

Chapter 10

Remote Monitoring to Improve Gestational Diabetes Care

By Margherita Grossi and Brian Pickering

10.1 Introduction

Rates of Gestational Diabetes Mellitus (GDM) have been rising worldwide over the past number of decades, and Ireland is no exception to this trend. The rising prevalence is a reflection of higher rates of obesity and advancing maternal age in pregnancy.

Epidemiologic studies quote rates of GDM in Ireland and the UK as 8%–24%, varying with study populations and diagnostic criteria used [1]. In our own unit, annual rates of GDM in 2015, 2016, and 2017 were 775, 975, and 974, respectively. This represents a significant increase in 10 years when compared with 2005, 2006, and 2007, when annual rates of GDM were 157, 133, and 153, respectively. The introduction of The International Association of Diabetes and Pregnancy Study Groups (IADPSG) thresholds for diagnosis of GDM in our unit in 2014 has contributed to this significant rise. However, other large contributors are most notably, advancing maternal age and rising rates of obesity in an increasingly more complex obstetric population.

The association between GDM, poor glycemic control, and adverse perinatal outcomes has been established many years ago [2]. Babies born to diabetic mothers have a higher incidence of macrosomia, [3] increased operative delivery rates, [4] increased rates of birth complications, and higher admission rates to the NICU to

correct metabolic imbalances [5]. The rising prevalence of GDM also represents a major public health concern as affected mothers are at increased risk of type 2 diabetes in later life [6]. In this chapter, we introduce a novel remote management pathway for diet-controlled gestational diabetes. The concept of an app-assisted lifestyle and blood sugar level monitoring program fosters patient-centered care. This surveillance approach has potential health and economic benefits for both the patient and the overall hospital infrastructure through a reduction in clinic waiting times, hospital attendances, administrative duties, and staff requirements. Ultimately, our aim is to move forward with a patient-oriented model of care for women with GDM. Patients who require prompt intervention beyond dietary and lifestyle changes to optimize their Blood Sugar Level (BSL) can be easily identified. Another anticipated benefit of this strategy in the long term will be reduced healthcare burden and cost. The application of technology in healthcare is a rapidly growing field that can be exploited to achieve and maintain positive healthcare behaviors.

10.2 Methods

10.2.1 Ethics and Privacy Procedures

Prior to the initiation of the study reported here, all necessary documentation to ensure full compliance with both EU and National Irish regulations in terms of data privacy and ethical aspects was fulfilled and met. There was particular attention paid to the GDPR requirements and security measures to avoid any ethical issues or data breaches. Ethical approval to proceed with the study was obtained from the Rotunda Hospital Research Ethics Committee.

As to the classification of the app, the Health Products Regulatory Authority (HPRA) oversees the regulation of medicines and devices used by the public in Ireland. We approached the HPRA for guidance on the classification of our proposed app and portal system to ensure compliance with national guidelines. Inclusion of certain functionalities within the system would have resulted in a requirement for the app to be classified as a medical device.

In our app development phase, we were mindful not to include any functionalities that would change the classification of the app to that of a medical device, as this would have resulted in lengthy delays in the project. The red flag monitoring system we developed within the hospital portal is buttressed by a safety mechanism in that all app-users are also required to report their actual glycemic indices during their scheduled virtual clinic reviews.

As to the consent procedure in the clinical phase, all GDM patients from Rotunda Hospital willing to participate in the study were asked to read our patient

information leaflet pertaining to the background and main objectives of the study, the expected benefits of participating, the right to withdraw at any time, the study's methodology, and our assurance of confidentiality. Following review of this document, interested parties were asked to sign our informed consent form.

Other related documents produced were a Data Protection Impact Assessment (DPIA), an internal validation document and risk assessment to evaluate the security by design approach adopted for the whole architecture and data flows, a data privacy policy for the app, the application for the clinical investigation (not needed in the final stage), and all the necessary GDPR agreements between the study participants.

Relevant standards/regulation of reference are:

- EU 93/42/EEC regulation on medical devices
- EN/IEC 62304: Medical device software – software life cycle processes
- ISO 14971: Application of risk management to medical devices
- ISO 13485:2016 “Medical devices – Quality management systems – Requirements for regulatory purposes”

10.2.2 Diagnostic Model

Remote monitoring of women with GDM holds the potential for decreasing pregnancy complications, improving patient quality of life, enhancing the efficiency of healthcare delivery, and reducing healthcare costs. In this paragraph, we present a simple, fast, and flexible method based on a fuzzy inference system for assessing risk levels given glucose readings from patients.

