Predicting perinatal health outcomes using smartphone-based digital phenotyping and machine learning in a prospective Swedish cohort (Mom2B): study protocol

Ayesha M Bilal,1,2 Emma Fransson,3,4 Emma Bränn,3 Allison Eriksson,2,3 Mengyu Zhong,2,5 Karin Gidén,3 Ulf Elofsson,1,2 Cathrine Axfors,3 Alkistis Skalkidou,3 Fotios C Papadopoulos1

ABSTRACT

Introduction Perinatal complications, such as perinatal depression and preterm birth, are major causes of morbidity and mortality for the mother and the child. Prediction of high risk can allow for early delivery of existing interventions for prevention. This ongoing study aims to use digital phenotyping data from the Mom2B smartphone application to develop models to predict women at high risk for mental and somatic complications.

Methods and analysis All Swedish-speaking women over 18 years, who are either pregnant or within 3 months postpartum are eligible to participate by downloading the Mom2B smartphone app. We aim to recruit at least 5000 participants with completed outcome measures. Throughout the pregnancy and within the first year postpartum, both active and passive data are collected via the app in an effort to establish a participant's digital phenotype. Active data collection consists of surveys related to participant background information, mental and physical health, lifestyle, and social circumstances, as well as voice recordings. Participants' general smartphone activity, geographical movement patterns, social media activity and cognitive patterns can be estimated through passive data collection from smartphone sensors and activity logs. The outcomes will be measured using surveys, such as the Edinburgh Postnatal Depression Scale, and through linkage to national registers, from where information on registered clinical diagnoses and received care, including prescribed medication, can be obtained. Advanced machine learning and deep learning techniques will be applied to these multimodal data in order to develop accurate algorithms for the prediction of perinatal depression and preterm birth. In this way, earlier intervention may be possible.

Ethics and dissemination Ethical approval has been obtained from the Swedish Ethical Review Authority (dnr: 2019/01170, with amendments), and the project fully fulfils the General Data Protection Regulation (GDPR) requirements. All participants provide consent to participate and can withdraw their participation at any time. Results from this project will be disseminated in international peer-reviewed journals and presented in relevant conferences.

INTRODUCTION

Optimal maternal health is important throughout pregnancy, childbirth and the postpartum period to ensure the full potential for the mother, infant and family to get a good start.1 Two health conditions that are important to address in order to reach a goal of good maternal health are perinatal depression (PND) and preterm birth (PTB), both affecting about every 10th pregnancy worldwide.23

Perinatal depression

PND is an episode of major depression with onset anytime during pregnancy and up to 4 weeks postpartum,3 although in research settings, a period of up to 1 year postpartum is often considered.1 Antenatal depression affects between 7% and 13% of pregnant women,6 and postpartum depression (PPD) is estimated to affect between 10% and 20% of all newly delivered mothers,2
while many women experience persistent depression throughout the perinatal period. PND is distinct from ‘baby blues’, which are commonly experienced symptoms of low mood and anxiety that subside within 2 weeks postpartum. PND is both emotionally and physically debilitating like major depression, with additional risks related to the pregnancy and birth, such as PTB, low birth weight, pre-eclampsia and placental abnormalities. Moreover, it is associated with retained maternal weight postpartum, decreased breast feeding, poor maternal sleep, and poor perinatal quality of life. PND can also compromise the critical mother–infant bond, as it affects the mother’s caregiving abilities and adaptation to the maternal role, and has a long-term impact on the child’s cognitive, emotional and behavioural development. Furthermore, PND can be characterised by the occurrence of self-harm thoughts, which are linked to long-term somatic and psychiatric morbidity, and increased maternal mortality from suicide in the first year postpartum.

The aetiology of PND is multifactorial, including biological, genetic, psychological and social factors, such as stressful life events, social support, domestic violence, childhood adversity, history of depression and anxiety, low self-esteem and even personality traits like resilience. Despite this knowledge, detecting PND remains a challenge for the healthcare system, with one review finding that around 30%–70% of cases go undetected and only 15% receive adequate treatment.

Current screening protocols include the Edinburgh Postnatal Depression Scale (EPDS) during postpartum visits to assess risk of PND. However, early detection of PND has remained challenging for many reasons, including inconsistencies in screening practices, and failure to distinguish depressive symptoms due to their overlap with typical somatic experiences in the early postpartum period. Furthermore, women may hesitate to seek care possibly because of the depression itself, but also stigma and fear of being judged as an imperfect mother, as well as concerns about antidepressant use during pregnancy and breast feeding. Besides, in-clinic screening frequently relies on retrospective self-reports of diagnostically relevant information, making it susceptible to errors and biases associated with autobiographical recollection.

