorithm/ OR exp regression analysis/) AND (prediction adj3 (metformin OR insulin OR glibenclamide OR glyburide OR drug OR pharmacological OR model or diet) [TEXT] OR ((insulin OR metformin OR glibenclamide OR glyburide) adj2 (failure OR pregan* OR need)) [TEXT] OR

	Medication or Drug or Pharmacolog* [TEXT] OR Drug therapy/ OR Metformin/ OR Insulin/ OR Glibenclamide OR Exp diet therapy/ OR Pharmacology/ OR Exp Medication Therapy Management/ OR exp glucose blood level/ OR exp blood level/ OR exp glucose level/) AND ((gestation* OR pregan* OR maternal) adj3 (diabete* OR hyperglyc?emia OR glucose) [TEXT] OR Fasting blood glucose* [TEXT] OR fgd [TEXT] OR bg [TEXT] OR gdm [TEXT] OR exp pregnancy diabetes mellitus/ OR pregnancy complication/ OR "pregnancy disorders of endocrine origin"/ OR maternal diabetes mellitus/) AND (limit to yr"2007-current")
IEEE Xplore	((machine OR *supervised) NEAR/2 learning OR ((artificial OR computer OR machine) NEAR/2 intelligence) OR ((logistic OR linear) NEAR/1 regression) OR Regression OR classification OR ai OR ml OR (decision NEAR/1 (forest* OR tree*)) OR Artificial intelligence [MESH] OR Exp machine learning [MESH] OR Exp algorithms [MESH] OR Exp sensitivity and specificity [MESH] OR Exp regression analysis [MESH]) AND (prediction NEAR/3 (metformin OR insulin OR glibenclamide OR glyburide OR drug OR pharmacological OR model OR diet) OR ((insulin OR metformin OR glibenclamide OR glyburide) NEAR/2 (failure OR pregan*OR need)) OR Medication OR Drug OR Pharmacolog* OR drug therapy/ [MESH] OR

	glycemic control [MESH] OR exp Metformin/[MESH] OR Insulin/[MESH] OR Glyburide/[MESH] OR exp Diet Therapy[MESH] OR Pharmacology/[MESH] OR Medication Therapy Management [MESH] OR Exp Blood glucose [MESH]) AND (gestation* OR pregan* OR maternal) NEAR/3 (diabete* OR hyperglyc?emia OR glucose) Fasting blood glucose* OR fgd OR bg OR gdm OR diabetes, gestational/ [MESH] OR pregnancy in diabetics/ [MESH])
Web of Science	(((TS=((machine OR supervised OR unsupervised) NEAR/2 learning) OR TS=(((artificial OR computer OR machine) near/2 intelligence))) OR TS=(((logistic OR linear) near/1 regression))) OR TS=(Regression OR classification OR ai OR ml)) OR TS=((decision near/1 (forest* OR tree*))) AND (TS=(prediction near/3 (metformin OR insulin OR glibenclamide OR glyburide OR drug OR pharmacological OR model OR diet)) or TS=(((insulin OR metformin OR glibenclamide OR glyburide) near (failure))) OR TS=(((insulin or metformin) near (pregan*))) OR TS=(((insulin or metformin OR glibenclamide OR glyburide) near (need))) OR TS=(Medication OR Drug OR Pharmacolog*)) AND (TS=((gestation* OR pregan* OR maternal) NEAR/3 (diabete* OR hyperglyc?emia OR glucose)) OR TS=(Fasting blood glucose* OR fgd OR bg OR gdm))

Appendix D *Ethical approval letter: Gestational Diabetes Research Database (GSH12DI001)*

NHS Greater Glasgow and Clyde
West of Scotland Safe Haven
2F ICE Building, QEUH, Glasgow

Thursday, 12 October 2023

To Whom It May Concern,

Safe Haven Application Feedback

Project ID: GSH23DI001

Project Name: Gestational Diabetes Research Database

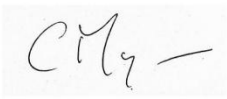
All applications to the Safe Haven are reviewed separately by the Safe Haven team, the R&D Peer Review Committee (if appropriate), the Local Privacy Advisory Committee group and measured against a set of pre-defined criteria.

Your application has been through all relevant review processes and found to be acceptable as appropriate research for involvement with the Safe Haven. This letter includes both R&D Management and REC approval.

It is a requirement of your approval that you acknowledge use of the NHS Greater Glasgow & Clyde Safe Haven in any publications resulting from this project, and that you inform the Safe Haven of any such publications.

I am pleased to inform you, therefore, that your project has been approved.

Yours Sincerely,



Charlie Mayor
Safe Haven Manager
e : safehaven@ggc.scot.nhs.uk

Appendix E *Volumn of missingness of variables from Badgernet dataset*

Variable/row name	**Cohort** (include GDM and non-GDM pregnancies)	**GDM** (complete cases)
Total:	10,638	996
GDM	0.00%	0.00%
BookingDate	0.00%	0.00%
DateOfDelivery	0.00%	0.00%
CareLocationIDAtBookingName	0.00%	0.00%
GestationAtBookingWeeks	0.00%	0.00%
GestationAtBookingDays	0.00%	0.00%
BookingAge	0.00%	0.00%
BookingBMI	0.00%	0.00%
FamilyOrigin_Mother	7.30%	0.00%
FamilyOrigin_Mother_SMR02	7.30%	0.00%
SIMD_2016_QUINTILE	6.20%	5.92%
SIMD_2016_DECILE	6.20%	5.92%
SIMD_2016_VIGINTILE	6.20%	5.92%
Smoker_AtBooking	0.00%	0.00%
PreviousCS	0.00%	0.00%
NumberOfPreviousCS	0.00%	0.00%
Gravida	0.00%	0.00%
Para	0.00%	0.00%
prevmacro	0.00%	0.00%
prevgdm	0.00%	0.00%
FamilyHistoryDiabetes	1.00%	>1.00%
FamilyHistoryDiabetes_binary	0.00%	0.00%
MotherFamilyHistoryDiabetes	0.00%	0.00%
Diabetes	>1.00%	>1.00%

DiabetesDisplay	>1.00%	>1.00%
dtype1	0.00%	0.00%
dtype2	0.00%	0.00%
dtype12	0.00%	0.00%
DiabetesGestational	0.00%	0.00%
GlucoseOffered	0.00%	0.00%
DateNoteTimeGlucoseOffered	54.30%	1.61%
GlucoseRefused	0.00%	0.00%
GTTPerformed	0.00%	0.00%
Result_Fasting	56.50%	0.00%
ResultFasting2HrDate	56.50%	0.00%
gestation_at_fasting_result_week	89.70%	0.00%
gestation_at_fasting_result_day	89.70%	0.00%
GTTResult	55.20%	>1.00%
GTTPositive	55.20%	>1.00%
Result_HbA1C_Earliest	90.10%	80.62%
Result_HbA1C_EarliestDate	90.00%	80.62%
Result_HbA1C_NearBooking	65.20%	30.32%
Result_HbA1C_Booking	56.80%	20.28%
Result_HbA1C_2428w	97.80%	96.69%
MetforminDispDate	97.00%	74.90%
MetforminBNFName	95.10%	68.98%
Metformin	0.00%	0.00%
GestationAtMetforminDispWeek	97.00%	74.90%
GestationAtMetforminDispDay	97.00%	74.90%
metformin_prior	0.00%	0.00%
InsulinDispDate	98.80%	92.47%
InsulinBNFName	97.70%	89.86%
Insulin	0.00%	0.00%
GestationAtInsulinDispWeek	98.80%	92.47%