Modern machine learning and artificial intelligence algorithms are used as a core element of Decision Support Systems (DSS), which are typically designed to integrate a medical knowledge base, patient data, and an inference engine to extract insights and generate personalized recommendations. Current DSS exploit embedded expert knowledge through expert system instantiations that eliminate the uncertainty and imprecision associated with the diagnosis of gestational diabetes, e.g., using fuzzy modeling and fuzzy inference [7, 8]. Typical implementations of such systems suffer from the lack of explainability and the long training time that is triggered whenever new patient data become available. Our approach, on the other hand, addresses real-time continuous glucose monitoring in gestational diabetes with a lightweight inference system (no training) that offers timely and interpretable output.

The fuzzy inference system consists of four generic parts (Figure 10.1): the fuzzification interface, which converts the input data into the internal format; the knowledge- and decision-based unit with rule evaluation; and, symmetrically, the defuzzification interface, which provides the estimated risks. In order to develop

to the clinician assessment, which focuses on the percentage of values above thresholds (for more information, please refer to [9]).

10.3 Results

10.3.1 Initial Evaluation

To prove the capability of the system, an evaluation was performed by simulating the model assessment for a larger artificially generated data set (obtained through data augmentation) and comparing the risk levels obtained by the medical criteria. The algorithm was able to achieve 94.68% accuracy compared to current monitoring methods, which rely on periodic face-to-face physician reviews. In particular, it was observed that the system indeed assigns higher risk values to patients having groups of weekly measurements with a very high glucose mean and standard deviation and, conversely, lower risk values to patients with a very low glucose mean. For the border cases, however, our algorithm is able to more accurately evaluate risks compared to hard thresholds. The risk system adapts better to each observation, giving, for example, a greater risk to patients with higher glucose levels on average, even if the number of values above threshold is below 30%, and vice-versa, a lower risk to patients with glucose levels only marginally above threshold.

To summarize, the advantages of using simple fuzzy thresholds include:

- simplicity: the system can be deployed without waiting for the collection of a larger amount of data, it does not require training, and the inference is very fast;
- interpretability: the system is explainable by design since the rules are evaluated and can be easily understood and trusted by the clinicians; and
- accuracy and personalization: the usage of fuzzy logic releases the necessity to have a fixed threshold for risk (different hospitals have slightly different thresholds) and mimics the real assessment of a clinician, which would assign different levels of risk whether the glucose measurement of a patient is slightly above the threshold or extremely high (different membership values in the fuzzy system). The system obtains comparable results in risk assessment for extreme cases (low and high) and can achieve personalized diagnosis in border cases.

10.3.2 Prognostic Model

Besides assisting healthcare professionals with a novel diagnostic model, in the last phases of the study, a new prognostic model was developed with the scope

of predicting the risk of abnormal fasting and postprandial glucose levels in the upcoming period of time. Since the analysis is mainly focused on the long-term monitoring of the patients, similar to what has been done for the fuzzy system, patients who show more than 30% of measurements above a given threshold (5.1 mmol/L for preprandial and 7.8 mmol/L for postprandial) within the next 7 days are classified as having a high risk.

Instead of using only the glucose levels like the previous fuzzy inference model, the prognostic model evaluates the risks taking into account personalized patient information, including features such as Body Mass Index (BMI), age, glucose tolerance test measurements, and parity, among others. One key challenge was to interpret, standardize, and then clean some of the values with the help of the medical team. Median imputation is used under the assumption that missing data were random, while features that presented too low variance in their distribution (e.g., ethnicity) were discarded since they would introduce a bias in the results and mislead some of the conclusions. Moreover, a major source of concern was the large number of unknown meal types in the glucose time series collected from the app (up to 40% of the total number of measurements). Since the normal glucose level ranges depend strongly on whether the values represent outcomes from a pre- or postprandial measurement, a correct assignment of those values is essential for the classification of risk.