Unquestionably, more efficient and effective methods for predicting PND in mothers are needed to enable early identification and intervention, thus improving prognosis, and reducing the burden of disease. Previous studies have attempted to develop predictive models of maternal depression primarily focus on the postpartum period only. Few have used social media fingerprints or biomarkers in their models, and these studies largely depended on psychometric self-reports and limited modalities. These drawbacks compromise the predictive power of the models and illustrate why multivariate, real-time and unobtrusive approaches to data collection and symptom monitoring must be encouraged to develop better predictive models.

Preterm birth

Among somatic pregnancy complications, PTB is a major cause of neonatal death, as well as of poor long-term health in children, affecting approximately 15 million babies worldwide each year. In Sweden, the PTB rate is about 6%, which is a relatively low number compared with the international average of over 10%. Like PND, the aetiology of PTB is multifactorial, including previous PTB, multifetal pregnancy, cervical insufficiency, intrauterine infections, vaginal bleeding in the second trimester, in-vitro fertilisation, primiparity, as well as maternal antenatal stress and depression. In fact, many risk factors overlap between PND and PTB, such as childhood traumatic events or maltreatment, stressful life events, being single or lacking social support, being overweight, smoking and low socioeconomic status. Inflammation has been suggested as a possible underlying pathway for both depression and preterm delivery.

There are evidence-based interventions for preventing or delaying PTB to optimise birth outcomes, such as smoking cessation, progesterone therapy, cerclage in women with cervical insufficiency or antibiotics. However, a major obstacle for the success of these interventions is the aetiological heterogeneity of PTB, which makes it extremely challenging to identify women at high risk. In fact, two-thirds of women who experience PTB do not present with any risk factors at all. Available biological diagnostic tests for PTB (such as fetal fibronectin) lack sufficient positive prediction values. Screening for cervical length is performed in Sweden for women with a history of PTB; however, this is not helpful in primiparous women.

It can be concluded that no single biomarker is sufficient for prediction; multimodal data, including psychosocial and behavioural factors, should, therefore, be the focus of prediction efforts.

Digital phenotyping and big data

Digital devices like smartphones allow us to capture moment-by-moment, objective data regarding the patient’s experiences and functions in non-clinical settings. This process, known as digital phenotyping, allows us to collect two kinds of data: active data and passive data. Active data refer to data that require user input, such as surveys and voice recordings. Passive data refer to automatically collected data from smartphone sensors and activity logs, which can be used to infer the user’s mobility and sleep patterns, digital social activity, smartphone usage patterns, and even affective and cognitive changes. The Mom2B smartphone app is developed using the Beiwe research platform (www.beiwe.org) from the Harvard School of Public Health. It can allow us to capture digital phenotyping data during the perinatal period with greater efficiency and temporal sensitivity as data collection occurs continuously and in real-world contexts, which minimises the risk of recall biases. Such apps could also be integrated into the mother’s perinatal care plan. One drawback of smartphone-based data
collection, in general, is that attrition rate increases with longer follow-up times; however, this can at least in part be compensated for by the continuous collection of passive data.

In fact, one of the biggest advantages of smartphone-based digital phenotyping is the ability to collect multi-variate, high-volume data, known commonly as big data. Big data are excellent for healthcare research since it can facilitate a unique insight into risk factors and the development of better diagnostic frameworks, however, the literature on big data approaches for psychiatric conditions, particularly perinatal mental health, is limited. Nordic countries are in the forefront in this respect—with the Danish National Birth Cohort, the Norwegian Mother, Father and Child Cohort (MoBa) and Autism Birth Cohort (ABC) studies, and the Swedish Biology, Affect, Stress, Imaging and Cognition (BASIC) cohort study—due to the availability of nationwide registers with comprehensive personal and medical information for all pregnant women in these countries. Nonetheless, register data, while valuable, lack the malleability, continuity and veracity offered by digital phenotyping. Furthermore, studies derived from these cohorts have largely relied on traditional statistical methods, which are limited in their ability to scale to large data sets and identify more subtle patterns in data.