GestationAtInsulinDispDay	98.80%	92.47%
insulin_prior	0.00%	0.00%
Diet	89.70%	0.00%
PharmacologicalTherapy	89.70%	0.00%
PharmacologicalTherapyDispDate	97.20%	72.09%
GestationAtPharmacologicalTherapyDispWeek	97.20%	72.09%
GestationAtPharmacologicalTherapyDispDay	97.20%	72.09%
treatment	89.70%	0.00%
TotalNumberOfBabiesBorn	0.00%	0.00%
GestationAtBirthWeeks	0.00%	0.00%
GestationAtBirthDays	0.00%	0.00%
Premature	0.00%	0.00%
OutcomeDisplay	0.00%	0.00%
TypeOfDelivery	0.00%	0.00%
FinalTypeOfDelivery	>1.00%	>1.00%
ModeOfDeliveryDisplay	1.50%	>1.00%
Elective_caesarean_section	0.00%	0.00%
Emergency_caesarean_section	0.00%	0.00%
InductionMethodsDisplay	65.20%	61.14%
TotalNumberOfInstrumentalDeliveries	0.00%	0.00%
Tear_Display	25.50%	27.41%
tear_any	0.00%	0.00%
TotalBloodLoss	0.00%	0.00%
TotalBloodLossAntenatal	0.00%	0.00%
TotalBloodLossPostantal	0.00%	0.00%
TotalBloodLossWithin24HrBirth	0.00%	0.00%
TotalBloodLossWithin24HrBirthOperative	0.00%	0.00%
bpiSex	1.40%	>1.00%
BirthWeight_Grams	>1.00%	0.00%
BirthweightMinCentileWHO	2.00%	>1.00%

BirthweightMaxCentileWHO	2.00%	>1.00%
BirthweightCentileIntergrowth	14.70%	15.46%
macrosomia	>1.00%	0.00%
LGA	2.00%	>1.00%
ApgarScore01	4.30%	2.31%
ApgarScore05	4.20%	2.41%
ShoulderDystocia	2.00%	>1.00%
ShoulderDystociaGrade	99.60%	99.60%
ShoulderDystociaBabyObs	0.00%	0.00%
BabyAdmittedNNU	0.00%	0.00%

Appendix F *Data variables included in the cohort after identification and cleaning*

Variable/row name	Data type			Format*	Percentage Missing	
	Continuous	Discrete	Categorical		Cohort (include GDM and non-GDM pregnancies)	GDM (complete cases)
Total:					10,694	996
GDM			x	2 Categories	0.00%	-
Booking Date		x		Date	0.00%	0.00%
Date of Delivery		x		Date	0.00%	0.00%
Care Location at Booking			x	5 Categories	0.00%	0.00%
Gestation at Booking (Weeks)		x		Digit	0.00%	0.00%
Gestation at Booking (Days)		x		Digit	0.00%	0.00%
Booking Age		x		Digit	0.00%	0.00%
Booking BMI[a]	x			Digit	0.00%	0.00%
Mother's Family Origin			x	24 Categories	7.30%	0.00%
Mother's Family Origin Grouped			x	7 Categories	7.30%	0.00%
SIMD[b] 2016 Quintile			x	5 Categories	6.20%	5.92%
SIMD[b] 2016 Decile			x	10 Categories	6.20%	5.92%

Variable/row name	Data type			Format*	Percentage Missing	
	Continuous	Discrete	Categorical		Cohort (include GDM and non-GDM pregnancies)	GDM (complete cases)
SIMD[b] 2016 Vigintile			x	20 Categories	6.20%	5.92%
Smoker at Booking			x	2 Categories	0.00%	0.00%
Number of Previous Caesarean births		x		Digit	0.00%	0.00%
Previous Caesarean births			x	2 Categories	0.00%	0.00%
Gravida		x		Digit	0.00%	0.00%
Para		x		Digit	0.00%	0.00%
Previous Macrosomia Baby			x	2 Categories	0.00%	0.00%
Previous Gestational Diabetes			x	2 Categories	0.00%	0.00%
Family History of Diabetes (detailed)			x	913 Categories	1.00%	>1.00%
Family History of Diabetes			x	2 Categories	0.00%	0.00%
Mother's Family History of Diabetes			x	2 Categories	0.00%	0.00%
Diabetes			x	10 Categories	>1.00%	>1.00%

Variable/row name	Data type			Format*	Percentage Missing	
	Continuous	Discrete	Categorical		Cohort (include GDM and non-GDM pregnancies)	GDM (complete cases)
Diabetes Display			x	11 Categories	>1.00%	>1.00%
Diabetes Type 1			x	2 Categories	0.00%	0.00%
Diabetes Type 2			x	2 Categories	0.00%	0.00%
Diabetes Type 1 or 2			x	2 Categories	0.00%	0.00%
Diabetes Gestational			x	2 Categories	0.00%	0.00%
OGTT[c] Offered			x	2 Categories	0.00%	0.00%
Date OGTT[c] Offered		x		Date	54.30%	1.61%
OGTT[c] Refused			x	2 Categories	0.00%	0.00%
OGTT[c] Performed			x	2 Categories	0.00%	0.00%
Fasting OGTT[c] Result (mmol/l)	x			Digit	56.50%	0.00%
Fasting 2-hour OGTT[c] Date		x		Date	56.50%	0.00%
Gestation at Fasting OGTT[c] Result (week)		x		Digit	89.70%	0.00%
Gestation at Fasting OGTT[c] Result (day)		x		Digit	89.70%	0.00%

Variable/row name	Data type			Format*	Percentage Missing	
	Continuous	Discrete	Categorical		Cohort (include GDM and non-GDM pregnancies)	GDM (complete cases)
Has an OGTT[c] Result			x	2 Categories	55.20%	>1.00%
Has a Positive OGTT[c]			x	2 Categories	55.20%	>1.00%
Earliest HbA1c[d] Result (mmol/L)		x		Digit	90.10%	80.62%
Date of Earliest HbA1c[d] Result		x		Date	90.00%	80.62%
Near Booking HbA1c[d] Result (mmol/L)		x		Digit	65.20%	30.32%
At Booking HbA1c[d] Result (mmol/L)		x		Digit	56.80%	20.28%
24/28 weeks Gestation HbA1c[d] Result (mmol/L)		x		Digit	97.80%	96.69%
Metformin Dispense Date		x		Date	97.00%	74.90%
Metformin BNF[e] Name			x	2 Categories	95.10%	68.98%
Metformin			x	2 Categories	0.00%	0.00%
Gestation at Metformin Dispense (week)		x		Digit	97.00%	74.90%
Gestation at Metformin Dispense (day)		x		Digit	97.00%	74.90%
Metformin 3-6 months before booking date			x	2 Categories	0.00%	0.00%
Insulin Dispense Date		x		Date	98.80%	92.47%
Insulin BNF[e] Name			x	9 Categories	97.70%	89.86%

Variable/row name	Data type			Format*	Percentage Missing	
	Continuous	Discrete	Categorical		Cohort (include GDM and non-GDM pregnancies)	GDM (complete cases)
Insulin			x	2 Categories	0.00%	0.00%
Gestation at Insulin Dispense (week)		x		Digit	98.80%	92.47%
Gestation at Insulin Dispense (day)		x		Digit	98.80%	92.47%
Insulin 3-6 months before booking date			x	2 Categories	0.00%	0.00%
Diet			x	2 Categories	89.70%	0.00%
Pharmacological Therapy			x	2 Categories	89.70%	0.00%
Pharmacological Therapy Dispense Date		x		Date	97.20%	72.09%
Gestation at Pharmacological Therapy Dispense (week)		x		Digit	97.20%	72.09%
Gestation at Pharmacological Therapy Dispense (day)		x		Digit	97.20%	72.09%
Treatment for gestational diabetes			x	3 Categories	89.70%	0.00%
Total Number of Babies Born			x	1 Category	0.00%	0.00%
Gestation at Delivery (week)		x		Digit	0.00%	0.00%
Gestation at Delivery (day)		x		Digit	0.00%	0.00%
Premature			x	2 Categories	0.00%	0.00%
Outcome			x	5 Categories	0.00%	0.00%

Variable/row name	Data type			Format*	Percentage Missing	
	Continuous	Discrete	Categorical		Cohort (include GDM and non-GDM pregnancies)	GDM (complete cases)
Type of Delivery			x	5 Categories	0.00%	0.00%
Final Type of Delivery			x	5 Categories	>1.00%	>1.00%
Mode of Delivery			x	13 Categories	1.50%	>1.00%
Elective Caesarean Section			x	2 Categories	0.00%	0.00%
Emergency Caesarean Section			x	2 Categories	0.00%	0.00%
Induced Labour			x	23 Categories	65.20%	61.14%
Instrumental Delivery			x	2 Categories	0.00%	0.00%
Tear (detailed)			x	9 Categories	25.50%	27.41%
Any Tear			x	2 Categories	0.00%	0.00%
Total Blood Loss (ml)	x			Digit	0.00%	0.00%
Total Blood Loss Antenatal (ml)	x			Digit	0.00%	0.00%
Total Blood Loss Postnatal (ml)	x			Digit	0.00%	0.00%
Total Blood Loss Within 24Hr Birth (ml)	x			Digit	0.00%	0.00%