After agreement with the medical experts, a pre-processing of the data was introduced, allowing the medical personnel to distinguish the patients who had overly high fasting levels and should be monitored further. With this data imputation, the total number of measurements above the normal threshold increased by about 16%, compatibly with the preference for higher recall values for this application. Finally, the quality of the data collected from the app is strongly affected by patient compliance. Sparse reporting or short time series should not be used for the model training and testing (at least 7 days of available measurements).

Before developing the classification model, a survival function for the time to the risk event was built from the data collected by the app up to August 2021 and from the historical data collected in 2019 using a Kaplan-Meier estimator. Comparing the two curves showed that the differences between the two populations were not significant (p-value of log-rank testⁱ: 0.46). Based on this analysis, the two data sets were merged in order to increase the number of available patients. Moreover, instead of considering each patient time series as a single training data point, a rolling window approach was adopted. Shifting a cut-out window over the data to create smaller time series from each user (from the beginning until a given

i. The log-rank is a non-parametric hypothesis test to compare the survival distributions of two samples.

point in time), it was possible to create a training data set of 16,900 data points. The rolling mechanism emulates the streaming of data coming from the app and was also applied to derive further features by aggregating the glucose measurements from these smaller time series and to investigate whether they could be good indicators of risk (e.g., in [10], the complexity of the time series is related to the risk of developing type 2 diabetes). The Python package “tsfresh” was used to extract the most descriptive characteristics of the glucose time series [11].

A grid search over a number of tree-based machine learning classifiers was performed to select the best model and its hyperparameters. In order to avoid overfitting, we used a k-fold cross validation ($k = 4$) on the training data, while the test data were obtained via stratified splitting of the users. Finally, a decision tree classifier was chosen as a predictive model because, while having performance comparable with other most sophisticated algorithms, it is easy to explain (the so called white box model) and robust against missing data and heterogeneous feature types. With respect to the model performance, we set the F1 score as the optimization metric and obtained a total accuracy of 97%, with 78% recall and 90% precision on the test data with the default probabilities threshold.

In order to analyze the reliability of the model, besides looking at its overall performance, different methods to interpret the results were applied, from a simple Pearson correlation coefficient to the more sophisticated Shapley Additive Explanation. As expected, it was found that a higher BMI has a very negative impact on the gestational diabetes risk, as well as a very high mean or median glucose value. The prognostic model results, together with the explanation of the risk factors contributing to the single measurement predictions, have been summarized in a dashboard for internal usage (Figure 10.2).

10.3.3 Evaluation Within the Clinical Setting

A virtual GDM clinic service was established to reduce the footfall in our outpatient departments as a measure of reducing the transmission and acquisition of COVID-19. The virtual clinic facilitated the provision of clinical care through telephonic review while simultaneously reducing the requirement for hospital attendance in a cohort of women with gestational diabetes by 90%. This had a positive impact on the “did not attend” rate and has resulted in a reduction of costs for the patient (absenteeism from work, transport and parking costs, and childcare costs).

While the hospital supplies 6 weeks’ worth of testing strips for all GDM patients, cost savings are envisaged with our new remote model through reduction of unit costs for all administrative staff, for the canteen where complimentary breakfast was previously provided, and for phlebotomy and laboratory staff who previously

Table 10.1. Characteristics of GDM patients based on treatment type (n ± SD).

Characteristic	Diet + Exercise (N = 180)	Metformin (N = 34)	Insulin (N = 67)	P-value
Maternal age (years)	32.0 ± 5.0	34.0 ± 6.0	33.0 ± 5.0	0.282
Maternal weight at booking (kg)	82.0 ± 17.0	86.0 ± 20.0	90.0 ± 22.0	0.015
GA at 3rd trimester scan	33.9 ± 1.7	33.9 ± 1.6	33.1 ± 1.6	0.002
AC centile	58 [30,81]	59 [32,87]	72 [39,95]	0.040
EFW centile	52 [33,74]	54 [39,73]	62 [45,84]	0.051
Gestational weight gain (kg)	6.2 ± 5.0	5.6 ± 4.7	6.2 ± 6.3	0.841
Gestational age at delivery (week)	39.1 ± 2.2	38.9 ± 1.0	38.5 ± 1.3	0.115
Birthweight (g)	3423 ± 603	3389 ± 435	3579 ± 588	0.136
Neonatal hypoglycemia	43 (24%)	18 (30%)	5 (15%)	0.271
NICU admission	28 (16%)	14 (23%)	2 (6%)	0.001