To date, few studies have applied digital phenotyping for prediction of psychiatric conditions, such as relapse in schizophrenia and severity of mood episodes in bipolar disorder. In the context of PND, while smartphone apps are widely used, their application has been largely focused on screening and intervention. Only two studies have applied digital phenotyping for predicting PPD. While these studies have reported encouraging results, their predictive ability is compromised due to limited modalities (using only active data in the form of questionnaires), infrequent measurement points and usage of more traditional statistical methods. The Mom2B study combines nationwide health and pregnancy register data with active and passive data collected through smartphone-based digital phenotyping to objectively monitor indicators of PND in non-clinical contexts.

In order to harness the full potential of big data, more advanced analytical methods, such as machine learning (ML) and deep learning (DL), are ideal. ML is an artificial intelligence approach that refers to various methods of enabling an algorithm to identify and learn intricate patterns in data to predict outcomes. Modern ML methods, such as deep neural networks (DNNs), are uniquely suited to analysing big data sets as they can detect complex, high-dimensional interactions and structured information, without guidance, that can then be used to train predictive algorithms. DNN models are comprised of multiple ‘hidden’ processing layers, inspired by biological neural networks, consisting of a series of interconnected nodes that resemble neurons.

Over the last decade, there has been a steady increase in the use of DL methods in medicine. However, few studies have used ML for diagnosis or risk assessment in psychiatry, and those that do are often limited by modest sample sizes and modalities, or from using only traditional ML techniques. To our knowledge, Mom2B is the first study to adopt a big data approach and use multimodal digital phenotyping with advanced ML techniques to develop predictive algorithms for PND and PTB.

**Objectives**

Using large-scale, multimodal data collected through the Mom2B smartphone app, together with health and pregnancy information from national registers, the primary aim of this study is to assess the accuracy of advanced ML and DL methods in predicting development of PND (1) in the third pregnancy trimester, using data from the first trimester, and (2) during the early and late postpartum period, using data collected throughout pregnancy and childbirth.

A secondary aim of this study is to apply advanced ML and DL techniques using the multimodal data set to predict the risk of PTB.

**METHODS AND ANALYSIS**

**Cohort description**

Mom2B (www.mom2b.se) is a national ongoing smartphone app-based study; the app was launched at the end of November 2019 to App Store and Google Play. All Swedish-speaking women above the age of 18 owning a smartphone, who are either pregnant or within 3 months postpartum, are eligible to participate by registering and providing consent in the Mom2B app. Participating women are also asked for optional consent to be contacted for additional research studies within and from outside the Mom2B project (see online supplemental appendix A). Participant data are then linked to psychiatric and somatic health-related and pregnancy-related information available from Swedish national registers.

We aim to recruit at least 5000 participants with completed outcome measures. Due to the complexity of ML methods, it is not possible to perform any traditional test of statistical power. However, based on previous studies, and conferring with experts in artificial intelligence, we estimated that this approximate number would give us enough material to build robust prediction models while accounting for attrition and the prevalence rates of the outcomes.

Information about the study is being disseminated on social media, and through posters and brochures sent to primary and maternal care centres across the country. Figure 1 illustrates Mom2B recruitment, data collection and opt-outs. Table 1 outlines participant characteristics of the existing Mom2B cohort based on users who have contributed relevant data, along with similar characteristics in the general Swedish population.
The Mom2B app collects three types of data: survey data, audio recordings, as well as passive data. Data can be collected from the first week of pregnancy, and up until week 52 after birth. Only data that participants have consented for are collected from the time they register to the study, and they can change their consent preferences anytime in the app if they wish to stop.

**Surveys and questionnaires**

The Mom2B app delivers a range of both validated and self-developed questionnaires two to three times per week on average, with a mean of five questions per survey. These questionnaires are used to collect information regarding the participant’s mental and physical well-being, and history, personality, relationships, as well as perinatal and parenthood experiences. They include the EPDS, a 10-item self-report screening tool with good psychometric properties, used as the primary outcome measure in this study to assess depressive symptoms throughout the study period. A summary of the timeline of validated and self-developed instruments, along with the number of occurrences of the survey throughout the study period, can be found in figures 2 and 3, respectively.

**Voice recordings**

Voice acoustic qualities, such as pitch, speed, timing and timbre, have been used in previous research to successfully distinguish depressed from non-depressed individuals. To collect voice data, the Mom2B app sends out a voice recording task asking the participant to record reading simple texts, numerical sequences or vocalisations every 2–4 weeks.