Variable/row name	Data type			Format*	Percentage Missing	
	Continuous	Discrete	Categorical		Cohort (include GDM and non-GDM pregnancies)	GDM (complete cases)
Total Blood Loss Within 24Hr Birth Operative (ml)	x			Digit	0.00%	0.00%
Baby's Sex			x	4 Categories	1.40%	>1.00%
Birth Weight (g)	x			Digit	>1.00%	0.00%
Birthweight Minimum Centile WHO[f]			x	11 Categories	2.00%	>1.00%
Birthweight Maximum Centile WHO[f]			x	11 Categories	2.00%	>1.00%
Birthweight Centile WHO[f] Intergrowth	x			Diget	14.70%	15.46%
Macrosomia (Birthweight ≥ 4500g)			x	2 Categories	>1.00%	0.00%
Large for Gestational age baby (Minimum Centile WHO[f] ≥ 90)			x	2 Categories	2.00%	>1.00%
1-minute Apgar Score			x	11 Categories	4.30%	2.31%
5-minute Apgar Score			x	12 Categories	4.20%	2.41%
Shoulder Dystocia			x	3 Categories	2.00%	>1.00%
Shoulder Dystocia Grade			x	4 Categories	99.60%	99.60%

Variable/row name	Data type			Format*	Percentage Missing	
	Continuous	Discrete	Categorical		Cohort (include GDM and non-GDM pregnancies)	GDM (complete cases)
Shoulder Dystocia Baby Observation			x	2 Categories	0.00%	0.00%
Baby Admitted to the Neonatal Intensive Care Unit			x	2 Categories	0.00%	0.00%
Comparable Public Health Scotland Year			x	2 Categories	0.00%	0.00%

x indicates the data type

**Format of the full cohort. The complete cases of GDM have fewer categories in some instances.*

[a]*BMI Body mass index*

[b]*SIMD Scottish index of multiple deprivation*

[c]*OGTT Oral glucose tolerance test*

[d]*HbA1c glycosylated haemoglobin*

[e]*BNF British National Formulary*

[f]*WHO World Health Organisation*

Appendix G *Descriptive statistical analysis of Diet vs Metformin treated GDM within the modelling data*

This analysis aimed to determine how different the women who did not receive insulin, i.e. those managed through diet or metformin, were from each other and is summarised in the **Table G.1** below. Of the 921 pregnancy episodes in the modelling data that did not require insulin, the majority were managed through diet (718, 72.1%), and the minority by metformin (203, 20.4%). Those treated with metformin, in comparison to those treated with diet, had a significantly higher age (diet mean[SD], 32[5.3] years, metformin: 33[4.8] years, p=0.002) and had more previous Caesarean births (Previous Caesarean birth: (N(%), 181 (25.2%), 66 (32.5%), p=0.047). Number of previous Caesarean births: 0.34[0.66], 0.44[0.74], p= 0.042), despite not having a significantly higher number of previous pregnancies (Gravida: 2.6[1.6], 2.7[1.5], p =0.256, Para: 2[1.1], 2.1[1.2], p=0.102). Other maternal characteristics were similar. The fasting glucose in the OGTT test was higher (5.2[0.5]mmol/L, 5.3[0.47]mmol/L, p=0.001) and taken earlier (26[6] weeks of gestation, 24 [5.6] weeks of gestation, p=0.002) in the metformin group compared to the diet group. But HbA1c at booking was similar for both metformin and diet-treated pregnancies (35[3.4]mmol/L, 35[3]mmol/L, p=0.931).

The significant differences between insulin and non-insulin-treated GDM may be influenced by the significant differences between diet and metformin-treated GDM. However, it is more helpful to HCPs to know who is at greater risk of requiring insulin, and hence it is still necessary to group diet and metformin together for the predictive model.

Table G.1 *Modelling data participant characteristics of non-insulin GDM treated pregnancies*

Characteristic	Diet (mean[standard deviation] or N(%))	Metformin (mean[standard deviation] or N(%))	p-value
Total	718 (72.1%)	203 (20.4%)	-
Gestational week at booking	10 [3]	10 [2.7]	0.190
Gestational day at booking	3.1 [2]	3.1 [1.9]	0.968
Booking age	32 [5.3]	33 [4.8]	0.002**
Booking BMI[a] (kg/m^2)	32 [7.3]	32 [7.5]	0.979
Family origin of mother[1]			0.074
... Group A – White	450 (62.0%)	110 (54.7%)	
... Group F – Other ethnic group	30 (4.0%)	10 (5.4%)	
SIMD 2016 quintile			
... 5	326 (45.4%)	85 (41.9%)	0.746
... 1	81 (11.3%)	27 (13.3%)	
... Unknown	42 (5.8%)	11 (5.4%)	
Smoker at booking	65(9.1%)	16 (7.9%)	0.704
Previous Caesarean birth	181 (25.2%)	66 (32.5%)	0.047**
Number of previous Caesarean births	0.34[0.66]	0.44 [0.74]	0.042**
Gravida	2.6 [1.6]	2.7 [1.5]	0.256
Para	2 [1.1]	2.1 [1.2]	0.102
Previous macrosomia baby	*	*	0.105
Previous gestational diabetes	118 (16.4%)	34 (16.7%)	1.00
Any family history of diabetes	410 (57.1%)	127 (62.6%)	0.189
Mother's family history of diabetes	340 (47.4%)	109 (53.7%)	0.129
Fasting glucose in OGTT[b] (mmol/L)	5.2 [0.5]	5.3 [0.47]	0.001**
Gestational week of OGTT[b]	26 [6]	24 [5.6]	0.002**
Gestational day of OGTT[b]	3.3 [1.8]	3.5 [1.8]	0.296
HbA1c[c] at booking (mmol/L)	35 [3.4]	35 [3]	0.931

[a]BMI Body mass index, [b]OGTT Oral glucose tolerance test, [c]HbA1c glycosylated haemoglobin. 1. Frequency rounded to the nearest 10. Redacted due to the potential risk of disclosure.** Statistically significant, p<0.05.*

Appendix H *User Needs Assessment for Gestational Diabetes Care and Management Digital Tool: A Qualitative Study [Published 35th Medical Informatics Conference 2025]*

The **Table H.1** below uses CRediT to show the author's contribution to the article entitled 'User Needs Assessment For Gestational Diabetes Care And Management Digital Tool: A Qualitative Study', published 35th Medical Informatics Conference 2025 (Kirkwood et al., 2025a).

Table H.1 *Author's contributor role, following CRediT, for the publication of 'User Needs Assessment For Gestational Diabetes Care And Management Digital Tool: A Qualitative Study', published 35th Medical Informatics Conference 2025*

Contributor Roles	**Author's initials**
Conceptualisation	JRK, RMR, DJW, RSL, AM
Data curation	JRK
Formal analysis	JRK, JD
Funding acquisition	RMR, DJW, RSL, AM
Investigation	JRK
Methodology	JRK
Project administration	JRK
Resources	RMR, RSL
Software	Not applicable
Supervision	RMR, DJW, RSL, AM
Validation	JD
Visualisation	Not applicable
Writing – original draft	JRK
Writing – review & editing	RMR, DJW, RSL, AM, JD, MS

JRK Jasmine R Kirkwood, JD Jane Dickson, AM Areti Manataki, RSL Robert S Lindsay, DJW Deborah J Wake, RMR Rebecca M Reynolds

Intelligent Health Systems – From Technology to Data and Knowledge
E. Andrikopoulou et al. (Eds.)

doi:10.3233/SHTI250511

User Needs Assessment for Gestational Diabetes Care and Management Digital Tool: A Qualitative Study

Jasmine R KIRKWOOD[a,1] Jane DICKSON[b], Areti MANATAKI[c], Robert S LINDSAY[d], Deborah J WAKE[e], and Rebecca M REYNOLDS[a]
[a] *Centre for Cardiovascular Science, University of Edinburgh,*
[b] *University of Dundee Medical School,*
[c] *School of Computer Science, University of St Andrews,*
[d] *School of Cardiovascular and Metabolic Health, University of Glasgow,*
[e] *Usher Institute, University of Edinburgh.*
ORCiD ID: Jasmine R Kirkwood https://orcid.org/0000-0002-4349-232X

Abstract. Gestational diabetes mellitus (GDM) is one of the most common pregnancy complications, with both short- and long-term consequences. We conducted a user needs assessment of GDM care and management through semi-structured interviews with healthcare professionals and women with GDM. Examining current GDM care, we found that time management, number and attendance of appointments, information, resources and education, and cultural social and language differences were affecting factors. Furthermore, participants suggested a digital tool to include notification and prompts, risk stratification, and information and resources. The results of this study form the foundation of a user-centred digital tool that has the potential to transform GDM care and management.