lifestyle, metformin monotherapy, insulin monotherapy, and combined insulin and metformin therapy. Baseline maternal characteristics and an Oral Glucose Tolerance Test (OGTT) glycemia indices at 0, 60, and 120 min were examined. Analysis of variance was used to compare groups and sensitivity to normality was assessed with non-parametric analysis. The latter was reported with the medians and the InterQuartile Range (IQR). Chi-square tests were used to compare categorical outcomes. Statistical significance was assumed for p-values <0.05. No adjustments for multiple tests were performed. A total of 197 women met the criteria for inclusion in the study. Appropriate glycemia control was attained by 73% (n = 144) of those managed with diet and lifestyle interventions alone. Insulin monotherapy was required in 13.7% (n = 27), metformin monotherapy was required in 10.15% (n = 20), and 6 (3.04%) required combined treatment with metformin and insulin. Maternal characteristics between the treatment groups are outlined in Table 10.1. Fasting glucose levels at OGTT were significantly associated with the subsequent requirement for insulin, either as a single agent or in combination with metformin (Table 10.2). Comparing the diet to the insulin group, a ROC analysis using the maximum Youden index suggests the optimal cut-off for fasting OGTT to be 97.2 g/dL (5.4 mmol/L), corresponding to a sensitivity of 81% and a specificity of 57% for insulin use. Elevated fasting glucose is a useful predictor of pregnancy in the GDM population, where insulin supplementation may subsequently be warranted and therefore considered at an earlier gestation.

Table 10.2. Comparison of GGT1, GTT2, and GTT3 between groups, Note: Mean $\pm$ SD or median [IQR] are presented for number characteristics. P-value is for any difference between the three groups using the chi-square test or Wilcoxon rank-sum test.

Group	Comparison Group	OGTT1 Time: 0 (fasting)		OGTT2 Time: after 60 min		OGTT3 Time: after 120 min	
		Difference		Difference		Difference	
		(95% CI)	p	(95% CI)	p	(95% CI)	p
Diet + Exercise	Insulin	−0.40 (−0.57, −0.24)	<0.001	−0.76 (−1.39, −0.14)	0.017	−1.29 (−2.78, 0.20)	0.089
Diet + Exercise	Metformin	0.20 (−0.40, 0.01)	0.064	−0.20 (−0.97, 0.58)	0.619	−0.48 (−1.27, 0.31)	0.236
Insulin	Metformin	0.21 (−0.04, 0.45)	0.097	0.57 (−0.35, 1.49)	0.224	0.82 (−0.81, 2.45)	0.324

The patient outcomes examined cover antenatal, delivery and postnatal areas of maternal and neonatal wellbeing. Mode of labor onset, subsequent mode of delivery, and gestational age at delivery were also collected. The birth weight was collected, and from this rate of macrosomia, a known complication of uncontrolled GDM, it could be calculated. Other surrogate markers of suboptimal glycemia control are neonatal hypoglycemia and hyperbilirubinemia (represented by jaundice), and these were also assessed. Any admission to the Neonatal Intensive Care Unit (NICU) was also documented to allow calculation of the overall NICU admission rate.

These parameters were then compared to historical data to demonstrate the non-inferiority of app-assisted care delivery compared with hospital-based care. We previously compared a telemedical approach to GDM follow-up with in-person hospital attendance through the analysis of 34,399 data points obtained from 283 patients attending either Hospital-based Clinics (HC) or Virtual Clinics (VC). The overall distribution of glucose levels was similar in both groups. The VC appeared to have greater sensitivity for detecting high blood glucose levels at lower thresholds. However, at a Fasting threshold (F) of 95 mg/dL and a PostPrandial threshold (PP) of 140 mg/dL, the results in both clinics were equivalent – 80% (F) and 96% (PP). The median birthweights were 3,350 g and 3,452 g for HC and VC, respectively, a difference that did not reach statistical significance ($p=0.241$). Birthweights > 4 kg were more frequent in the VC group (15%) than the group receiving hospital-based GDM surveillance (10%), but this trend did not reach statistical significance ($p = 0.322$). An abnormal fasting glucose was associated with Operative Vaginal Delivery (OVD) ($p = 0.032$) in the VC group, but not in the HC group. Aberrations in postprandial control were less likely to be associated with macrosomia (body weight > 4 kg) and OVD or C-section than aberrations in fasting control.

