Tablet computers for implementing NICE antenatal mental health guidelines – Feasibility study

Study Protocol:

Version 1.5

19/01/2017

Authors:

Dr José S Marcano-Belisario

Kevin Doherty

Dr Cecily Morrison

Dr John O'Donoghue

Dr Ajay K Gupta

Dr Gavin Doherty

Prof Azeem Majeed

Dr Josip Car

Principal Investigator

Name: Dr Josip Car

Director, Global eHealth Unit & Reader in eHealth, Imperial College London

Address: Department of Primary Care and Public Health, Room 326, Reynolds Building, St Dunstan's Road

London, W6 8RP

Telephone: 020 7594 0799

Email: josip.car@imperial.ac.uk

Co-investigators

Name: Dr José S Marcano-Belisario

PhD Candidate & Research Assistant, Global eHealth Unit, Imperial College London

Address: Department of Primary Care and Public Health, Room 304, Reynolds Building, St Dunstan's Road

London, W6 8RP

Telephone: 020 7594 0771

Email: jose.marcano@imperial.ac.uk

Name: Kevin Doherty

PhD Candidate, Department of Computer Science and Statistics, Trinity College Dublin

Address: Room 2.3, Westland Row, Trinity College Dublin, College Green

Dublin 2, Ireland

Telephone: +353 860524486

Email: dohertkc@tcd.ie

Name: Dr Cecily Morrison

Honorary Research Fellow, Global eHealth Unit, Imperial College London & Post-Doctoral Researcher, Microsoft Research

Address: Department of Primary Care and Public Health, 3rd Floor, Reynolds Building, St Dunstan's Road

London, W6 8RP

Email: c.morrison@imperial.ac.uk

Name: Dr John O'Donoghue

Deputy Director, Global eHealth Unit & Senior Lecturer in eHealth, Imperial College London

Address: Department of Primary Care and Public Health, Room 326, Reynolds Building, St Dunstan's Road

London, W6 8RP

Telephone: 020 7594 0799

Email: j.odonoghue@imperial.ac.uk

Name: Dr Gavin Doherty

Associate Professor in Computer Science, Trinity College Dublin

Address: School of Computer Science and Statistics, Trinity College Dublin

Dublin 2, Ireland

Telephone: +353 1 8963858

Email: Gavin.Doherty@scss.tcd.ie

Name: Dr Ajay K Gupta

Clinical Research Fellow, National Heart & Lung Institute

Address: International Centre for Circulatory Health, First Floor, 59 – 61 North Wharf Road

London, W2 1LA

Telephone: 020 7594 3437

Email: a.k.gupta@imperial.ac.uk

Name: Professor Azeem Majeed

Chair in Primary Care and Public Health, Imperial College London

Address: Department of Primary Care and Public Health

3rd Floor, Reynolds Building, St Dunstan's Road

London, W6 8RP

Telephone: 020 7594 3368

Email: a.majeed@imperial.ac.uk

Name: Professor Paul Ramchandani

Professor of Child and Adolescent Mental Health

Address: The Centre for Mental Health, Department of Medicine, Imperial College London

Commonwealth Building, Hammersmith Campus

London, W12 0NN

Telephone: 020 3313 4161

Email: p.ramchandani@imperial.ac.uk

Table of Contents

List of Tables and Figures

List of Abbreviations

Apps	Software applications
BYOD	Bring Your Own Device
CKS	Clinical Knowledge Summaries
CSO	Clinical Studies Officer
EPDS	Edinburgh Postnatal Depression Scale
EMA	Ecological Momentary Assessment
FOR	False Omission Rate
GCP	Good Clinical Practice
ICT	Information and Communication Technologies
ICL	Imperial College London
IG	Information Governance
JMIR	Journal of Medical Internet Research
NICE	National Institute for Health and Care Excellence
NHS	National Health Service
NPV	Negative predictive value
NRES	National Research Ethics Service
PHQ-9	Patient Health Questionnaire 9-item
PND	Perinatal Depression
PPI	Patient and Public Involvement
PPV	Positive predictive value
QR	Quick Response
RCI	Reliable Change Index
RCP	Royal College of Psychiatrists
RCT	Randomised Controlled Trial
SOP	Standard Operating Procedures
SSL	Secure Socket Layer
UK NSC	UK National Screening Committee
VAS	Visual Analogue Scale

1. Background

Clinical and public health decisions are often based on evidence that has been constructed by collecting data from multiple sources and through different methods. Quantitative survey methods offer the means of collecting large amounts of highly structured data that can be standardised across collection sites (Bowling 2005; Boynton & Greenhalgh 2004; Carter, Shaw & Thomas 2000; Groves, Fowler, Couper et al. 2009; Hosking, Newhouse, Bagniewska et al. 1995). For this reason, they are often used in population screening programmes, and for the monitoring of symptoms, including those related to disability and mental health (Bowling 2009). Screening programmes however, can be resource intensive. Therefore, their implementation must be preceded by an assessment of their effectiveness and appropriateness against criteria related to the importance of the medical condition, the characteristics of the screening test, and the availability of treatment options (see Table 1) (Bowling 2009; UK NSC 2015a). In addition, there should be clear evidence that the screening programme is effective in reducing morbidity and/or mortality, and that it is acceptable to healthcare professionals and patients alike (UK NSC 2015a). Similarly, the ongoing monitoring of symptoms can present a considerable number of logistical issues.

Table 1. Criteria for appraising the viability, effectiveness and appropriateness of a screening programme (adapted from the UK National Screening Portal)

Domain	Effectiveness and appropriateness criteria
Medical Condition	Should be an important problem
	Well-understood epidemiology and natural history
	Detectable risk factor, disease marker, latent period or early symptomatic stage
	Cost-effective primary prevention interventions should have been implemented as far as practicable
Screening Test	Simple, safe, precise and validated
	Distribution of test values in the target population should be known and a suitable cut-off level defined and agreed
	Acceptable to the population
	Agreed policy on the further diagnostic investigation of individuals with a positive test result and on the choices available to those individuals
Treatment	Effective treatment available
	Clear evidence of early treatment leading to better outcomes than late treatment
	Evidence-based policies covering which individuals should be offered treatment, and what treatment they should receive
	Clinical management of the condition and patient outcomes should be optimised in all health care providers prior to participation in a screening programme
Screening Programme	High quality evidence that the screening programme is effective in reducing mortality and/or morbidity
	Evidence that the complete screening programme is clinically, socially, and ethically acceptable to health professionals and the public
	The benefits from the screening programme should outweigh its psychological and physical harms
	The opportunity cost of the screening programme should be economically balanced in relation to expenditure on medical care as a whole
	All options for managing the medical condition should have been considered, to ensure that no more cost effective intervention could be introduced or current interventions increased within the resources available
	There should be a plan for managing and monitoring the screening programme and an agreed set of quality assurance standards
	Adequate staffing and facilities for testing, diagnosis, treatment and programme management should be available prior to the commencement of the screening programme
	Evidence-based information, explaining the consequences of testing, investigation and treatment, should be made available to potential participants
	Public pressure for widening the eligibility criteria for reducing the screening interval, and for increasing the sensitivity of the testing process, should be anticipated

Maternal mental health is one of the key health priorities in the UK, given the prevalence and impact of conditions such as perinatal depression (PND). PND can affect up to 20% of women during pregnancy or within a year of giving birth (RCP 2012; RCP 2014; Bauer, Parsonage, Knapp et al. 2014). This condition affects many aspects of a woman's life, including her quality of life and long-term mental health, as well as her child's long-term outcomes (including his/her emotional, behavioural, cognitive and social development) (Abdollahi, Zarghami, Azhar et al. 2014; Bauer, Pawlby, Plant et al. 2015; Thomas, Komiti & Judd 2014). These effects coupled with the direct and indirect financial costs to the health system account for the majority of the overall costs to society of perinatal mental disorder (Thomas, Komiti & Judd 2014). These costs have been estimated at £8.1 billion for each one-year cohort of birth (with the average cost of one case of PND estimated at £74k), 72% of which could be attributed to the adverse effects experienced by the child (Bauer, Parsonage, Knapp et al. 2014).