Keywords. Gestational diabetes, mHealth, telemedicine, qualitative study, user-centred design

1. Introduction

Gestational diabetes mellitus (GDM) is one of the most common pregnancy complications [1], and due to rising maternal age when pregnant and obesity, it is expected to increase [2,3]. Women who have GDM, have a 10-fold increase in developing subsequent type 2 diabetes [4], and a 2-fold increase in developing early cardiovascular disease [5]. Treatment for GDM requires modifications to diet and exercise, medication, and self-monitoring of blood glucose levels which are reviewed at a multidisciplinary clinic [6].

Making lifestyle adjustments for GDM can be difficult and complex [7]. It has been reported that digital tools for GDM can lead to improvement in achievement of blood glucose targets reducing adverse pregnancy outcomes and in-person appointments [8].

Involving end-users in the development of a new digital tool improves the effectiveness of the end result [9]. Previous GDM digital tools have not been developed

[1] Corresponding Author: Jasmine R Kirkwood; E-mail: J.R.Kirkwood@sms.ed.ac.uk.

with both HCPs and women with GDM simultaneously. To address this gap and intending to develop a digital tool for GDM care and management, we sought the views of end users, HCPs and women with GDM. We aimed to understand how a digital tool could improve GDM care services, through semi-structured interviews. We focused on two research questions: 1. What is end-user's experience of GDM care currently, and 2. What would an end-user want from a digital tool for GDM care?

2. Methods

Our study consists of semi-structured interviews of potential end-users, women with GDM and HCPs, from Edinburgh and Glasgow. Prior to the interviews, participants were given a patient information sheet and gave written and verbal consent. The lead author (JRK) had an 'outsider' perspective and conducted the interviews following a topic guide. Interviews were undertaken in-person and online between January-July 2023 and lasted around 30 minutes. As interviews are known to provide a rich source of information, a minimum of six to ten participants is deemed sufficient [10]. Accordingly, a total of 11 participants were recruited (6 HCPs, 5 women with GDM).

Thematic analysis of transcripts was carried out using the Braun and Clarke framework [11]. Interviews were audio-recorded and transcribed, transcriptions were manually and iteratively coded using NVivo 12 software V3, with 10% double coded by JD. Codes and final themes were refined through discussion between JRK and JD. Results are reported using the Standard for Reporting Qualitative Research (SRQR) [12]. These results focus on current GDM care and how a digital tool could improve it. Themes are highlighted in bold, and quotes are presented with participant type and number with nonrelevant sections denoted […]. Ethical approval was granted from the Research Ethics Committee (reference number 22/NW/0164), and co-sponsored by NHS Lothian and the University of Edinburgh through the Academic and Clinical Central Office for Research and Development.

3. Results

There was a total of 11 interviews, (6 HCPs, 5 Women with GDM). A summary of the themes is shown in Table 1. Participants were in Edinburgh (82% 9/11, HCPs: 67%, 4/6, Women with GDM: 100%, 5/5), and Glasgow (18%, 2/11, HCPs: 33%, 2/6). HCPs' roles included consultant physician (50%, 3/6), consultant obstetrician (17%, 1/6), dietitian (17%, 1/6) and diabetes specialist nurse (17%, 1/6). Women with GDM were between 21-40 years, had an ethnic background of white (40%, 2/5), mixed (40%, 2/5) or black (20%, 1/5) with 80% (4/5) having had previous GDM.

3.1. What is an end-user's experience of GDM care currently?

Current care and experience of GDM was discussed with participants. **Time management**, was most frequently raised. HCPs had concerns about limited specialised staff and restricted clinical hours. While women with GDM discussed the challenges of making the necessary lifestyle changes because of their other parental, work and life commitments.

"It is tiring having to keep track of everything and I can't be bothered, especially with work and being a mum already, I'm quite busy, so it is trying to balance my diet is quite hard" (Woman with GDM 5)

This was coupled with **number and attendance of appointments**.

"Because of the higher BMI, I already had extra appointments. And now it's like gestational (diabetes) on top of that. […] so, this week, I've already got three appointments. So, I have this one today (metabolic-antenatal), the midwife, and I've got the nutritionist again this week." (Woman with GDM 1)

Additionally, **Information, resources and education** was discussed. There can be a steep learning curve when diagnosed with GDM, with comprehending blood glucose readings alongside understanding the risks and complications that are associated with GDM. This is further complicated by different guidelines. Women with GDM often reported using internet searches rather than the information provided by the hospital, HCPs were sympathetic towards this.

"I found that most of the NHS trusts have different rules and different (blood glucose) readings are high, different (blood glucose) readings are low and these are okay and these aren't okay." (Woman with GDM 2)

Finally, GDM affects a diverse population, and it was highlighted that **cultural, social and language differences** could affect treatment and compliance.

"Asking people to change very basic habits, that's difficult, and people of course don't only eat as individuals, they eat as families, therefore changing things can be difficult for that reason" (HCP 2)

3.2. What would an end-user want from a digital tool for GDM care?

Our driving aim is to understand what end-users would want from a digital tool to improve GDM care and management. Participants expressed a desire for a system of **notification and prompts** to help both HCPs manage clinics and help to facilitate easier self-management for women with GDM.

"I think a dashboard would be amazing, rather than them (diabetic specialist nurse or midwife) having, so, at the end of the clinic they make a list of all women in they have to phone in the virtual clinic." (HCP 5)

"Built into some algorithm that alerts the patient to phone for advice." (HCP 5)

Furthermore, there was seen to be a potential to personalise care through **risk stratification**, particularly from HCPs.

"A big enough dataset to try to then personalise those targets a bit more, pick out women who we know are at more risk and alter that." (HCP 2)

Finally, there was a need to have reliable **information and resources**, all in one place.

"There is an opportunity both to improve availability of information" (HCP 6)

Table 1. Summary of themes and subtheme with participant mentioning percentages.

Theme	Subtheme	Participant mentioning (%)
What is an end-user's experience of GDM care currently?	Time management	Overall: 73%, 8/11, HCPs: 83%, 5/6, Women with GDM: 60%, 3/5
	Number and attendance of appointments	Overall: 64%, 7/11, HCPs: 67%, 4/6, Women with GDM: 60%, 3/5
	Information, resources and education	Overall: 36%, 4/11, HCPs: 33%, 2/6, Women with GDM: 40%, 2/5
	Cultural, social and language differences	Overall: 55%, 6/11, HCP: 100%, 6/6

What would an end-user want from a digital tool for GDM care?	Notification and prompts	Overall: 82%, 9/11, HCPs: 100%, 6/6, Women with GDM: 60%, 3/5
	Risk stratification	Overall: 64%, 7/11, HCPs: 100%, 6/6, Women with GDM: 20%, 1/5
	Information and resources	Overall: 64%, 7/11, HCPs: 67%, 4/6, Women with GDM: 60%, 3/5

4. Discussion

Aiming to understand how a digital tool could improve GDM care services, we conducted semi-structured interviews with HCPs and women with GDM. We found that current GDM care is time consuming and challenging, and a digital tool for GDM would need to help mitigate these issues. This was suggested through risk stratification and notification or prompts alongside credible information.

A digital tool could allow for remote monitoring and consultation, reducing the burden of the additional appointment for women with GDM and improve attendance. It has already been shown to reduce clinical visits and improve quality of life [8].

Notification and prompt to highlight to HCPs women who need assistance was reported in a qualitative study by Safiee et al. [13], as well. They found there was the potential for better self-management and suggested the timesaving benefits. Furthermore, blood glucose reminder systems for GDM were associated with a statistically significant increase in patient compliance and a decrease in mean blood glucose values [14].