Admissions to the NICU were higher in the HC group (20%) compared to the VC group (14%), but this trend did not reach statistical significance ($p = 0.288$).

We also devised a patient satisfaction survey exploring feasibility, functionality, and utility of the remote monitoring solution – GDMapp – that has been developed in collaboration with all three project partners. While the responses to this survey will not be analyzed until all participants have had the opportunity to complete it, a general overview of the user experience from weekly phone calls with recruited participants has been positive. In general, we have a 90% recruitment rate with most eligible candidates motivated to trial this new eHealth initiative. Out of 115 recruited participants, only two (1.7%) have withdrawn from the study – both due to self-reported increased stress levels attributable to the diagnosis of GDM.

10.4 Discussion

This project created a clinical decision support tool that enables self-management and remote monitoring of GDM, in the form of a patient-facing smartphone application (GDMapp) linked to a medical web portal for use by the obstetric diabetes team. In addition, a diagnostic model based on fuzzy inference systems was developed in the initial stage of the study to support healthcare personnel in monitoring GDM by assigning risk scores based on the glucose levels collected from the app. In the last stage of the project, the model evolved into a prognostic model based on a variety of data, enabling the prediction of those patients whose likelihood of having levels above the established threshold is high.

By study completion, 150 women had contributed to the design of the portal in the pre-study stages of the study, with a further 200 enrolling in app-assisted care using GDMapp, giving scope for validation of the product.

The outcomes of the study on GDM may be summarized as follows:

- Enhancement of patient engagement and education of GDM.
- Promotion of patient-centered care.
- Optimization of compliance with GDM management strategies.
- Elimination of the need for the majority of women with GDM to attend additional hospital appointments.
- Establishment of a telemedical service and incorporation of information and communication technology to complement care provision for obstetric patients with GDM.
- Glycemic data generated from this study will facilitate scalability of the risk assessment algorithm created by Huawei leading to the generation of a clinical

decision-making tool, further enhancing the remote capabilities of this AI solution for gestational diabetes care.

- The introduction of a fully functional solution in the standard GDM care has a positive impact compared to previous practices in terms of: accuracy and quality of data, reduction in the number of admin duties, and staff hours at the hospital.
- Overall, there is a positive tendency in patients in using the mobile app rather than continuous visits to the hospital, particularly in during the COVID-19 epidemic.

At the same time, there were various challenges in collecting and exploiting the data during the study.

- Cleaning, formatting, interpreting, and analyzing real medical data sets from historical data were shown to be a difficult task, due to inconsistencies in the information provided, missing information in some relevant parameters and data privacy issues due to personal information.
- Depending on the granularity and the information contained in the historical data, pre-processing was necessary for integrating those information with the one provided from the app and using them in the training of the machine learning model. This issue was encountered in other studies as well.
- The reliability of the user data collected from the app strongly depends on the correct usage of the device from the end user.
 - In particular, the meal type/timing information, which should have been provided by the patients when taking each measurement, was not done in 40% of the cases (less than 10% of the patients always set the meal type).
 - Postprocessing of the data retrieved from the app in order to impute the missing information has been necessary in order to evaluate the risk, since the threshold is different for post and preprandial glucose measurements.
- Other common challenges for using the app data include:
 - Syntax analysis of unstructured or not standardized input data (especially from the user personalized record) and sparse time series (only 60 patients had at least 1 week of glucose measurements).
 - In a consistent number of cases, moreover, the measurements were regularly repeated twice, so that the frequency of daily measurements increased from 4 to 8.
- Unfortunately, AI applications require many medical sensor datasets that are rarely available and lack diversity. It was not possible to use some of the more complex machine learning models because of the patient statistics and some

of the features (e.g., ethnicity) since they lacked the diversity necessary for generalization.