Compared to other stages of life, women are at an increased risk of developing depression 3 and 6 months after childbirth (Abdollahi, Zarghami, Azhar et al. 2014; Yazici, Kirkan, Aslan et al. 2015). Although a recent longitudinal study suggests that symptoms of depression may be more prevalent at 4 years postpartum (Woolhouse, Gartland, Mensah et al. 2014). Several factors are known to increase the likelihood of developing postpartum depression (see Table 2). Depression during pregnancy however, has been recognised as a very important factor and one that is susceptible to intervention. The point prevalence of depression during the first, second and third trimesters of pregnancy has been estimated at 7.4%, 12.8% and 12%, respectively (Bennett, Einarson, Taddio et al. 2004). Treating depression during these stages can reduce the likelihood of developing postpartum depression, prevent more severe forms of this disorder, and improve a woman's general health status (RCP 2012; RCP 2014; Thomas, Komiti & Judd 2014; Yazici, Kirkan, Aslan et al. 2015). Moreover, treating depression during pregnancy may also help to reduce the intergenerational impact of PND. For these reasons, the screening for and monitoring of depressive symptoms in pregnancy should be considered a priority (O'Connor, Rossom, Henninger et al. 2016; Parker, Hegarty, Paterson et al. 2015).

Table 2. Risk factors for perinatal depression (adapted from Abdollahi, Zarghami, Azhar et al. 2014; Mohamad Yusuff, Tang, Binns et al. 2015; Parker, Hegarty, Paterson et al. 2015)

Depression during pregnancy, particularly in the first and second trimesters
Previous history of depression
Psychiatric disorders other than depression during pregnancy (including antenatal anxiety)
Marital difficulties, including intimate partner abuse
Stressful life events
Lack of social or family support
Socioeconomic disadvantage
Unwanted pregnancy
Preference of infant's gender
Gestational diabetes
Recurrent urinary infection

The use of validated screening tools, such as the Edinburgh Postnatal Depression Scale (EPDS) (Hewitt, Gilbody, Brealey et al. 2009), can facilitate the early identification of depressive symptoms during pregnancy, offering the opportunity for prompt intervention. Despite the repeated interactions with the health system that pregnant women may experience as part of their antenatal care, depressive symptoms during this period often go undetected (Tsai, Tomlinson, Dewing et al. 2014). Moreover, the UK National Screening Committee (UK NSC) currently does not recommend screening for PND on the basis of the (i) lack of clarity on the population to be identified by screening, (ii) lack of evidence that current screening tools can accurately identify risk, and (iii) lack of evidence demonstrating the beneficial effect of routine screening on clinical outcomes (UK NSC 2015b). In a recent cost-

effectiveness analysis, Paulden and colleagues (2009) concluded that the routine identification of PND does not seem to offer value for money to the National Health Service (NHS) given the potential additional costs of managing women incorrectly identified as being depressed. However, a recent systematic review conducted by the US Preventive Services Task Force found direct and indirect evidence suggesting that screening pregnant and postpartum women for depression may be beneficial in reducing depressive symptoms and the prevalence of depression (O'Connor, Rossom, Henninger et al. 2016). This situation resembles the debate surrounding the population screening for depression (Gilbody, House, Sheldon 2005; Thombs & Ziegelstein 2014), for which there is currently insufficient evidence to recommend it (which is not the same as having sufficient evidence to oppose it).

Antenatal care provides a good opportunity for the screening and monitoring of depressive symptoms. These repeated interactions between pregnant women and the health system usually take place with the same healthcare provider, which could foment a sense of familiarity. In addition, the timeframe for these interactions is clearly delineated, and users of antenatal services are already engaged with the health system and are more likely to be highly motivated. Depending on the nature of their pregnancy, women accessing antenatal services are not seen as patients in the traditional sense (i.e., someone presenting with an illness) and, in certain appointments, there is scope for a broad discussion around overall well-being and psychosocial determinants of health.

In this setting, the introduction of electronic devices could facilitate depression case-finding and monitoring of symptoms. Electronic delivery modes have been proposed as a way of maximising the scalability and speed of data collection while reducing costs (Lampe & Weiler 1998; Shih & Fan 2009; Groves, Fowler, Couper et al. 2009; Lane, Heddle, Arnold et al. 2006). In recent years, the use of software applications (apps) running on smartphones or tablet computers has received special attention due to the perceived advantages of these devices (i.e., providing continuous connectivity and being almost always on a person). For PND, this could increase the coverage (both in terms of number of women screened and geographic coverage) of screening while reducing its administrative burden (e.g., time needed to score a survey questionnaire) and implementation costs. It could also facilitate the continuous monitoring of women during pregnancy, increase access to treatment options, and enable a prompt and proactive response by the clinical care team. Therefore, electronic screening and monitoring programmes for PND have the potential to be more cost-effective than traditional, paper-based screening programmes.

Researchers from the School of Social Work at the University of Illinois and staff members at Champaign-Urbanara Public Health District have already started exploring the possibility of providing perinatal depression screening to pregnant women attending maternal clinics using tablet computers (MedicalXpress 2015). In the UK, although formal screening is not currently recommended, mobile devices could be used to implement the National Institute for Health and Care Excellence (NICE) recommendations for recognising depression in pregnancy, as outlined in the *Antenatal and postnatal mental health: clinical management and service guidance* (NICE 2014) guideline. The implementation of these recommendations prior a consultation could release some of the consultation time for a discussion of mental health-related issues, thus adding to the cost-effectiveness of this screening process. Similarly, mobile devices could be used to assess depressive symptoms in pregnant women remotely.

Therefore, it is important to determine the feasibility of delivering screening and monitoring programmes via smartphones or tablets in antenatal settings. This needs to take into account the potential impact that mobile devices could have on the quality of the responses collected through screening surveys (i.e., a delivery mode effect), as responses are the result of a complex interaction between survey questionnaires, respondents, and the delivery mode (Bowling 2005; Coons, Gwaltney,

Hays et al. 2009; Fan & Yan 2010; Groves, Fowler, Couper et al. 2009; Manfreda, Bosnjak, Berzelak et al. 2008; Tourangeau, Rips & Rasinski 2000; Wells, Bailey & Link 2014). Delivery mode effects in this context can have significant implications, as screening results will inform diagnostic, referral and treatment decisions.

Similarly, we need to take into account the role of mobile phones in enabling techniques, such as ecological momentary assessment (EMA), that are better suited for the remote monitoring of mood and its associated symptoms. EMA methods support the collection of momentary, experiential and ecologically valid data; they can address some of the biases affecting self-reported retrospective instruments, such as recall bias; and are more sensitive to change, thus facilitating the study of dynamic processes in real-life contexts (Moskowitz & Young 2006; Shiffman, Stone & Hufford 2008). Complemented with retrospective reports, EMA data have been described as holding great promise for the provision of a new perspective on psychological wellbeing (Kahneman, Diener & Schwarz 1999). However, EMA methods have the potential to overburden respondents, thus reducing their long-term engagement and the quality of their responses.

We have conducted preliminary work to explore some of these issues. Through a Cochrane systematic review, we evaluated the impact that apps as a delivery mode have on the quality (defined here as data equivalence, data accuracy, data completeness, adherence to sampling protocols, response rates, time taken by respondents to complete a survey questionnaire, and acceptability to respondents) of self-reported survey responses (Marcano-Belisario, Jamšek, Huckvale et al. 2015). Our results (based on a narrative synthesis of the evidence) suggest that, in clinical settings, survey responses collected via apps are equivalent to those collected via other modes, and that apps might improve data completeness. Our review also highlighted the importance that data completeness and adherence to sampling protocols have outside clinical settings, particularly when repeated measurements are needed. Through this review, we also identified the multiple decisions involved in the adaptation of a validated or non-validated survey for use with an app (e.g., questionnaire design, questionnaire layout, consideration of the original validation circumstances, and data collection technique). These decisions can result in delivery mode effects that are unique to this mode, and which have been the focus of survey methodology research. For example, Wells and colleagues (2014) found that the quality of responses collected via apps were susceptible to experimental manipulations in frequency scales and in the size of text boxes. Similarly, Mavletova and Couper (2014) found that the layout (i.e., scrolling – all questions presented on a single screen – and paging – questions presented one page at a time) of mobile surveys can influence the ease of, and engagement with, survey completion. However, none of the studies included in our systematic review (predominantly from health sciences) explored these factors.

A systematic review has also been conducted concerning conceptions and measures of engagement across the field of computer science, which included 351 studies (under review). Along with the identification of 102 unique definitions of engagement, this review showed that work on engagement has largely been driven by situated experiments, by Flow Theory, by forward-looking developmental work and by gaming and learning applications. This review therefore highlighted the need for longitudinal and ecologically valid studies of user engagement with technology. A second narrative review of the theory behind, applications of, and design lessons to be garnered from previous implementations of systems for ecological momentary assessment was also conducted; the insights from which are applied in the development of the application at the centre of this study. Related work has focused on the differences between momentary and retrospective forms of self report from both the user's perspective and with respect to the objective analysis of data collected by such means (under review).