**Passive data collection**

Passive data that the user has provided consent for are continuously collected via the Mom2B app throughout the study period, and are used to infer the user’s behavioural patterns. Some of these features are collected differently for iOS and Android users. The feature modalities are briefly explained below.

**Mobility**

Correlations have been demonstrated between a patient’s geographical movement patterns and changes in depressive symptoms. Sixty seconds of Global Positioning System (GPS) data are continuously collected every 10 min. Accelerator data are collected when the motion exceeds a certain threshold; motion activities, including being stationary, walking, running, cycling and movement in a vehicle, are recorded when the state changes. For iOS, we collect device motion to provide more detailed motion sensor data.

**Data usage**

Internet usage is the main feature of data usage. Various patterns of internet usage have been identified in relation to depression in different populations, but not among women with PND. The Mom2B app records the accumulated upload and download rates together with timestamps. Another feature, reachability, records time-stamped smartphone connectivity—whether the phone is connected to cellular network, Wi-Fi or neither. It also records a Wi-Fi log for Android phone, which includes anonymised Media Access Control (MAC) addresses’ frequencies and Received Signal Strength Indicator (RSSI) of available wireless networks in the area.

**Smartphone usage**

General smartphone use has been found to correlate with sleep quality and depression. Phone power state, combined with mobility parameters, can reflect individual needs and activities.
behaviour like sleep patterns. To keep track of the use of the smartphone device, data are collected on screen activity, charging status and device reboot.

**Social media activity**

Social media behaviour has also been proven useful in detecting mental states. It has also been shown that reduced social activity on Facebook predicted symptoms of PPD. Collected data consist of simple behavioural measures, such as posting, commenting or liking, together with their timestamps. Notably, we only measure activity levels, not information related to the text or image content of that activity, and participants are made aware of this when providing consent.

**Survey metadata**

App-based surveys make it possible to also collect metadata. This kind of behavioural metadata may contain clinically relevant information related to attention, processing speed and working memory capacity, and even any deterioration of psychiatric symptoms. We collect data on the time a survey was opened, time taken to answer each question and fully complete a survey, as well as any changes made in survey responses.

---

**Table 1** Sociodemographic characteristics, pregnancy history and birth outcomes on participants in the Mom2B study and the general population of pregnant women in Sweden

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Mom2B (2020–2022) (n=3909)*</th>
<th>Missing data (n)</th>
<th>n (%) or mean±SD</th>
<th>Sweden (2019)†</th>
<th>Available data (n)</th>
<th>n (%) or mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maternal age (years)</td>
<td>3430</td>
<td>479</td>
<td>31.2±4.4</td>
<td>113 816</td>
<td>30.7</td>
<td></td>
</tr>
<tr>
<td>Country of origin</td>
<td>3441</td>
<td>468</td>
<td></td>
<td>112 530</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sweden</td>
<td>3177 (92.3)</td>
<td></td>
<td></td>
<td>78 033 (69.3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nordic countries except Sweden</td>
<td>40 (1.2)</td>
<td></td>
<td></td>
<td>1280 (1.1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Europe except Nordic countries</td>
<td>116 (3.4)</td>
<td></td>
<td></td>
<td>9172 (8.2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Outside Europe</td>
<td>108 (3.1)</td>
<td></td>
<td></td>
<td>24 045 (21.4)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Education</td>
<td>3444</td>
<td>465</td>
<td></td>
<td>107 711</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤12 years</td>
<td>744 (21.6)</td>
<td></td>
<td></td>
<td>48 793 (45.3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post-secondary education</td>
<td>2700 (78.4)</td>
<td></td>
<td></td>
<td>58 918 (54.7)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Employment before pregnancy</td>
<td>1677</td>
<td>2232</td>
<td></td>
<td>113 147</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Working/student/ parental leave</td>
<td>1626 (97)</td>
<td></td>
<td></td>
<td>103 967 (91.9)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unemployed/sick leave</td>
<td>51 (3)</td>
<td></td>
<td></td>
<td>9180 (8.1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Smoking 3 months before pregnancy</td>
<td>3041</td>
<td>868</td>
<td>441 (14.5)</td>
<td>110 991</td>
<td>11 765 (10.6)</td>
<td></td>
</tr>
<tr>
<td>BMI before pregnancy (kg/m²)</td>
<td>3353</td>
<td>556</td>
<td>25.5±5.3</td>
<td>108 929</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;18.5</td>
<td>70 (2.1)</td>
<td></td>
<td>2783 (2.5)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18.5–25</td>
<td>1815 (54.1)</td>
<td></td>
<td>59 384 (54.6)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25–&lt;30</td>
<td>923 (27.5)</td>
<td></td>
<td>29 636 (27.2)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥30</td>
<td>545 (16.3)</td>
<td></td>
<td>17 126 (15.7)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primiparous</td>
<td>3268</td>
<td>641</td>
<td>1188 (36.4)</td>
<td>113 816</td>
<td>48 473 (42.5)</td>
<td></td>
</tr>
<tr>
<td>Caesarean section</td>
<td>1356</td>
<td>639‡</td>
<td>238 (17.5)</td>
<td>114 757</td>
<td>20 312 (17.7)</td>
<td></td>
</tr>
<tr>
<td>Preterm delivery (&lt;week 37)</td>
<td>3311</td>
<td>598</td>
<td>190 (5.7)</td>
<td>116 071</td>
<td>6502 (5.6)</td>
<td></td>
</tr>
</tbody>
</table>