- It required large data sets to develop robust analytical/prediction model to obtain relevant and accurate outcomes to support the medical staff in their decision-making process.
- Challenges in acquiring a relevant volume of data via the app due to both technical and legal aspects:
 - Extremely lengthy process in obtaining the approval from the Apple Store caused a significant delay in facilitating the last version of the mobile app, as most of the users used an iPhone.
 - Moreover, it was required to put additional efforts to provide a mobile app iOS version fully compliant with Apple requirements, which resulted in additional efforts to change the code in the Android version to obtain an exact mirror of the iOS version.
 - Lack of data from recruited patients due to pairing/connectivity problems caused by a number of unexpected events: a cyber-attack at the hospital that caused failure of critical IT systems at the hospital and network disruptions, making data unavailable.
 - Challenges in collecting a larger amount of personal data from wearable sensors due to some constraints introduced by used protocols (Bluetooth), the main instable connection. In addition, the transfer of data from the edge (e.g., a mobile phone) to the server can be disturbed, requiring flexible methods for checking the consistency of the data transfer.

10.5 Conclusion

In this chapter, we presented the results of exploiting the use of data from patient self-monitoring at home for the prevention and treatment of GDM and its complications. Although the initial aim was purely to demonstrate the feasibility in facilitating appropriate care for this cohort (mothers with a diagnosis of diabetes during pregnancy), which might help relieve the burden on existing hospital provision, the COVID-19 pandemic created an unexpected benefit. Not only is it possible to develop and deploy a monitoring-based patient care process, but this can also support unexpected challenges such as the need for patients to avoid hospital visits unless absolutely necessary. Given the increased incidence of GDM, we believe that our experience can be extended to other facilities to enable and support the effective remote treatment of GDM mothers during pregnancy.

References

- [1] Farrar D, Simmonds M, Bryant M, Sheldon TA, Tuffnell D, Golder S, Dunne F, Lawlor DA. Hyperglycaemia and risk of adverse perinatal outcomes: systematic review and meta-analysis. *BMJ*. 2016. 354.
- [2] Billionnet C, Mitanchez D, Weill A, *et al.* Gestational diabetes and adverse perinatal outcomes from 716,152 births in France in 2012. *Diabetologia*. 2017. 60: 636–644. <https://doi.org/10.1007/s00125-017-4206-6>.
- [3] Kamana KC, Shakya S, Zhang H. Gestational diabetes mellitus and macrosomia: a literature review. *Annals of Nutrition and Metabolism*. 2015. 66 Suppl 2: 14–20. DOI: 10.1159/000371628.
- [4] Zeki R, Wang AY, Lui K, Li Z, Oats JJN, Homer CSE, Sullivan EA. Neonatal outcomes of live-born term singletons in vertex presentation born to mothers with diabetes during pregnancy by mode of birth: a New South Wales population-based retrospective cohort study. *BMJ Paediatr Open*. January 2018. 30;2(1): e000224. DOI: 10.1136/bmjpo-2017-000224.
- [5] Al-Khalifah R, Al-Subaihin A, Al-Kharfi T, Al-Alaiyan S, Al-Faleh KM. Neonatal short-term outcomes of gestational diabetes mellitus in Saudi mothers: a retrospective cohort study. *Journal of Clinical Neonatology*. 2012. 1(1): 29.
- [6] Bellamy L, Casas J-P, Hingorani AD, Williams D. Type 2 diabetes mellitus after gestational diabetes: a systematic review and meta-analysis. *The Lancet*. 2009. 373(9677): 1773–1779.
- [7] Okpor MD. Prognostic diagnosis of gestational diabetes utilizing fuzzy classifier. *International Journal of Computer Science and Network Security*. 2015. 15.6: 44–48.
- [8] Übeyli ED. Automatic diagnosis of diabetes using adaptive neuro-fuzzy inference systems. *Expert Systems*. 2010. 27.4: 259–266. DOI: 10.1111/j.1468-0394.2010.00527.x.
- [9] Sánchez CS, *et al.* Fuzzy inference system for risk evaluation in gestational diabetes mellitus. *IEEE 19th International Conference on Bioinformatics and Bioengineering (BIBE)*. 2019.
- [10] Rodriguez de Castro C, *et al.* Glucose time series complexity as a predictor of type 2 diabetes. *Diabetes/metabolism research and reviews*. 2017. 33.2: e2831.
- [11] Christ M, Braun N, Neuffer J, Kempa-Liehr AW. Time series Feature extraction on basis of scalable hypothesis tests (tsfresh A Python package). *Neurocomputing*. 2018. 307: 72–77.