The aim of the current study is twofold: (1) assessing the feasibility of using tablet computers for implementing the recommendations of the *Antenatal and postnatal mental health: clinical management and service guidance* NICE guideline for recognising depression in pregnancy in the waiting area of antenatal clinics in general practices, midwifery services or secondary care sites; and (2) assessing the feasibility of using an app for the repeated and longitudinal assessment of depressive symptoms and mood-related factors (through EMA techniques) during pregnancy using a bring-your-own-device (BYOD) approach. Additionally, it will explore the feasibility of having a member of the clinical care team providing regular feedback to patients on their data.

Concerning the first aim, this study will evaluate the positive predictive value (PPV), the negative predictive value (NPV), and the false omission rate (FOR) of the Whooley questions (as recommended by NICE). In addition, this study will assess if the findings from the survey methodology literature (Wells, Bailey & Link 2014; Mavletova & Couper 2014) are relevant to validated screening instruments in clinical settings. Therefore, we will assess the impact that iPads as a delivery mode and survey questionnaire layout (i.e., scrolling versus paging) could have on the quality of the responses given to the Whooley Questions and the Edinburgh Postnatal Depression Scale (EPDS) when completed in the waiting area of general practices, midwifery services or secondary care sites prior an antenatal appointment. PPV, NPV and FOR will be assessed by comparing the answers given to the Whooley questions against women's scores on the EPDS. Data quality will be defined in terms of data equivalence (i.e., distribution of overall survey questionnaire scores) between the two layouts (i.e., scrolling and paging) of the EPDS. We will also compare the two layouts (i.e., scrolling and paging) in relation to: (i) acceptability to pregnant women; (ii) the time required to complete survey questionnaires; and (iii) data completeness. We will document the changes made to the original version of the EPDS during its adaptation for use with an app, and we will log the problems experienced by our participants while completing this screening tool. The latter has significant implications for the implementation of screening programmes, as it can pose an additional administrative burden to already busy staff members.

Concerning the second aim, this study will assess response rates and differences in (1) drop-out rates, (2) adherence to sampling protocols, (3) patterns of engagement, (4) qualitative user impressions, (5) timeliness of data collection, and (6) data completion between a longitudinal administration of the EPDS during pregnancy (i.e., retrospective assessment), and the longitudinal administration of this instrument complemented with the EMA of mood, sleep, energy, enjoyment and worry (i.e., retrospective plus momentary assessment). In the latter condition, we will evaluate if the EMA assessments affect the EPDS scores. This study will also explore the feasibility of having a member of the clinical team (e.g., a midwife) providing remote feedback to participants, and the impact that this could have on participant engagement and on the perceived usefulness of this approach (to both participants and practitioners).

2. Aim & objectives

2.1. Tablet computers during antenatal clinics

The aim of this study is to determine the feasibility of implementing the recommendation of the *Antenatal and postnatal mental health: clinical management and service guidance* NICE guideline for recognising depression in pregnancy using iPads in the waiting area of general practices, midwifery services and/or secondary care sites during antenatal clinics. This study will also assess the positive predictive value (PPV), the negative predictive value (NPV) and the false omission rate (FOR) of the Whooley questions, which are recommended by this guideline, by comparing the answers given to them against women's scores on the Edinburgh Postnatal Depression Scale (EPDS). In addition, we will manipulate the survey questionnaire layout (i.e., scrolling vs. paging) of the EPDS in order to replicate the findings from Mavletova and Couper (2014) in a healthcare setting, and determine if this choice affects the quality of responses.

2.1.1. Primary objectives

- To determine the feasibility of implementing the recommendations for recognising depression in pregnancy outlined in the *Antenatal and postnatal mental health: clinical management and service guidance* NICE guideline by using iPads in the waiting area of general practices and/or midwifery services during antenatal clinics

2.1.2. Secondary objectives

- To determine the positive predictive value of the Whooley questions recommended by the *Antenatal and postnatal mental health: clinical management and service guidance* NICE guideline. For this, we will calculate the number of pregnant women who answered *Yes* to any of these questions AND scored 10 points or higher on the EPDS as a proportion of all the pregnant women who answered *Yes* to any of these questions regardless of their scores on the EPDS
- To determine the negative predictive value of the Whooley questions recommended by the *Antenatal and postnatal mental health: clinical management and service guidance* NICE guideline. For this, we will calculate the number of pregnant women who answered *No* to both questions AND scored 9 points or less on the EPDS as a proportion of all the pregnant women who answered *No* to both questions regardless of their scores on the EPDS
- To determine the false omission rate (FOR) of the Whooley questions recommended by the *Antenatal and postnatal mental health: clinical management and service guidance* NICE guideline. For this, we will calculate the number of pregnant women who answered *No* to both questions AND scored 10 points or higher on the EPDS as a proportion of all the pregnant women who answered *No* to both questions regardless of their scores on the EPDS
- To determine the data equivalence between the two layouts (i.e., scrolling versus paging) of the EPDS by calculating the mean difference in the overall EPDS scores between the two groups
- To assess differences between the two layouts (i.e., scrolling versus paging) of the EPDS in the time taken by respondents to complete the EPDS
- To evaluate differences in data completeness between the two layouts (i.e., scrolling versus paging) of the EPDS by comparing the proportion of complete survey questionnaires (i.e., those with no questions left unanswered) as indicated by breakoff rates
- To assess differences in acceptability to patients between the two layouts (i.e., scrolling versus paging) of the EPDS by evaluating (i) the technical problems experienced by respondents during survey completion, and (ii) differences in breakoff rates

2.1.3. Hypotheses

We expect that tablet computers are a feasible way of implementing NICE recommendations in the waiting areas of general practices and/or midwifery services during antenatal clinics. We do not have any hypotheses regarding the PPV, NPV and FOR of the Whooley questions. We expect that differences in the mean overall scores of the EPDS between the two layouts (i.e., scrolling versus paper) will not exceed the minimal clinically significant difference of 4 points (Matthey 2004). As reported by Mavletova and Couper (2014) we expect the paging layout of the EPDS to be more acceptable to respondents and to take less time to complete than the scrolling layout. We do not expect to find differences in the proportion of complete records between the two EPDS layouts (i.e., scrolling versus paging).

2.2. App through a BYOD approach

The aim of this study is to determine the feasibility of using an app running on participants' own devices (i.e., through a BYOD approach) for the monitoring of depressive symptoms and mood-related constructs during pregnancy. We will assess the overall response rates of potential participants who are approached and asked to take part in this study. We will also compare two monitoring strategies, *retrospective* (using the EPDS) versus *retrospective plus momentary* (EPDS complemented with the EMA of mood, sleep, enjoyment, energy and worry), in terms of (1) drop-out rates, (2) adherence to sampling protocols, (3) patterns of engagement, (4) acceptance to participants, (5) timeliness of data collection, and (6) data completeness. We will assess if the implementation of a EMA technique affects the EPDS scores of participants allocated to the *retrospective plus momentary* condition. Additionally, we will assess if feedback by a member of the clinical care team affects engagement with the app, and participants' and clinicians' perceived usefulness of this approach.

2.2.1. Primary Objectives

- To determine the feasibility of using an app running on participants' own smartphones for the remote and longitudinal monitoring of depressive symptoms and mood-related constructs during pregnancy using retrospective assessment or a combination of retrospective and momentary assessment.

2.2.2. Secondary Objectives

- To assess differences in drop-out rates between the participants allocated to the retrospective assessment strategy (i.e., EPDS), and those allocated to the retrospective plus momentary assessment strategy (i.e., EPDS complemented with the EMA of mood, sleep, energy, enjoyment and worry).
- To assess differences in participants' adherence to sampling protocols between a retrospective assessment strategy (i.e., EPDS), and a retrospective plus momentary assessment strategy (i.e., EPDS complemented with the EMA of mood, sleep, energy, enjoyment and worry). We will achieve this by comparing the proportion of participants in each group that complete 100% of the expected assessments.
- To assess differences in usage patterns of the app between a retrospective assessment strategy and a retrospective plus momentary assessment strategy, including additional voluntary self-reporting, time spent using the app, number of interactions with other sections of the app, time to complete the self-assessments, and other ancillary usage data
- To assess differences in participants' acceptance of the app between a retrospective assessment strategy and a retrospective plus momentary assessment strategy through a post-study self-administered questionnaire

- To assess differences in timeliness of data collection between a retrospective assessment strategy and a retrospective plus momentary assessment strategy. We will achieve this by comparing the proportion of participants completing their assessments in response to a prompt displayed on their mobile devices between the two groups
- To assess differences in data completeness between a retrospective assessment strategy and a retrospective plus momentary assessment strategy. We will achieve this by comparing the proportion of complete records between the two groups
- To evaluate the impact of a momentary assessment technique of the EPDS scores by comparing the EPDS scores obtained at the start of the sampling period and the EPDS scores obtained at the end of this period. This will be assessed in the *retrospective plus momentary assessment* condition only
- To assess if feedback provided by a healthcare practitioner influences participants' and clinicians' engagement, acceptance and perceived usefulness of an app used for the remote and longitudinal monitoring of mood symptoms throughout pregnancy
- To describe the trend in EPDS scores and the variability in mood, sleep, energy, enjoyment and worry throughout the duration of this study
- To examine the correlation between variability in mood, sleep, energy, enjoyment and worry, and depressive symptoms (as assessed by the EPDS)
- To examine the association between contextual information and variability in mood, sleep, energy, enjoyment and worry, and the EPDS scores
- To assess safety and rates of adverse events due to the use of the app. Whilst these are expected to be non-existent or rare, our study provides a novel mode of depression screening, with potentially more frequent testing.