Percentages are given in relation to available data from women.

*Data downloaded on 1 February 2022.
†Data retrieved from the Swedish Medical Birth Register and Swedish National Board of Health and Welfare from 2019.
‡Calculated using the confirmed number of women in the postpartum period only.
BMI, body mass index.
Figure 2  Timeline of validated instruments administered in the Mom2B study during pregnancy and postpartum periods, and the number of occurrences for each instrument throughout the study period. Surveys become available to users for varying periods of time and will disappear once completed or when their period of availability is over. EPDS, Edinburgh Postnatal Depression Scale107; WHO-5, WHO-5 Well-Being Index106; DSM-screening, Diagnostic and Statistical Manual of Mental Disorders, 5th Edition, criterion for depression; DSM-screening short is a shortened version of the DSM-screening with selected questions chosen by the research team; HADS, Hospital Anxiety and Depression Scale107; PSS, Perceived Stress Scale108; SLE, Stressful Life Events109; LITE, Lifetime Influence of Traumatic Experiences110; FOBS, Fear of Birth Scale111; IPAQ, International Physical Activity Questionnaire112; FSFI, Female Sexual Function Index113; ISI, Insomnia Severity Index114; RS-14, Resilience Scale115; SOC, Sense of Coherence116; VPSQ, Vulnerability Personality Style Questionnaire117; ECRS, Experience in Close Relationships Scale118; Valentine Scale (relationship with your partner)119; S-MIBS, Swedish Mother to Infant Bonding Scale120; PBQ, Postpartum Bonding Questionnaire121; IBQ, Infant Behavior Questionnaire122; SPSQ, Swedish Parenthood Stress Questionnaire123; LMUP, London Measure of Unplanned Pregnancy124; ICQ, Infant Characteristics Questionnaire.125

Figure 3  Timeline of self-developed surveys administered in the Mom2B study during pregnancy and postpartum periods, and the number of occurrences for each instrument throughout the study period. Surveys become available to users for varying periods of time and will disappear once completed or when their period of availability is over.
The app fetches surveys and voice recording tasks, together with the decryption key, where another layer of encryption is added. Passive data collected from the phone follow the same path.

▲ From MinIO, all data are sent to Bianca, a private offline server, in both encrypted and decrypted forms for storage and analysis, respectively.

▲ The app provides weekly reports based on participant activity and fetches health-related information relating to the perinatal period the user is in, as well as frequently asked questions about the study and perinatal health.

**Preliminary data analysis strategy**

The Mom2B data set contains different modalities, including audio data, sensor data and survey data, which will be analysed separately and then combined. We plan to use both traditional ML and DL techniques in order to determine reliable predictors of PND and develop accurate predictive algorithms, and will report our findings following the best fit current guidelines, such as the Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD) statement.

**Feature engineering**

To handle these multimodal data, we will extract features for these modalities separately using traditional feature engineering as well as DL techniques.

An example of traditional feature engineering uses trajectories, which use mobility features, such as number of significant places visited, maximum distance and SD. However, DL can be used for extracting features in many other modalities, such as social media and audio data, which are not investigated widely in the area.