Information and education provided to women with GDM needs to be reliable, accessible and concise. A digital tool could be a platform for this, which aligns with previous studies [15].

To further aid HCPs in managing clinics, risk stratification was explored. Risk stratification models [16], including prediction models of GDM [17], medication [18] or complications [19], have been developed but there is a lack of qualitative work on how HCPs and women with GDM would perceive them within GDM care.

Finally, our study was limited geographically to two major cities in Scotland. To strengthen our findings, further work in different locations and with a diverse range of HCPs' roles and women with GDM backgrounds would be beneficial.

5. Conclusions

Our study found that current GDM care can be time consuming and education resources could be improved. A digital tool for GDM care would need to be simple and integrate easily into both HCPs' clinics and women with GDM's lives. It could include notification and prompts to better self-manage GDM and the running of the GDM clinics and provide information and diet resources. In addition, care could be personalised through risk stratification of woman with GDM that need addition help.

There is an opportunity for further qualitative research with HCPs' and women with GDM's on risk stratification within care. Ultimately, these results form a basis for the development of a user-centred digital tool that could revolutionise GDM care.

References

[1] International Diabetes Federation, IDF Diabetes Atlas Tenth Edition 2021: 2021; 2021 [10:[Available from: https://diabetesatlas.org/idfawp/resource-files/2021/07/IDF_Atlas_10th_Edition_2021.pdf.

[2] Johns EC, Denison FC, Norman JE, Reynolds RM. Gestational Diabetes Mellitus: Mechanisms, Treatment, and Complications. Trends in Endocrinology & Metabolism. 2018;29(11):743-54.

[3] Ferrara A. Increasing Prevalence of Gestational Diabetes Mellitus. Diabetes Care. 2007;30(Supplement_2):S141-S6.

[4] Vounzoulaki E, Khunti K, Abner SC, Tan BK, Davies MJ, Gillies CL. Progression to type 2 diabetes in women with a known history of gestational diabetes: systematic review and meta-analysis. BMJ. 2020:m1361.

[5] Daly B, Toulis KA, Thomas N, Gokhale K, Martin J, Webber J, et al. Increased risk of ischemic heart disease, hypertension, and type 2 diabetes in women with previous gestational diabetes mellitus, a target group in general practice for preventive interventions: A population-based cohort study. PLOS Medicine. 2018;15(1):e1002488.

[6] National Institute for Health and Care Excellence (NICE). Diabetes in pregnancy: management from preconception to the postnatal period, 2015, 25 February [updated 2020, 16 December. Available from: https://www.nice.org.uk/guidance/ng3/resources/diabetes-in-pregnancy-management-from-preconception-to-the-postnatal-period-pdf-51038446021.

[7] He J, Chen X, Wang Y, Liu Y, Bai J. The experiences of pregnant women with gestational diabetes mellitus: a systematic review of qualitative evidence. Reviews in Endocrine and Metabolic Disorders. 2021;22(4):777-87.

[8] Rasekaba TM, Furler J, Blackberry I, Tacey M, Gray K, Lim K. Telemedicine interventions for gestational diabetes mellitus: A systematic review and meta-analysis. Diabetes Research and Clinical Practice. 2015;110(1):1-9.

[9] Chamberlain S, Dutt P, Mitra R, Godfrey A, Lefevre AE, Scott K, et al. Lessons learnt from applying a human-centred design process to develop one of the largest mobile health communication programmes in the world. BMJ Innovations. 2022:bmjinnov-2021-0.

[10] Terry G, Hayfield N. Essentials of thematic analysis / Gareth Terry, Nikki Hayfield. Washington, DC: American Psychological Association; 2021.

[11] Braun V, Clarke V. Using thematic analysis in psychology. Qualitative Research in Psychology. 2006;3(2):77-101.

[12] O'Brien BC, Harris IB, Beckman TJ, Reed DA, Cook DA. Standards for Reporting Qualitative Research. Academic Medicine. 2014;89(9):1245-51.

[13] Safiee L, Rough DJ, Whitford H. Barriers to and Facilitators of Using eHealth to Support Gestational Diabetes Mellitus Self-management: Systematic Literature Review of Perceptions of Health Care Professionals and Women With Gestational Diabetes Mellitus. Journal of Medical Internet Research. 2022;24(10):e39689.

[14] Horgan R, Pierce-Williams R, Saccone G, Berghella V. Reminder Systems to Increase Compliance With Glucose Logging in Gestational Diabetes: A Systematic Review & Meta-analysis. American Journal of Obstetrics & Gynecology MFM. 2022:100586.

[15] Garnweidner-Holme L, Hoel Andersen T, Sando MW, Noll J, Lukasse M. Health Care Professionals' Attitudes Toward, and Experiences of Using, a Culture-Sensitive Smartphone App for Women with Gestational Diabetes Mellitus: Qualitative Study. JMIR mHealth and uHealth. 2018;6(5):e123.

[16] Harrison CL, Lombard CB, East C, Boyle J, Teede HJ. Risk stratification in early pregnancy for women at increased risk of gestational diabetes. Diabetes Research and Clinical Practice. 2015;107(1):61-8.

[17] Mennickent D, Rodríguez A, Farías-Jofré M, Araya J, Guzmán-Gutiérrez E. Machine learning-based models for gestational diabetes mellitus prediction before 24–28 weeks of pregnancy: A review. Artificial Intelligence in Medicine. 2022;132:102378.

[18] Alvarez-Silvares E, Bermúdez-González M, Vilouta-Romero M, García-Lavandeira S, Seoane-Pillado T. Prediction of insulin therapy in women with gestational diabetes: a systematic review and meta-analysis of observational studies. Journal of Perinatal Medicine. 2022;50(5):608-19.

[19] Cooray SD, Wijeyaratne LA, Soldatos G, Allotey J, Boyle JA, Teede HJ. The Unrealised Potential for Predicting Pregnancy Complications in Women with Gestational Diabetes: A Systematic Review and Critical Appraisal. International Journal of Environmental Research and Public Health. 2020;17(9):3048.

Appendix I *Supplementary material of The User-Centered Design of a Clinical Dashboard and Patient-Facing App for Gestational Diabetes [Published in the Journal of Diabetes Science and Technology]*

SUPPLEMENTARY MATERIAL 1: INTERVIEW GUIDES

Health care professionals

Pre-interview

- Greet and introduction.
- Explain the interview - the interview will last around 30 minutes and the questions will involve four topics – about you and your experiences with providing care for gestational diabetes and the use of digital tools in the services you provide.
- Gain verbal consent for audio-recording the interview and remind participants that they do not have to answer any questions if they do not feel comfortable and can stop the interview at any time.

Underlined questions are key questions asked in every interview.

Research Question	Key Concepts	Example Interview Questions
What is current care like that health care professionals provide to women with gestational diabetes?	Current care Improvements Caregiving	Describe to me how you care for women with gestational diabetes. How often will you see a patient with gestational diabetes during their pregnancy? Do you provide a woman with gestational diabetes any education on their condition? If so, how do you provide this? Do you recommend/do you patients uses any non-hospital resources for education or advice

		for women with gestational diabetes? (e.g. eLearning course, websites) <u>What are the barriers and enablers for women to comply with recommended lifestyle changes or blood glucose monitoring/recording?</u> Do you think/feel there could be any improvements in the area of care you provide for women with gestational diabetes? What would these improvements be? What system is in place for getting women who had gestational diabetes to return for their 12-week postpartum and yearly glucose screening?
What technology/digital tools and types of data are captured in the care of women with gestational diabetes?	Technology Data type/capture Current care	<u>Explain to me the type of technology, if any, you use in the care you provide for women with gestational diabetes.</u> Do you suggest or recommend any technology or digital tools to your patients with gestational diabetes that are not provided by the hospital? If so, please explain to me what they are.