2.2.3. Hypotheses

We expect that an app running of participants' own smartphones is a feasible way of remotely monitoring mood symptoms during pregnancy. We do not have any hypotheses regarding the difference in adherence to sampling protocols, timeliness of data collection and data completeness between the *retrospective* and the *retrospective plus momentary* assessments strategies.

3. Methods

3.1. Study Design

We will use a parallel, randomised control trial (RCT) study design. For the study assessing the use of tablet computers in antenatal clinics, participants will be randomly assigned to complete (i) an app version of the Whooley questions and the EPDS in which the questionnaires are presented using a scrolling layout (i.e., App screening – Scrolling), or (ii) an app version of the Whooley questions and the EPDS in which the questionnaires are presented using a paging layout (i.e., App screening – Paging).

Participants will also be asked if they wish to take part in the study evaluating an app through a BYOD approach. Those consenting to take part in this study will be randomly assigned to (i) complete the EPDS once a month for 6 months (i.e., retrospective assessment), or to (ii) complete the EPDS, plus 4 days of EMA of mood, sleep, worry, enjoyment and energy, plus EPDS once a month for 6 months (i.e., retrospective plus momentary assessment).

In selected sites, we will recruit an additional 30 participants into the app through a BYOD approach study. Participants recruited from these sites will receive feedback on their data from designated members of the clinical team.

3.2. Participant recruitment

We will recruit pregnant women attending antenatal clinics in general practices, midwifery services or secondary care NHS centres. Participant recruitment will be conducted by members of the clinical care team who also have a research role, or by Clinical Studies Officers (CSO) who have an honorary contract with or a Letter of Access from the participating NHS centre. Participants will be assessed against our inclusion and exclusion criteria (Table 3) by a member of the clinical care team through a review of medical notes, by the CSO through discussion with the clinical care team, or by asking potential participants directly. Potential participants will be approached opportunistically in the waiting area of participating centres and asked about participation in both studies. If potential participants choose to take part in the study evaluating the tablet computers in antenatal clinics, they will be provided with a copy of the patient information sheet [*Antenatal Mental Health Screening programme – Patient Information Sheet, Version 1.1, dated 06 July 2015*]. If potential participants choose to take part in the study evaluating an app through a BYOD approach, they will be provided with a copy of the patient information sheet [*mHealth for the EMA of mood during pregnancy – Patient Information Sheet, Version 1.2, dated 19 January 2017*]. Participants will be made aware that they can ask any questions about either study, and that they have at least 24 hours before deciding on participation. However, if participants wish to take part in the study on the same day they were approached by a member of the clinical team or a CSO, they will have the opportunity to be consented immediately. In addition, we will explain to participants that refusal to take part in any of these studies or withdrawing from them will not have an impact on their legal rights, medical care or their relationship with those providing medical care. Written informed consent will be taken from those participants who, after receiving all the relevant study information and asking study-related questions, still express their willingness to take part in either study (provided that they meet our participant inclusion and exclusion criteria). Participants can choose to take part in either study (tablet computers in antenatal clinical or app through a BYOD approach), or in both of them.

3.2.1. Participant inclusion and exclusion criteria

We followed the recommendations made by Thombs and Ziegelstein (2014) to inform our participant inclusion and exclusion criteria (see Table 3). For these feasibility studies, we will exclude individuals who are not comfortable reading and writing in English, as the EPDS would need to be re-validated for use in any language other than English to ensure that their psychometric properties remain unchanged.

In addition to the criteria listed in *Table 3*, participants will only be eligible to take part in the study evaluating an app through a BYOD approach if they are in their first trimester of pregnancy and if they own an Apple or Android mobile phone.

Table 3. Participant inclusion and exclusion criteria.

Inclusion Criteria	Exclusion Criteria
18 years old or older	Diagnosis of any common mental health disorder (i.e., depression or anxiety disorders) as specified in the Diagnostic and Statistical Manual of Mental Disorders, 4th Edition, Text Revision
Attending antenatal clinics in participating GP practices	Receiving treatment for any common mental health disorder
	Recent personal history of any common mental health disorder (i.e., within the past 12 months)
	Not comfortable reading and writing in English

3.3. Measurements

After obtaining written informed consent, we will ask participants to complete a brief survey questionnaire asking for personal (not identifiable) demographic information (Appendix 1). This applies to both studies. For participants in the study using tablet computers in antenatal clinics, we will then conduct a single administration of the Whooley questions, followed by a single administration of the EPDS.

Participants in the study using an app through a BYOD device approach, will be asked to complete the EPDS and, depending on the experimental condition, EMA assessments of mood, sleep, energy, enjoyment and worry for 6 months using their own mobile phones. After 6 months, we will ask them to complete a post-study acceptance survey.

3.3.1. Endpoints

3.3.1.1. Tablet computers in antenatal clinics

3.3.1.1.1. Positive predictive value of the Whooley questions

Overall, we will calculate the number of pregnant women who answered *Yes* to any of the Whooley questions AND who scored 10 points or higher on the EPDS, as a proportion of the total number of pregnant women who answered *Yes* to any of the Whooley questions regardless of their EPDS scores.

3.3.1.1.2. Negative predictive value of the Whooley questions

Overall, we will calculate the number of pregnant women who answered *No* to both Whooley questions AND who scored 9 points or lower on the EPDS, as a proportion of the total number of pregnant women who answered *No* to both Whooley questions regardless of their EPDS scores.

3.3.1.1.3. False omission rate of the Whooley questions

Overall, we will calculate the number of pregnant women who answered *No* to both Whooley questions AND who scored 10 points or higher on the EPDS, as a proportion of the total number of pregnant women who answered *No* to both Whooley questions regardless of their EPDS scores.

3.3.1.1.4. Mean overall scores on the EPDS

The score of each participant on the EPDS will be calculated automatically after survey completion. We will then compute the mean overall score on the EPDS for each experimental group.

3.3.1.1.5. Breakoff rates

For each experimental group, we will report the proportion of participants who interrupt the survey completion process before reaching the end of the survey questionnaires (i.e., breakoff rates). In such eventualities, we will document the reason for breakoff.

3.3.1.1.6. Time needed to complete survey questionnaires

This will be the time elapsed (in seconds) between the participants starting to read the basic instructions on how to complete the survey questionnaires and the participant completing the EPDS. These actions will be indicated by participants pressing the *Start* button on the device screen and when a message acknowledging that they have completed the survey questionnaires is displayed on the device screen. This component will be broken down further into the individual survey questionnaires that participants will be required to complete (i.e., personal demographic information, Whooley questions, and EPDS). In addition, we will account for the time that participants spend experiencing problems, distractions or making requests for help or clarification.

3.3.1.1.7. Proportion of complete EPDS records

A complete EPDS record will be defined as one in which no question was left unanswered. For each experimental group, we will compute the proportion of complete EPDS records.

Leaving unanswered questions in the *App Screening – Scrolling* and in the *App Screening – Paging* groups will not be possible due to the implementation of validation procedures. Therefore, data completeness will be indicated by the breakoff rates.

3.3.1.1.8. Proportion of participants requesting help

We will document the requests for help made by each participant, as well as the nature of the request. We will categorise participants into 3 groups: (i) one request for help; (ii) between 2 and 4 requests for help; and (iii) 5 or more requests for help. Subsequently, for each experimental group, we will compute the proportion of participants falling within each category.