**Feature selection and model selection**

To analyse the multimodal Mom2B data set, we will start with each modality separately. To reduce the possibility of potential overfitting, given the numerous features in our data set, we will use recursive feature elimination to obtain the optimal set of variables for further model development.

Previously, logistic regression, support vector machine, random forests, XGBoost and neural networks have been the most commonly used and efficient ML algorithms for prediction of PND.

An advantage of using such traditional ML methods is to give us a feature importance ranking, allowing us to identify stronger predictors. Using DL to analyse digital phenotyping data for evaluating risk of depression is a relatively novel approach compared with traditional ML models. DL models have been shown to outperform traditional ML in various tasks involving complex data sets, and can be combined with traditional ML in multimodal data mining tasks to further improve performance. We will test and select the best performing ML models for each modality and determine strong predictors of PND.

**National register data**

Supplementary information will be accessed via the following Swedish national health and quality registers: the Medical Birth Register, the Pregnancy Register, the National Patient Register, the Prescribed Drug Register and population censuses from Statistics Sweden. The accessed information includes records of perinatal complications such as PND, PTB or any other complications considered important risk factors for the study outcomes, such as gestational diabetes, gestational hypertension, pre-eclampsia, prolonged delivery, severe lacerations, postpartum haemorrhage, induction of delivery, instrumental vaginal delivery, caesarean section and small for gestational age.

Further, the mother’s weight at enrolment in maternity care and at aftercare visits; calculated date of birth from last menstrual period and from ultrasound; and information on previous miscarriages, previous abortions, chronic diseases, fear of childbirth, involuntary infertility, gestational age at enrolment in maternity care and fetal diagnostics will be obtained from the Medical Birth Register. Retrieved information also includes variables regarding the background, health and lifestyle of the participant for validation purposes of our self-report questionnaires, as well as psychiatric and somatic morbidity for up to 15 years after childbirth.

**Data flow and storage**

Figure 4 illustrates the data flow and storage process as follows:

▲ Participants register to the study via the Mom2B app using their Swedish Social Security number, which is encrypted in the device using a private key provided by the Beiwe backend server, and replaced by a random, pseudonymised code number.

▲ The decryption key, together with the participant consent information and electronic signatures, is stored in a private, write-only server at Uppsala University.

▲ The app fetches surveys and voice recording tasks, and uploads data from participants to the backend server, where it is encrypted and sent to MinIO, a secure, cloud-based storage, where another layer of encryption is added. Passive data collected from the phone follow the same path.
Figure 5 A multimodal machine learning model for peripartum depression (PPD) diagnosis. The extracted features can be classified into three categories: acoustic signals, time series features and categorical features. We can then determine the most suitable model for each category. For example, for acoustic signals, we would apply convolutional neural network (CNN); for time series data, we would apply recurrent neural network (RNN) such as long short-term memory (LSTM); and for numerical variables, we would apply deep neural networks (DNNs) such as transformers, or traditional models like extremely randomised trees (XRT), gradient boosted trees, etc. These models can yield high-dimension representations of multimodal features. After feature fusion, the integrated features will be fed into another neural network for prediction.

Multimodal computational model
The multimodal data we collect are in different scales, dimensions and formats, which need to be harmonised before prediction.103 Different models are better suited to perform on different modalities. To handle this complexity in multiple data modalities, we consider modality fusion during the development phase.

One example of a multimodal ML model is shown in figure 5. The model is designed to detect potential depressive episodes based on multimodal data collected in the Mom2B app. Preprocessed data in three modalities are fed into models and intermediate representations are then fused together and fed in as input features of a classification model.

Evaluation metrics
Our data will be split into a training data set, for analysis, and a test data set, to assess model performance. We consider using multiple evaluation metrics including area under the receiver operating characteristic curve, specificity, sensitivity, positive predictive value, negative predictive value, balanced accuracy and F1 score, as these measurements vary in importance according to the setting and goal of the final algorithm. Thereby, we can compare the performance of traditional ML with DL, and the different assemblies of models, from different perspectives in the context of prediction of PND.

Patient and public involvement
A qualitative study is planned for exploring the attitudes and concerns of participating women towards the Mom2B app. Furthermore, an online survey will be sent to women who have had no recent activity on the app or withdrawn participation from the study. Direct contact with end users and the ability to make changes to the app based on their feedback can enhance user experience and increase engagement. A representative from Mamma till Mamma, a non-profit organisation in Sweden focused on perinatal mental well-being, serves on our advisory board. The organisation has been involved in the piloting of our study and design of questionnaires, and currently supports us with recruitment. We plan to involve them in the dissemination of study results as well.