		Does the technology used in gestational diabetes care you provide capture any form of data, if so, what does it capture?
What are the gestational diabetes health care professional's acceptability of digital tools and computer decisions/support in gestational diabetes care?	Digital tool/technology Acceptability	Explain to me how confident you would be if computer decision-making was part of the care pathway for gestational diabetes. Would you be confident with a digital tool giving you or women with gestational diabetes automated advice or suggestions that were generated through machine learning? Such as highlighting women who have consistently had high blood glucose readings or giving women with gestational diabetes feedback on their glucose readings. How helpful do you think a tool that predicts women's pregnancy complications or outcomes would be? For example, predicting which women are more likely to need mediation or have a Caesarean section. What risks do you see with having a computer assistant with

		decisions in gestational diabetes care? What benefits or limitations do you think there could be for both you and your patients for the use of digital tools or computer decisions?
What would healthcare professionals want from a digital tool that helps with gestational diabetes care and management?	Digital tool Platform/interface Feature	<u>How would you like a digital tool to help you with the care and management of women with gestational diabetes?</u> If a digital tool were developed for healthcare professionals who care for gestational diabetes, what features would you want it to have? How do you anticipate the support system could support HCPs in providing care? How would you like to access this digital tool for healthcare professionals? If a digital tool was developed for the women with gestational diabetes that you care for, what features would you want it to have for them?

		How would you like your patients with gestational diabetes to access this digital tool?

Women with GDM

Pre-interview

- Greet and introduction.
- Explain the interview - the interview will last around 10/30 minutes and the questions will involve four topics about you and your experiences with gestational diabetes and the use of digital tools in the services you receive.
- Gain verbal consent for audio-recording the interview and remind participants that they do not have to answer any questions if they do not feel comfortable and can stop the interview at any time.

Underlined questions are key questions asked in every interview.

Research Question	**Key Concepts**	**Example Interview Questions**
What was the care like for women with gestational diabetes?	Current care Improvements Experience of GDM	<u>Describe your experience with gestational diabetes.</u> - How was it taking BG to read regularly - How was going to extra appointments - How was changing diet or increasing exercise What treatment plan are you on? (diet, metformin, insulin) How is this treatment? Explain to me how you managed your gestational diabetes.

		How often did you have to go to the clinic or have appointments? What was this like? How did you find the education for gestational diabetes? Did you use only the hospital resources, or did you search for more information as well? Was it/do you think/ do you feel easy or hard to remember everything you were told? Was/do you think/do you feel the advice and education on diet tailored to your cultural or food preferences? No? How could this be improved? <u>How was the adaption to lifestyle changes, and what were the barriers or enablers to these changes?</u> What were your main worries or concerns about having gestational diabetes and its possible problems, e.g. large baby, type 2 diabetes? What did you like and not like about gestational diabetes care?

		What improvements do you think there could be? /What would you change? <u>Postpartum only.</u> How were you reminded about your 12-week postpartum OGTT? And did you go to that appointment?
What technology/digital tools and type of data is captured in the care of women with gestational diabetes?	Technology Data type/capture Current care	<u>Did you use any technology for your gestational diabetes (or pregnancy) and if so, what were they?</u> Was this suggested or supplied by the hospital? How easy or hard was it to use? Did it capture any data? If so, what did it record? Did you use any other technology to help you manage your gestational diabetes that wasn't recommended by the hospital? Please explain to me what this was. How did you find this, was it beneficial or not?
What are women with gestational diabetes'	Digital tools/technology	<u>Do you think/feel that if there was a digital tool that had some form</u>

acceptability to digital tools or technology for gestational diabetes care?	Acceptability	of automated advice or suggestion on gestational diabetes management you would find this helpful? Such as giving instant feedback on your blood sugar reading letting you know if you're in the right range or not and how to make the corrections if needed. Explain to me how confident would you be if computer-based decision-making was used in your gestational diabetes care. For example, the computer could make the decision based on the blood sugar reading you've recently had that you should increase your physical activity. What about computer assistance with medication? Administrating? Predicting if you were to need it? How comfortable would you be with a computer predicting that you are more likely to have complications associated with GDM? What complications do you think it would be helpful to predict?

What would women with gestational diabetes want from a digital tool that helped with gestational diabetes care and management?	Digital tool Platform/interface Features	What would you like a digital tool to help you within gestational diabetes care and management? If a digital tool was made for gestational diabetes, how would you like to access it? What would you want from a digital tool for gestational diabetes care?

SUPPLEMENTARY MATERIAL 2 – FEEDBACK SESSIONS

Feedback session guide

Prototype Demonstration

Demonstrate both the clinical dashboard and patient app to the feedback group.

Dashboard (Figma prototype demonstration)

- Home page
- Home page icons
- Example patients, 7,3 and 1
- Message to patients
- Add new patients

Slides (PowerPoint)

- Example of medication prediction outputs

App (Figma prototype demonstration)

- Home page
- Home page, add notes and/or treatments
- Menu
- Request a call-back
- Menu – food log main page (including recipes and GDM community)
- Home -> education
- Medication
- Medication quiz

Slides (PowerPoint)

- Example of notifications for patients

Discussion points/starting questions

Key Concepts	Example Questions
Notification of off-target blood glucose	1. What do you think about the dashboard notifying you of off-target patients? 2. What do you think of the in-app notifications and reminders for your patient if they are off-target?
Education, information and resources	1. What do you think about the education resources for your patients are they [helpful, useful, appropriate]?
Usability	1. Do you think this would increase/decrease workload? And why? 2. Could you see this being beneficial to your clinic and/or patients? And why? 3. Did the dashboard (and app for patients) seem easy or hard to use? And why? 4. Do you feel that the app is socially and culturally appropriate for your patient's demographic?
Predictive modelling/risk stratification	1. What do you think/feel about the risk stratification of patients through predicted medication needs?
Overall	1. [Likes/dislikes], [Improvements/missing] of any section. 2. How do you see the dashboard and app being used in your clinics? 3. What do you think of patients being able to request a callback? 4. What do you think about being able to send patients a message? 5. What do you think of setting up/adding new patients?

Characteristics of feedback session participants

Characteristics of 31 participants in the three feedback sessions that evaluated the MyGDM prototype.

	Role	Number
Feeback session 1	Clinical consultant and academic	2
	Research midwife	7
	Clinical research facilitators	4
	PhD Students and clinical research fellows	3
Feeback session 2	Consultant Obstetrician	3
	Consultant Diabetologist	4
	Diabetes Specialist Nurse	4
Feeback session 3	Diabetic specialist midwife	4

Iterative changes from feedback sessions

Elements that were iteratively added to the MyGDM prototype from each feedback session (*SMBG: Self-monitoring blood glucose)*

Feedback session	Elements added to the MyGDM prototype
1	Call back information: • Specify who is going to be called. • Add office hours.

	• Explain what concerns the function should be used for. • Add a triage number for other concerns.
	Add a technical support section.

2	Clinical dashboard sort/filter function.
	• Add the ability to sort clinical dashboard:
	  • Off-target SMBG most to least, and least to most.
	  • Call back request only.
	  • Warning only.
	  • Alphabetically.

	Visualization of patient's information:
	• Add visualization:
	  • Line graph of SMBG readings.
	  • Pie chart of on- and off-target SMBG.

3	Clinical dashboard sort/filter function: • By clinical location. • By medication type.
	Visualization of patient's information: • Format the patient's SMBG in a logbook style, grouping readings by day.
	Highlight which SMBG readings had been edited or added manually.

SUPPLEMENTARY MATERIALS 3 – QUESTIONNAIRE

Questionnaire

Participants will be asked to watch a short video that demonstrates the MyGDM prototype app and then to complete the questionnaire.

Questionnaire

Welcome

Thank you for taking part in the study to assess the prototype MyGDM app.

In this study, we wish to find out what women with gestational diabetes first impressions are of the MyGDM app.

All the information we collect during the research will be kept confidential and there are strict laws which safeguard your privacy at every stage. However, NHS Lothian cannot guarantee the safety of your information when you are using your own devices (e.g., your own mobile or laptop).