3.3.1.1.9. Proportion of each type of request for help

We will document the requests for help made by each participant. We will conduct a content analysis of these requests in order to categorise them according to their nature. For each experimental group, we will compute the proportion of participants making each type of request.

3.3.1.2. App through a BYOD approach

3.3.1.2.1. Response rates

The proportion of potential participants who agree to take part in this study out of the total number of potential participants who are approached about participation in this study.

3.3.1.2.2. Drop-out rates

The proportion of participants who stop completing the retrospective or the retrospective plus momentary assessments before the end of the study (i.e., before 6 months). For those participants taking part in the app through a BYOD approach, we will consider they have dropped out if they miss 2 consecutive assessments. In this case, a notification on their phones will ask them to complete a question about the reason for not completing the assessments.

3.3.1.2.3. Adherence to sampling protocols

The proportion of participants who complete 100% of the expected assessments over the duration of the study.

3.3.1.2.4. Timeliness of data collection

Each pre-specified assessment will be preceded by a prompt. Timeliness of data collection will be defined as the proportion of participants who complete the pre-specified assessments immediately after receiving this prompt.

3.3.1.2.5. Data completeness

The implementation of validation procedures will prevent users from leaving unanswered questions. However, they will be able to terminate the current assessment (leaving a log trail of this). Therefore, data completeness is the proportion of assessment records with complete answers out of the total number of complete plus attempted records.

3.3.1.2.6. Mean overall scores on the EPDS

The score of each participant on the EPDS will be calculated automatically after each pre-specified assessment. We will then compute the mean overall score on the EPDS for each experimental group overall and for each month.

3.3.1.2.7. Mean overall score on the EMA assessments

Mood, energy, enjoyment, sleep and worry will be assessed through a visual analogue scale (VAS) over-imposed on a 5-point Likert scale. For participants allocated to the retrospective plus momentary assessment condition, we will compute the overall score on each domain for each period of each assessment day.

3.4. Materials and Equipment

3.4.1. Survey questionnaires

3.4.1.1. Non-validated survey questionnaires

3.4.1.1.1. Personal demographic information

We will administer a survey (Appendix 1) to collect information on a woman's age group, ethnic background, marital status, employment status, level of education, smartphone or tablet computer ownership, pregnancy history, and previous history of depression.

3.4.1.1.2. EMA assessments

We will assess mood, energy, enjoyment, sleep and worry using VAS scales, over-imposed on a 5-point Likert scale (Appendix 2).

3.4.1.1.3. Contextual information

Following the EMA assessments, we will administer 2 questions asking about the respondents' current location (framed in semantic terms: home, work) and her current activity (e.g., working, shopping) (Appendix 3).

3.4.1.1.4. Post-study acceptance survey

Participants taking part in the *App through a BYOD approach* study, will be asked to complete a post-study survey at the end of their participation (Appendix 4).

3.4.1.2. Validated survey questionnaires

3.4.1.2.1. Whooley questions

The Whooley questions (Appendix 5) were developed by Whooley and colleagues (1997) as a case-finding instrument for depression in primary care. This 2-question instrument screens for depressed

mood and anhedonia that have been present during the past month. Respondents are required to answer Yes or No to each of these questions. An affirmative answer to any of them should be followed by further assessment, including the use of a validated screening instrument (such as the EPDS or the Patient Health Questionnaire 9-item) or referral to a general practitioner or a mental health practitioner. The sensitivity of the Whooley questions has been estimated at 96% and their specificity at 57% (Whooley, Avins, Miranda et al. 1997).

3.4.1.2.2. Edinburgh Postnatal Depression Scale (EPDS)

The EPDS (Appendix 6) is a 10-item self-administered survey questionnaire that was developed to screen for PND in the community (Cox, Holden & Sagovsky 1987). This instrument screens for clinical symptoms such as feelings of guilt, sleep disturbance, reduced energy levels, anhedonia and suicidal ideation that have been present during the 7 days preceding the administration of this survey questionnaire. Each question is scored on a 4-point scale ranging from 0 to 3 points. A final overall score is calculated by adding the scores from each question. Overall scores between 10 and 12 points suggest increased risk for depression; scores of 13 points or above indicate that the diagnostic criteria for major depression disorder have probably been met (Allbaugh, Marcus, Ford et al. 2015). In addition, those administering the EPDS should always pay attention to the answers given to item 10, as it deals with suicidal thoughts. In both instances, a clinical assessment for depression should be provided.

The EPDS is a valid and reliable tool for identifying women at risk of PND, both during pregnancy and postpartum, and is sensitive to changes in the severity of depression over time (Cox, Holden, Sagovsky 1987). The EPDS can be reproduced without further permission provided that the original source of the scale is cited in each reproduced copy.

3.4.2. Software

For the study using tablet computers in antenatal clinics, we will use SNAP Surveys to create the app versions of the (i) personal demographic information survey questionnaire, (ii) the Whooley questions, and (iii) the EPDS.

For the study using an app through a BYOD approach, we will use SNAP Surveys to present the (i) informed consent form [*mHealth for the EMA of mood during pregnancy – Informed Consent Form, Version 1.2, dated 19 January 2017*], (ii) the personal demographic information survey questionnaire, (iii) the Whooley questions, and (iv) the EPDS. In addition, we will use a bespoke app developed by Kevin Doherty to collect (i) the EPDS, (ii) EMA assessments, (iii) the Contextual Information survey, and (iv) the post-study survey. This app will also include additional features providing additional information about mood during pregnancy and a list of available resources.

3.4.3. Mobile devices

For the assessment of tablet computers in antenatal clinics, we will use iPads to deliver the survey questionnaires. For the assessment of an app using a BYOD approach, we will use any tablet computer to collect the baseline assessments, and any Android or iOS handset to install the bespoke app.

3.5. Data collection

3.5.1. Pilot phase & Co-design sessions

3.5.1.1. Tablet computers in antenatal clinics

We will conduct a heuristic evaluation of the user interface of the two layouts (i.e., scrolling versus paging) of the survey questionnaires in order to identify any usability issues prior administration in general practices or midwifery services. During this process, a member of the research team will

evaluate the user interface of the survey questionnaires against the 10 usability principles proposed by Jakob Nielsen (Nielsen 1995).

In addition, we will pilot the administration of both layouts (i.e., scrolling versus paging) of the survey questionnaires with female staff members at the Department of Primary Care and Public Health, Imperial College London (ICL). If possible, we will involve members of the public or members of relevant patient groups in the pilot phase. We expect that this would allow us to identify any usability issues or problems with the study design.

Lastly, during the pilot phase we will develop a script of the instructions that will be given to participants during the study, as well as our standard operating procedures (SOP).

3.5.1.2. App through a BYOD approach

As part of a patient and public involvement and a service development initiative, we will conduct workshops with potential end users of the bespoke app and with interested stakeholders (including general practitioners, nurses, midwives, consultants, health visitors, carers, and members of charity or advocacy groups) to gather feedback on the initial prototype of the app. This feedback will be incorporated into the final working version of the bespoke app.

3.5.2. Study Phase

3.5.2.1. Tablet computers in antenatal clinics

After obtaining written informed consent from our participants, a member of the research team will impart verbal instructions on how to complete the survey questionnaires. Participants will be informed that their results on the EPDS will be communicated to her clinician and that they will be available for discussion during her antenatal appointment.

Moreover, the member of the research team imparting the instructions will explain to participants that they can ask for help at any point while completing the survey questionnaires, and that all questions should be directed to a member of the research team. We will offer help with the use of the devices. However, we will not offer help reading questions out loud to participants, as this corresponds to an interview-administered survey questionnaire in which different factors could influence the quality of the survey responses.

Participants will be asked to complete all the survey questionnaires while waiting for their appointment in the waiting area of participating general practices, midwifery services or hospitals.

3.5.2.1.1. App Screening – Scrolling layout

Participants allocated to this intervention arm will receive an iPad running the survey questionnaires (personal demographic information, Whooley questions, and EPDS). We will consider that a participant has started completing the survey questionnaires as soon as they press the button *Start* displayed on the screen. Immediately after, a screen will be displayed which will contain an outline of the survey questionnaires to be completed and basic instructions on how to complete them. In this experimental group each survey questionnaire will be presented on a single screen. Therefore, participants will need to scroll vertically in order to answer all the questions. They will be able to scroll up and down as they wish and to modify the answers to previous question as long as they do it before submitting their answers. However, once they have submitted their responses, a message acknowledging their participation will be displayed on the screen of the iPad Air and they will no longer be able to modify their answers. Validation procedures will ensure that participants will not be able to submit a survey with unanswered questions. Once the surveys have been completed, the participants will return the tablet to a member of the research team. Subsequently, the participant will be given printed material

containing additional information on maternal mental health, and a printed letter [*Antenatal mental health screening – Communicating results to the clinician,* version 1.0, dated 28 April 2015] with her score on the EPDS will be given to her clinician.