Substudies
In addition to predicting PND and PTB, the rich data collected from the Mom2B cohort will also be used to investigate further questions, mainly regarding the health of pregnant and postpartum women. Other planned areas of research are regarding the impact of early mother-infant separation and neonatal intensive care of the baby on the well-being of the mother; and sexual function and its potential correlates to depression and anxiety in the perinatal period.

Maternal depression and well-being during the COVID-19 pandemic
In the beginning of 2021, data collected from 1577 participants were used to assess depressive and anxiety symptoms, as well as well-being and life changes in pregnant women in Sweden during the COVID-19 pandemic (from February 2020 to March 2021).102 The Mom2B app enabled gathering psychiatric information at a national level during the pandemic, as well as passive data on mobility. Levels of perinatal affective symptoms and low well-being were elevated compared with previous years and to months with fewer cases. Similar apps can help healthcare providers and governmental bodies to monitor high-risk groups during crises in real time, as well as to adjust measures and the support offered.

ETHICS AND DISSEMINATION
Participants are informed about the aims of the study, and that the confidentiality and security of their data will be assured. All participants provide their consent to participate while registering to the study, and are informed that they can withdraw their participation at any time without giving a reason. Ethical approval has been obtained from the Swedish Ethical Review Authority (dnr: 2019/01170, with amendments) and the project fulfils General Data Protection Regulation (GDPR) requirements, including the processing, storage and protection of all data. Results will be continuously disseminated through international peer-reviewed journals, the project’s website and social media channels, and presented in national and international conferences.
international conferences. All publications will be open access.

DISCUSSION
Strengths and limitations
Besides the utility of digital phenotyping in combination with the advanced analytical methods planned to be used, other strengths of the Mom2B study include the involvement of participants. Statistics based on the WHO-5 Well-Being Index and behavioural data (movement, internet usage, sleep, etc) collected from participants are sent to the user, allowing them to follow their well-being and activity as an incentive for continued participation. Weekly informational reports regarding common experiences and concerns for both the mother and the child for that particular week of the perinatal period, based on information taken from 1177.se (Swedish health-care service), are available to users and allow them to easily stay informed. As per standard guidelines, if participants receive a high score on the EPDS, they are prompted to contact their healthcare provider or emergency support services for support, and if unsure, they can contact the research team, which will help them find appropriate support for their needs. Continuous contact is maintained with participants until they find support.

The involvement of user organisations and an international advisory board further strengthens the study by increasing the feasibility, the use of state-of-the-art methods and the potential for high acceptance by the end users, which is especially important for future integration in regular clinical practice.

However, there are some limitations to acknowledge. Weekly reports and statistics are important in supporting and incentivising users, but it is possible they may influence users’ responses to certain questionnaires. To account for this, we consider including how often they are checked by users as a feature within our models. Furthermore, our app is available only in Swedish, which excludes a number of otherwise eligible participants, and the high costs for maintaining the technical infrastructure in the frontend and backend of the app require considerable funding. Attrition is also an issue, especially with data that require active input from users. While we can attempt to combat this by improving the app based on user feedback, it is important to consider that attrition might also reflect the worsening of symptoms and be a predictor per se of clinical deterioration. It will be important to distinguish such participants and determine how to use attrition as a predictor variable.

Future perspectives
We are at the beginning of the smartphone-based research era, and future possibilities seem numerous. We intend to develop the Mom2B app in other languages, including English, to expand to a more diverse and wider population. If the app succeeds in developing good predictive models for PND, the research team anticipates that the app could be further developed to include evidence-based interventions. Furthermore, since PND is much less understood in co-parents and improving the other parent’s mental well-being is conducive to the health of the mother and the children as well, the app could be further developed to study co-parental PND. The Mom2B research team plans to further adapt the app to other research topics such as teenage and student mental health, and prediction of new episodes or self-harm in major depression.
REFERENCES


73 Graham S, Depp C, Lee EE. Artificial intelligence for mental health and mental health services: an overview: current medicine group LLC 1, 2019: 116–16.


96 Saqib K, Khan AF, Butt ZA. Mobile learning methods for predicting postpartum depression: Scoping review. JMIR Ment Health 2021;8:e29938.