Background

1. **Age**
    a. Under 16
    b. 16-20
    c. 21-25
    d. 26-30
    e. 31-35
    f. 36-40
    g. 41-45
    h. 46-50
    i. 50 +
    j. Prefer not to answer
2. **Ethnic identity**
    a. Asian/Asian British
        i. Indian

ii. Pakistani
iii. Bangladeshi
iv. Chinese
v. Any other Asian background, please describe

b. Black / African / Caribbean / Black British
   i. African
   ii. Caribbean
   iii. Any other Black / African / Caribbean background, please describe

c. Mixed / Multiple ethnic groups
   i. White and Black Caribbean
   ii. White and Black African
   iii. White and Asian
   iv. Any other Mixed / Multiple ethnic background, please describe

d. White
   i. English / Welsh / Scottish / Northern Irish / British
   ii. Irish
   iii. Gypsy or Irish Traveller
   iv. Any other White background, please describe

e. Other:
   i. Please state

f. Prefer not to say

3. When were you diagnosed with gestational diabetes?

a. 0-7 days ago
b. 8-14 days ago
c. 15-30 days ago
d. 1 to 2 months ago
e. 3 to 6 months ago
f. 7 to 9 months ago

4. Do you have any other children?

a. Yes

b. No

5. **Have you had gestational diabetes in a previous pregnancy?**
    a. Yes
    b. No
6. **What medication did/do you take for your gestational diabetes?**
    a. None
    b. Metformin
    c. Insulin
    d. Don't know
    e. Prefer not to say
7. **Would you like to be updated with a short report once the study is completed?**
    a. Yes
    b. No

Questionnaire

1) **Do you think that the app presented here would have helped manage your gestational diabetes?** [select one]
    a. Yes, it would have helped manage my gestational diabetes.
    b. No, it would not have helped manage my gestational diabetes.
    c. Comments
2) **Could you see this app fitting in with your lifestyle to help manage your gestational diabetes?**
    a. Yes, it would have fitted in with my lifestyle,
    b. No, it would not have fitted in with my lifestyle.
    c. Comments
3) On a scale of 1 to 5, 1 being very unlikely and 5 being likely.
    How likely, if you had need, would you have used the 'request a call' feature?
    (very unlikely) 1 2 3 4 5 (very likely)
4) On a scale of 1 to 5, 1 being very unlikely and 5 being likely.

How likely, would you have used the educational resources?

(very unlikely) 1 2 3 4 5 (very likely)

5) On a scale of 1 to 5, 1 being very unlikely and 5 being likely.

   How likely, would you have used the course/quiz?

   (very unlikely) 1 2 3 4 5 (very likely)

6) **What do you like the most about the digital tool?** [short answer]
7) **What do you like the least about the digital tool?** [short answer]
8) **Is there anything that you expected the prototype type to have that you think is missing?**

   [short answer]

9) **Do you have any other comments that you would like to add about this digital tool prototype?** [short answer]

   a. Yes

      Please state:

   b. No

Thank you!

Thank you for taking the time to complete the questionnaire.

If you have any further questions please contact Jazz Kirkwood, j.r.kirkwood@sms.ed.ac.uk

Characteristics of questionnaire participants

Characteristics of 13 participants who took the questionnaire to evaluate the MyGDM prototype app

Overview	Value	Number (percentage response, %)
Age (years)	21-25	2 (15.4)
	26-30	2 (15.4)
	31-35	4 (30.8)
	36-40	4 (30.8)
	41-45	1 (7.7)
Ethnicity	Asian/Asian British	2 (15.4)
	Mixed / Multiple ethnic groups	2 (15.4)
	White	9 (69.2)
Gestational diabetes diagnosis before taking the questionnaire	8 to14 days ago	1 (7.7)
	15 to 30 days ago	2 (15.4)
	1 to 2 months ago	2 (15.4)
	3 to 6 months ago	3 (23.1)
	7 to 9 months ago	5 (38.5)
Other children	Yes	7 (53.9)
	No	6 (46.2)
Gestational diabetes in a previous pregnancy	Yes	6 (46.2)
	No	7 (53.9)
Treatment for Gestational diabetes at the time of the questionnaire	Diet	5 (38.5)
	Metformin	5 (38.5)
	Metformin and insulin	3 (23.1)

Appendix J *Ethical approval letter: User needs assessment of computer-decision-making digital tools for gestational diabetes care (22/NW/0164)*

NHS
Health Research Authority

North West - Greater Manchester Central Research Ethics Committee
3rd Floor
Barlow House
4 Minshull Street
Manchester
M1 3DZ

16 June 2022

Professor Rebecca Reynolds
University Of Edinburgh
The Queens Medical Research Institute, 47 Little France Crescent
Edinburgh
EH164TJ

Dear Professor Reynolds

Study title:	**User needs assessment of computer-decision making digital tools for gestational diabetes care.**
REC reference:	**22/NW/0164**
Protocol number:	**AC22044**
IRAS project ID:	**310656**

Thank you for your submission, responding to the Proportionate Review Sub-Committee's request for changes to the documentation for the above study.

The revised documentation has been reviewed and approved on behalf of the PR sub-committee.

Confirmation of ethical opinion

On behalf of the Research Ethics Committee (REC), I am pleased to confirm a favourable ethical opinion for the above research on the basis described in the application form, protocol and supporting documentation as revised.

Good practice principles and responsibilities

The UK Policy Framework for Health and Social Care Research sets out principles of good practice in the management and conduct of health and social care research. It also outlines the responsibilities of individuals and organisations, including those related to the four elements of research transparency:

1. registering research studies
2. reporting results
3. informing participants
4. sharing study data and tissue

Conditions of the favourable opinion

The REC favourable opinion is subject to the following conditions being met prior to the start of the study.

Confirmation of Capacity and Capability (in England, Northern Ireland and Wales) or NHS management permission (in Scotland) should be sought from all NHS organisations involved in the study in accordance with NHS research governance arrangements. Each NHS organisation must confirm through the signing of agreements and/or other documents that it has given permission for the research to proceed (except where explicitly specified otherwise).

Guidance on applying for HRA and HCRW Approval (England and Wales)/ NHS permission for research is available in the Integrated Research Application System.

For non-NHS sites, site management permission should be obtained in accordance with the procedures of the relevant host organisation.

Sponsors are not required to notify the Committee of management permissions from host organisations.

Registration of Clinical Trials

All research should be registered in a publicly accessible database and we expect all researchers, research sponsors and others to meet this fundamental best practice standard.

It is a condition of the REC favourable opinion that **all clinical trials are registered** on a publicly accessible database within six weeks of recruiting the first research participant. For this purpose, 'clinical trials' are defined as:

- clinical trial of an investigational medicinal product
- clinical investigation or other study of a medical device
- combined trial of an investigational medicinal product and an investigational medical device
- other clinical trial to study a novel intervention or randomised clinical trial to compare interventions in clinical practice.

Failure to register a clinical trial is a breach of these approval conditions, unless a deferral has been agreed by the HRA (for more information on registration and requesting a deferral see: Research registration and research project identifiers).

If you have not already included registration details in your IRAS application form you should notify the REC of the registration details as soon as possible.

Publication of Your Research Summary

We will publish your research summary for the above study on the research summaries section of our website, together with your contact details, no earlier than three months from the date of this favourable opinion letter.

Should you wish to provide a substitute contact point, make a request to defer, or require further information, please visit: https://www.hra.nhs.uk/planning-and-improving-research/application-summaries/research-summaries/

N.B. If your study is related to COVID-19 we will aim to publish your research summary within 3 days rather than three months.

During this public health emergency, it is vital that everyone can promptly identify all relevant research related to COVID-19 that is taking place globally. If you haven't already done so, please register your study on a public registry as soon as possible and provide the REC with the registration detail, which will be posted alongside other information relating to your project. We are also asking sponsors not to request deferral of publication of research summary for any projects relating to COVID-19. In addition, to facilitate finding and extracting studies related to COVID-19 from public databases, please enter the WHO official acronym for the coronavirus disease (COVID-19) in the full title of your study. Approved COVID-19 studies can be found at: https://www.hra.nhs.uk/covid-19-research/approved-covid-19-research/

It is the responsibility of the sponsor to ensure that all the conditions are complied with before the start of the study or its initiation at a particular site (as applicable).

After ethical review: Reporting requirements

The attached document "After ethical review – guidance for researchers" gives detailed guidance on reporting requirements for studies with a favourable opinion, including:

- Notifying substantial amendments
- Adding new sites and investigators
- Notification of serious breaches of the protocol
- Progress and safety reports
- Notifying the end of the study, including early termination of the study
- Final report
- Reporting results

The latest guidance on these topics can be found at https://www.hra.nhs.uk/approvals-amendments/managing-your-approval/.