3.5.2.1.2. App Screening – Paging layout

Participants allocated to this intervention arm will receive an iPad Air running the survey questionnaires (personal demographic information, Whooley questions, and the EPDS). We will consider that a participant has started completing the survey questionnaires as soon as they press the button *Start* displayed on the screen. Immediately after, a screen will be displayed which will contain an outline of the survey questionnaires to be completed and some basic instructions on how to complete them. In this experimental group each survey questionnaire will be presented using a paging layout. Therefore, participants will only see one question on the screen at any given time. They will need to select and submit their answer to the current question before being able to move on to the next question. Participants will not be allowed to re-visit already answered questions. Validation procedures will ensure that participants will not be able to submit an unanswered question. Once the surveys have been completed, the participants will return the tablet to a member of the research team. Subsequently, the participant will be given printed material containing additional information on maternal mental health, and a printed letter [*Antenatal mental health screening – Communicating results to the clinician,* version 1.0, dated 28 April 2015] with her score on the EPDS will be given to her clinician.

3.5.2.1.3. Communicating answers to the Whooley questions and EPDS scores to clinicians

For participants in both experimental groups (i.e., *App Screening – Scrolling,* and *App Screening – Paging*) a member of the research team will complete the *Antenatal mental health screening – Communicating results to the clinician* (Version 1.0, dated 28 April 2015) form immediately after a participant has completed the study. In this letter, the researcher will write down the participant's answers to the Whooley questions and her score on the EPDS; a legend in this letter will aid with the interpretation of scores. In addition, the researcher will be able to tick a box that indicates that the participant is experiencing suicidal thoughts (as assessed by question 10 of the EPDS). This letter will be handed over to the clinician before the antenatal consultation for that participant takes place.

3.5.2.1.4. Duty of care

All participants will receive a sheet of paper containing additional information about mental health during pregnancy and a list of useful resources (i.e., links to the information leaflets produced by the Royal College of Psychiatrists (RCP).

In addition, we will agree a protocol with the clinicians in our participating general practices, midwifery services or secondary care sites to ensure that those participants who receive a score suggestive of depression are assessed and managed as recommended by the *Antenatal and postnatal mental health: clinical management and service guidance* NICE guideline (NICE 2014) and their local pathways.

3.5.2.2. App through a BYOD approach

After obtaining written informed consent from our participants, a member of the research team, a member of the clinical care team, or a CSO will ask the participant to complete a personal demographic survey, the Whooley Questions and the EPDS using a tablet provided by the research team. At this point, participants will be randomised to one of two experimental conditions: *retrospective assessment*, or *retrospective plus momentary assessment*. After random allocation has taken place, participants will be given a Quick Response (QR) code to scan with their own smartphone, or otherwise directed to download and install the app. This app will be free of charge and it will **not** offer any in-app purchases. After installing this app, participants will need a code in order to activate their account and

be able to take part in the study. This activation code will be generated in advance by the study coordinator. We will communicate a participant's baseline EPDS score to her clinician (i.e., midwife or consultant) if it is 10 points or higher, or if she has been identified as having increased risk of self-harm as measured by question 10 of the EPDS. We will do this through the document *Antenatal mental health screening – Communicating results to the clinician (Version 1.0, dated 28 April 2015*). Participants will be required to take part for 6 months.

3.5.2.2.1. Retrospective assessment

Participants in this condition will be asked to complete the EPDS once a month for 6 months. The day on which participants are required to complete the EPDS will be selected at random from each calendar month, and with no less than 2 weeks in between 2 consecutive sampling periods. On that day, we will select a random time between 10.00 and 22.00 to administer the EPDS. At that time, a notification will be displayed on a participant's phone asking her to complete the EPDS. They will then be able to accept the request or reject it. Outside the sampling period, participants will be allowed to complete as many assessments as they wish to. However, during the sampling period, participants will not be allowed to complete any additional assessments.

3.5.2.2.2. Retrospective plus momentary assessment

Participants in this condition will be required to complete a sampling period of 4 consecutive days once a month for 6 months; this period will be randomly selected from each calendar month, with no less than 2 weeks in between 2 consecutive sampling periods. The sampling period will start on day 1 with an administration of the EPDS and will conclude with another administration of the EPDS on day 4. In between (i.e., days 1 to 4 of the sampling period) participants will be asked to complete EMA assessments on mood, sleep, energy, enjoyment and worry plus 2 contextual information questions 3 times a day. The sampling hours will run between 10.00 and 22.00. This 12-hour window will be divided into 3 periods of 4 hours each, and with no less than one hour in between 2 consecutive assessments. Participants will receive a notification prompting them to complete the assessment at random times within each 4-hour period. Participants will have the option to accept or reject the request in the notification. Outside the sampling period, participants will be allowed to complete as many assessments as they wish to. During the sampling period however, participants will not be allowed to complete any additional assessments.

3.5.2.2.3. Duty of care

The EPDS is a screening tool that indicates the likely risk of depression during pregnancy or postpartum. Scores between 10 and 12 points indicate an increased risk of depression; whereas a score of 13 points or more indicates that the diagnostic criteria for major depression have probably been met. In both cases, further assessment is warranted to confirm a diagnosis of depression. In addition, question 10 of the EPDS deals with a woman's risk of self harm. Scores of between 1 and 3 points on this question warrant further assessment.

For every participant, we will collect at baseline the following contact details:

- GP: telephone number, email address and postal address
- Designated clinician: telephone number, email address and postal address

The system will send an email notification to the study coordinator if any participant:

- Scores 10 points or more on the EPDS
- Scores 1 point or more on Question 10 of the EPDS regardless of her overall score on the EPDS
- Completes more than 2 instances of the EPDS in one day

Upon receiving any notification, the study coordinator will notify the participant's risk to all of her contacts via the telephone. This will be followed up by a written notification using the documents *Antenatal mental health screening – Letter to the GP (Version 1.1, dated 06 December 2015*) or *Antenatal mental health screening – Communicating results to the clinician (Version 1.0, dated 28 April 2015*) being posted to the participant's GP and her designated clinician. The study coordinator will action this within 24 hours.

3.5.2.2.4. Feedback by a human supporter

Sites choosing to take part in this part of the study, will identify a member of the clinical care team who will be in charge of reviewing participants' data, and of providing feedback to them through a desktop web interface that will be developed by Kevin Doherty (they will not be using an app on their personal devices for this purpose). The type of feedback that participants will receive will be through templates that the designated member of staff will tailor to each participant. The most appropriate content for these templates will be identified through the co-design sessions outlined in *Section 3.5.1.2*. The designated member of the clinical care team **will not** replace the person conducting the procedures outlined in *Section 3.5.2.2.2. Duty of care*. In addition, the provision of feedback to participants **is not** intended to be an emergency or out-of-hours service. No personal contact details will be exchanged between participants and the designated member of the clinical team.

The designated members of the clinical team will be required to spend no more than 30 minutes once a week reviewing participants' data and providing feedback. After the study, they will be asked to take part in semi-structured interviews to assess their general sense of usefulness of this method and to complete a post-study survey (Appendix 7).

For participants in this experimental condition, we will assess their sense of connectedness and general sense of usefulness through additional questions in the post-study acceptance survey (Appendix 4).

3.6. Sample size

According to Matthey (2004), the minimal clinically significant difference for the EPDS (calculated using a reliable change index (RCI) and a standard deviation of 4) is 4 points. Therefore, for the study using tablet computers in antenatal clinics, in order to demonstrate a one-sided equivalence between the two layouts (i.e., scrolling versus paging) of the EPDS using an additive model (i.e., the mean score difference between the two groups not exceeding 3 points) with an alpha level of 0.025, power of 0.80 and a standard deviation of 4 we would need 126 participants. In addition, in order to demonstrate the PPV, PNV and FOR of the Whooley questions we would need at least 30 events (i.e., women scoring 10 points or higher on the EPDS). Assuming a prevalence of depression in pregnant women of 12%, we would need to recruit 250 participants. For these reason, we will aim to recruit a total sample of at least 300 participants.

For the app through a BYOD approach study, we aim to determine adherence to sampling protocols. We have chosen to relate the proposed sample size to the 95% confidence interval for the rate, as recommended by the Research Design Service in London. Therefore, we would need 96 participants in each arm with a 95% confidence level and a confidence interval of 10. This translates into a total sample of 192 (i.e., 96 participants in each experimental group). For this reason, we will aim to recruit at least 200 participants.

For the additional condition where participants will receive feedback from a designated member of the clinical care team, we aim to recruit 30 participants.

3.7. Randomisation

For each of the studies, we will utilise a block randomisation procedure to allocate participants to one of 2 experimental arms. Random numbers will be generated using Stata 13.0 and each consecutive number that is generated will be printed and put inside an opaque envelope, which will then be sealed. Envelopes will be numbered in an ascending sequence. Once a participant has provided informed consent, the researcher will take the relevant envelope, open it, and use the number contained in it to allocate the participant to one of the 2 experimental groups. Researchers conducting participant recruitment will not be involved in this randomisation procedure in order to avoid recruitment bias.

3.8. Statistical Analysis

3.8.1. Descriptive statistics

We will report the number of participants approached that did not meet our participant inclusion criteria, as well as the main reason for exclusion (as a fraction of the total number of participants that were excluded). In addition, we will report the proportion of participants who were approached and refused to take part in our study as a fraction of the total number of participants approached (i.e., refusal rates).

Moreover, we will report the following information for each experimental group and for each study:

- Main demographic characteristics: percentage of participants according to age group, ethnic group, employment status, educational attainment, marital status, number of children, smartphone or tablet ownership, and previous history of depression
- Proportion of participants answering *Yes* to any of the Whooley questions
- Proportion of participants scoring at each interval of the EPDS: (i) 0 and 9 points; and (ii) 10 and 12 points; and (iii) 13 points or above
- Proportion of participants scoring 2 or 3 points on question 10 of the EPDS
- Mean time spent on each survey questionnaire (i.e., personal demographic information, Whooley questions, and the EPDS)
- Proportion of participants making each type of request for help
- Mean time spent dealing with requests for help
- Breakoff rates and reasons for breakoff

3.8.2. Inferential statistics

3.8.2.1. Positive predictive value (PPV)

We will calculate the PPV of the Whooley questions by dividing the total number of pregnant women who answered *Yes* to any of the Whooley questions and scored 10 points or higher on the EPDS by the total number of women who answered *Yes* to any of the Whooley questions. We will present this value in percentage form.

3.8.2.2. Negative predictive value (NPV)

We will calculate the NPV of the Whooley questions by dividing the total number of pregnant women who answered *No* to both Whooley questions and scored 9 points or lower on the EPDS by the total number of women who answered *No* to both Whooley questions. We will present this value in percentage form.

3.8.2.3. False omission rate (FOR)

We will calculate the FOR of the Whooley questions by dividing the total number of pregnant women who answered *No* to both Whooley questions and scored 10 points or higher on the EPDS by the total

number of women who answered *No* to both Whooley questions. We will present this value in percentage form.

3.8.2.4. Data equivalence

We will determine data equivalence between the two EPDS layouts (i.e., scrolling and paging) by testing for statistically significant differences in the mean overall scores on the EPDS. For this, we will conduct a t-test.

In addition, we will compare the proportion of women scoring (i) between 10 and 12 points, and (ii) 13 or more points between the 2 experimental groups using a t-test.

3.8.2.5. Time needed to complete the survey questionnaires

We will compare the mean time needed by participants to complete the survey questionnaires between the experimental groups. For this we will conduct a t-test.

3.8.2.6. Data completeness

We will compare the proportion of complete records between the experimental groups using a t-test.

3.8.2.7. Proportion of requests for help

We will compare the proportion of participants making requests for help between the experimental groups. For this, we will conduct a t-test for each category in this endpoint: (i) one request; (ii) between 2 and 4 requests; and (iii) 5 or more requests.

3.8.2.8. Drop-out rates

We will compare the proportion of potential participants who drop-out the study before 6 months (out of the total number of participants enrolled in the study) between the experimental groups.

3.8.2.9. Adherence to sampling protocols

We will compare the proportion of participants who completed 100% of the expected assessments between the experimental groups using a t-test.

3.8.2.10. Timeliness of data collection

We will compare the proportion of participants who completed the pre-specified assessments after receiving a prompt between the experimental groups using a t-test.

3.9. Timetable

Recruitment of participants began on 19 October 2015 (for the study using tablet computers in antenatal clinics). Recruitment for the study using an app through a BYOD approach is expected to start in January 2017. Overall recruitment is expected to conclude on 31 December 2017.

4. Patient and Public Involvement

For the tablet computers in antenatal clinics study, we will aim to recruit two pregnant women to advise us during the design and piloting of the survey questionnaires. We will also aim to involve these members of the public in the overall management of our study, and during the interpretation of our results. We will recruit them from the same general practices, midwifery services and hospitals where this study will be conducted. We will apply for the *Enabling Involvement Fund* offered by the Research Design Service London in order to reimburse these members of the public for travel expenses and their time.

For the app through a BYOD approach study, we will engage patients and the general public through our co-design sessions.

We expect that through patient and public involvement (PPI), we will be able to identify issues that are relevant to the target population so that our findings will become more relevant, and will have a more significant impact.

5. Regulatory Issues

5.1. Ethics approval

Ethics approval was obtained from the National Research Ethics Service (NRES) Committee South East Coast – Surrey.

5.2. Consent and withdrawal

Local research teams or CSOs will be responsible for taking written, informed consent at the participating general practices, midwifery services or secondary care sites. This will be done by adequately trained research or clinical staff using the documents *Antenatal mental health screening programme – Consent Form* (Version 1.1, dated 06 July 2015) or *mHealth for the EMA of mood during pregnancy – Informed Consent Form* (Version 1.2, dated 19 January 2017) as relevant. Participants will be given the opportunity to ask any questions prior to giving consent. In addition, participants will be made aware that they have at least 24 hours to decide whether or not they wish to take part in this study. However, if they wish to take part in the study, they will have the opportunity to consent immediately. Furthermore, the right of participants to withdraw will be clearly explained. Participants may withdraw at any stage without having to offer an explanation. Refusal to take part in any of these studies or withdrawing from them will not have an impact on the patients' legal rights, medical care or their relationship with those providing medical care.

5.3. Anonymity and Confidentiality

We will ensure the anonymity and confidentiality of the data collected from our participants through a number of steps. For the study using tablet computers in antenatal clinics, we will not collect personal identifiers from our participants. The data generated as part of this study will be stored in an anonymised database. This database will be hosted by the ICT Department at Imperial College London and will be located within the Imperial College London secure network. For this study, participant information will not be stored in any of the devices used for data collection (i.e., iPads); instead, these data will be transferred automatically to our secure database via a secure connection using secure socket layer (SSL) certificates.

For the study using an app through a BYOD approach, we will implement PIN authentication procedures and any data stored in participants' handsets will be encrypted. Transmission of data to our database will also take place via a secure connection using SSL certificates. We will implement two separate databases: (i) an anonymised database to store the data generated during the study, and (ii) a secondary database where we will store participants' details and their clinicians' details (which will be collected during the baseline assessment). The app will only need to connect to the anonymised database. The secondary database will be accessed by the researchers conducting this study, only in cases where a participant's EPDS score needs to be communicated to her clinician.

For both studies, access to the study databases will be restricted to researchers working on this project, and it will be enforced through administration of user rights and authentication of Imperial College London username and password.

5.4. Indemnity

Indemnity cover will be arranged through the Imperial College Joint Research Compliance Office (JRCO).

5.5. Sponsor

Imperial College London will act as the sponsor for this study.

5.6. Assurance

We will seek NHS assurance through the relevant Comprehensive Research Networks and Local R&D Departments.

JC will act as guarantor for this work. JC, CM, JOD, GD and AKG are all experienced researchers who will ensure the quality standards of this work.

We will adhere to principles of GCP and all relevant ICL and Trinity College policies.

5.7. Funding

The funding needs of this study have been met through the Imperial Biomedical Research Centre.

5.8. Audits and Inspection

This study may be subject to inspection and audit by the Joint Research Compliance Office under the remit as sponsor and other regulatory bodies to ensure adherence to Good Clinical Practice and the Research Governance Framework.

6. Study management

6.1. Data handling

We will ensure compliance with all the relevant NHS policies (i.e., Data Protection Policy, Freedom of Information Policy, Confidentiality Policy, Information Security Policy, Document and Records Management Policy, Information Sharing Policy, and Information Governance Policy). In addition, we will comply with the Information Governance (IG) policies of participating general practices, midwifery services or hospitals, as well as with the principles of IG outlined by the NHS Health Research Authority. In addition, we will observe all the relevant ICL policies (i.e., Data Quality Policy, and Information Systems Security Policies). Lastly, all the staff involved in this project will complete the Good Clinical Practice (GCP) training provided by ICL, and the IG training provided by the NHS.