Ethical review of research sites

The favourable opinion applies to all NHS/HSC sites taking part in the study, subject to management permission being obtained from the NHS/HSC R&D office prior to the start of the study (see "Conditions of the favourable opinion" above).

Approved documents

The documents reviewed and approved by the Committee are:

Document	*Version*	*Date*
Copies of materials calling attention of potential participants to the research [Flyer GDM women v1]	1	06 May 2022
Copies of materials calling attention of potential participants to the research [Flyer HCP v1]	1	06 May 2022
Evidence of Sponsor insurance or indemnity (non NHS Sponsors only) [Insurance]		01 August 2021
Interview schedules or topic guides for participants [Interview guide GDM women v3]	3	06 May 2022
Interview schedules or topic guides for participants [Interview guide HCP v3]	3	06 May 2022
IRAS Application Form [IRAS_Form_03052022]		03 May 2022
Letter from funder [Funding letter]		22 February 2021
Letters of invitation to participant [Postpartum study invite letter v1]	1	06 May 2022
Letters of invitation to participant [HCP study invite email v0]	0	06 May 2022
Non-validated questionnaire [Questionnaire GDM women v3]	3	06 May 2022
Non-validated questionnaire [Questionnaire HCP v2]	2	06 May 2022
Non-validated questionnaire [Screening questions GDM women v1]	1	06 May 2022
Non-validated questionnaire [screening question HCP v1]	1	06 May 2022
Other [Application Clarification]		09 May 2022
Other [Second academic supervisor cv v0]	0	09 May 2022
Other [REC response]		27 May 2022
Other [Example digital tool v0]	0	27 May 2022
Participant information sheet (PIS) [PIS GDM women v4]	4	06 May 2022
Participant information sheet (PIS) [PIS HCP v4]	4	06 May 2022
Research protocol or project proposal [Protocol v3]	3	06 May 2022
Research protocol or project proposal [Protocol v4]	4	27 May 2022
Summary CV for Chief Investigator (CI) [CI CV v0]		11 January 2022
Summary CV for supervisor (student research) [CV of J Kirkwood]		09 May 2022
Summary, synopsis or diagram (flowchart) of protocol in non technical language [Study flow diagram v2]	2	06 May 2022

Statement of compliance

The Committee is constituted in accordance with the Governance Arrangements for Research Ethics Committees and complies fully with the Standard Operating Procedures for Research Ethics Committees in the UK.

User Feedback

The Health Research Authority is continually striving to provide a high quality service to all applicants and sponsors. You are invited to give your view of the service you have received and the application procedure. If you wish to make your views known please use the feedback form available on the HRA website: http://www.hra.nhs.uk/about-the-hra/governance/quality-assurance/

HRA Learning

We are pleased to welcome researchers and research staff to our HRA Learning Events and online learning opportunities– see details at: https://www.hra.nhs.uk/planning-and-improving-research/learning/

IRAS project ID: 310656	**Please quote this number on all correspondence**

With the Committee's best wishes for the success of this project.

Yours sincerely

Rebecca Throup
Approvals Specialist
PP

Dr George Gkimpas
Chair

Email: gmcentral.rec@hra.nhs.uk

Enclosures: "After ethical review – guidance for researchers"
Non CTIMP Standard Conditions of Approval]

Copy to: Mr Chris Coner

Lead Nation Scotland: gram.nrspcc@nhs.scot

Appendix K *Ethical approval letter: Gestational Diabetes Data for Research Database (23/NW/0262)*

NHS
Health Research Authority

North West - Haydock Research Ethics Committee
2 Redman Place
Stratford
London
E20 1JQ

18 October 2023

Ms Jasmine Kirkwood
University Of Edinburgh
The Queens Medical Research Institute, 47 Little France Crescent
Edinburgh
EH164TJ

Dear Ms Kirkwood

Title of the Research Database:	**Gestational Diabetes Data for Research Database.**
REC reference:	**23/NW/0262**
IRAS project ID:	**333082**

Thank you for your letter of 16th October 2023, responding to the Committee's request for further information on the above research database and submitting revised documentation.

The further information has been considered on behalf of the Committee by the Chair.

Confirmation of ethical opinion

On behalf of the Committee, I am pleased to confirm a favourable ethical opinion of the above research database on the basis described in the application form and supporting documentation as revised.

Publication of Your Research Summary

We will publish your research summary for the above study on the research summaries section of our website, together with your contact details, no earlier than three months from the date of this favourable opinion letter.

Should you wish to provide a substitute contact point, make a request to defer, or require further information, please visit:
https://www.hra.nhs.uk/planning-and-improving-research/application-summaries/research-summaries/

N.B. If your study is related to COVID-19 we will aim to publish your research summary within 3 days rather than three months.

During this public health emergency, it is vital that everyone can promptly identify all relevant research related to COVID-19 that is taking place globally. If you haven't already done so, please register your study on a public registry as soon as possible and provide the HRA with the registration detail, which will be posted alongside other information relating to your project. We are also asking sponsors not to request deferral of publication of research summary for any projects relating to COVID-19. In addition, to facilitate finding and extracting studies related to COVID-19 from public databases, please enter the WHO official acronym for the coronavirus disease (COVID-19) in the full title of your study. Approved COVID-19 studies can be found at: https://www.hra.nhs.uk/covid-19-research/approved-covid-19-research/

Duration of ethical opinion

The favourable opinion is given for a period of five years from the date of this letter provided that you comply with the standard conditions of ethical approval for Research Databases set out in the attached document. You are advised to study the conditions carefully. The opinion may be renewed for a further period of up to five years on receipt of a fresh application. It is suggested that the fresh application is made 3-6 months before the 5 years expires, to ensure continuous approval for the research database.

Approved documents

The documents reviewed and approved at the meeting were:

Document	*Version*	*Date*
Other [List of data items v0]		10 August 2023
Other [CV of applicant]		10 August 2023
Other [CI CV RR]		16 August 2023
Other [DW_DOI]	v0	27 September 2023
Other [Provisional opinion response]	v0	13 October 2023
Protocol for management of the database [Protocol GDM data v2]	2.0	16 August 2023
REC Application Form [RD_Form_25082023]		25 August 2023
Summary of research programme(s) [Summary of research]		10 August 2023
Summary of research programme(s) [summary of research program v0]		10 August 2023

Research governance

Under the UK Policy Framework for Health and Social Care Research, there is no requirement for NHS research permission for the establishment of research databases in the NHS. Applications to NHS R&D offices through IRAS are not required as all NHS organisations are expected to have included management review in the process of establishing the database.

Research permission is also not required by collaborators at data collection centres (DCCs) who provide data under the terms of a supply agreement between the organisation and the database. DCCs are not research sites for the purposes of the RGF.

Database managers are advised to provide R&D offices at all DCCs with a copy of the REC application for information, together with a copy of the favourable opinion letter when available.

All DCCs should be listed in Part C of the REC application.

NHS researchers undertaking specific research projects using data supplied by a database must apply for permission to R&D offices at all organisations where the research is conducted, whether or not the database has ethical approval.

Statement of compliance

The Committee is constituted in accordance with the Governance Arrangements for Research Ethics Committees and complies fully with the Standard Operating Procedures for Research Ethics Committees in the UK.

After ethical review

Reporting requirements

The attached standard conditions give detailed guidance on reporting requirements for research databases with a favourable opinion, including:

- Notifying substantial amendments
- Submitting Annual Progress reports

The HRA website also provides guidance on these topics, which is updated in the light of changes in reporting requirements or procedures.

User Feedback

The Health Research Authority is continually striving to provide a high quality service to all applicants and sponsors. You are invited to give your view of the service you have received and the application procedure. If you wish to make your views known please use the feedback form available on the HRA website:
http://www.hra.nhs.uk/about-the-hra/governance/quality-assurance/

HRA Learning

We are pleased to welcome researchers and research staff to our HRA Learning Events and online learning opportunities– see details at:
https://www.hra.nhs.uk/planning-and-improving-research/learning/

IRAS project ID: 333082	Please quote this number on all correspondence

Yours sincerely

p.p.

Mr Stephen Edgar
Chair

E-mail: haydock.rec@hra.nhs.uk

Enclosures: (RD) Conditions of Approval

Research Database – Conditions of Approval]

Copy to: Mr Charlie Mayor University of Edinburgh