6.2. Data Storage

We will keep one master file for signed consent forms. This file will be kept in a locked cabinet located in a secure room at the Department of Primary Care and Public Health, Imperial College London (ICL), Reynolds Building, St Dunstan's Road, London. Access to this room will be restricted to members of the research team working on this project.

We will not store participants' responses in the iPad devices used for the tablet computers in antenatal clinics study; instead, their responses will be transferred to a secure, anonymised database hosted in the ICL secure network. Data will be transmitted over a secure connection with Secure Socket Layer (SSL) certificates.

Data relevant to this study will be anonymised and stored in a secure database hosted by the Information and Communication Technologies (ICT) Department at ICL and located within the ICL secure network. This database will be encrypted using an AES 256 certificate (or equivalent). Access to

this database will be limited to ICL staff members working on this research project, which will be controlled through a policy of authorised usernames and passwords.

We will not store any personal identifiers for our participants in the study using tablets in antenatal clinics. For the study assessing an app through a BYOD approach, we will keep a secondary database containing participants' and their clinicians' contact details. This database will only be accessed if a participant's EPDS score needs to be communicated to her clinician.

No personal equipment will be used throughout the duration of this study.

After concluding this study, we will archive all data for up to 10 years according to Imperial College London policies.

6.3. Sensitivity issues

It is possible that participants with no previous history of common mental health disorder might receive a score highly suggestive of depression or anxiety. In these cases, we will ensure that patients are referred to the relevant mental health services linked to participating general practices, midwifery services or hospitals.

6.4. Dissemination plan and publication policy

The findings from different stages of this study formed part of José S Marcano Belisario's and will form part of Kevin Doherty's PhD Theses. One of the study protocols has been published in the BMJ Open (Marcano-Belisario, Gupta, O'Donoghue et al. 2016), and we will submit another study protocol for publication in the BMJ Open, Journal of Medical Internet Research (JMIR Research Protocols (www.researchprotocols.org) or similar journals. The findings of these studies will be disseminated through papers submitted to academic peer-reviewed journals, presentations and posters submitted to professional and academic conferences, and discussion with peers.

6.5. Future research

With these studies we expect to demonstrate the feasibility of implementing NICE guidelines for antenatal mental health care in general practices or midwifery services using tablet computers, and the feasibility of using a BYOD approach to support at-home collection of patient self-reported data. We also expect to ensure that the electronic administration (through iPads) of the EPDS does not affect the psychometric properties of this instrument. In addition, we will explore factors that affect patient engagement with longitudinal, repeated data collection.

The results from these studies will answer key questions about the impact of mobile technology on the quality of clinical data. These findings will inform a clinical trial evaluating the cost-effectiveness of mobile-based depression screening and monitoring in the UK, including its potential impact on clinical outcomes (i.e., postpartum depression). In addition, we will use these findings to inform a clinical trial led by the Nanyang Technological University in Singapore.

References

Abdollahi, F., Zarghami, M., Azhar, M. Z., Sazlina, S-G., Lye, M-S. (2014). Predictors and incidence of post-partum depression: A longitudinal cohort study. *The Journal of Obstetrics and Gynaecology Research; 40(12):* 2191 – 2200. doi: 10.1111/jog.12471

Allbaugh, L. J., Marcus, S. M., Ford, E. C., Flynn, H. A. (2015). Development of a screening and recruitment registry to facilitate perinatal depression research in obstetrics settings in the USA. *International Journal of Gynocology and Obstetrics; 128:* 260 – 263.

Bauer, A., Parsonage, M., Knapp, M., Iemmi, V., Adelaja, B. (2014). *The cost of perinatal mental health problems*. London: Centre for Mental Health, London School of Economics

Bauer, A., Pawlby, S., Plant, D. T., King, D., Pariante, C. M., Knapp, M. (2015). Perinatal depression and child development: exploring the economic consequences from a South London cohort. *Psychological Medicine; 45:* 51 – 61. doi: 10.1017/S0033291714001044

Bennett, H. A., Einarson, A., Taddio, A., Koren, G., Einarson, T. R. (2004). Prevalence of depression during pregnancy: systematic review. *Obstetrics and gynecology; 103(4):* 698 – 709.

Bowling, A. (2005). Mode of questionnaire administration can have serious effects on data quality. *Journal of Public Health; 27(3):* 281 – 291

Bowling, A. (2009). *Research Methods in Health* (3rd Edition). McGraw Hill: New York, USA.

Boynton, P. M., Greenhalgh, T. (2004). Selecting, designing, and developing your questionnaire. *BMJ; 328:* 1312 – 1313

Carter, Y., Shaw, S., Thomas, C. (2000). *The Use and Design of Questionnaires*. Master Classes in Primary Care Research No 2. The Royal College of General Practitioners: London

Coons, S. J., Gwaltney, C. J., Hays, R. D., Lundy, J. J., Sloan, J. A., Revicki, D. A., et al. (2009). Recommendations on evidence needed to support measurement equivalence between electronic and paper-based patient-reported outcome (PRO) measures: ISPOR ePRO Good Research Practices Task Force report. *Value in Health; 12(4):* 419 – 429. doi: 10.1111/j.1524-4733.2008.00470.x

Cox, J. L., Holden, J. M., Sagovsky, R. (1987). Detection of Postnatal Depression. Development of the 10-item Edinburgh Postnatal Depression Scale. *British Journal of Psychiatry; 150:* 782 – 786

Fan, W., Yan, Z. (2010). Factors affecting response rates of the web survey: A systematic review. *Computers in Human Behaviour; 26(2):* 132 – 139

Gilbody, S., House, A., Sheldon, T. (2005). Screening and case finding instruments for depression. Cochrane Database of Systematic Reviews; Issue 4: Art. No CD002792. doi: 10.1002/14651858.CD002792.pub2

Groves, R. M., Fowler, F. J., Couper, M. P., Lepkowski, J. M., Singer, E., Tourangeau, R. (2009). *Survey Methodology* (2nd Edition). WILEY: London.

Hewitt, C. E., Gilbody, S. M., Brealey, S., Paulden, M., Palmer, S., Mann, R., Green, J., Morrell, J., Barkham, M., Light, K., Richards, D. (2009). Methods to identify postnatal depression in primary care: an integrated evidence synthesis and value of information analysis. *Health Technology Assessment; 13: 36*.

Hosking, J. D., Newhouse, M. M., Bagniewska, A., Hawkins, B. S. (1995). Data Collection and Transcription. *Controlled Clinical Trials; 16:* 66S – 103S

Kahneman, D., Diener, E., Schwarz, N. (1999). *Well-being: Foundations of Hedonic Psychology.* New York: Russell Sage Foundation.

Lampe, A., Weiler, J. (1998). Data Capture from the Sponsors' and Investigators' Perspectives: Balancing Quality, Speed, and Cost. *Drug Information Journal; 32:* 871 – 886

Lane, S. J., Heddle, N. M., Arnold, E., Walker, I. (2006). A review of randomized controlled trials comparing the effectiveness of hand held computers with paper methods for data collection. *BMC Medical Informatics and Decision Making; 6(23):* 1 – 10. doi: 10.1186/1472-6947-6-23

Manfreda, L., Bosnjak M., Berzelak, J., Haas, I., Vehovar, V. (2008). Web Surveys versus Other Survey Modes – A Meta-Analysis Comparing Response Rates. *International Journal of Market Research; 50(1):* 79 – 104

Marcano Belisario, J. S., Gupta, A. K., O'Donoghue, J., Morrison, C., Car, J. (2016). Tablet computers for implementing NICE antenatal mental health guidelines: protocol of a feasibility study. *BMJ Open; 6:* e009930.

Marcano Belisario, J. S., Jamšek, J., Huckvale, K., O'Donoghue, J., Morrison, C. P., Car, J. (2015) Comparison of self administered survey questionnaire responses collected using mobile apps versus other methods. *Cochrane Database of Systematic Reviews; Issue 7.* Art No: MR000042. DOI: 10.1002/14651858.MR000042.pub2.

Matthey, S. (2004). Calculating clinically significant change in postnatal depression studies using the Edinburgh Postnatal Depression Scale. *Journal of Affective Disorders; 78(3):* 269 – 272.

Mavletova, A., Couper, M. P. (2014). Mobile Web Survey Design: Scrolling Versus Paging, SMS Versus E-Mail Invitations. *Journal of Survey Statistics and Methodology; 2(4):* 498 – 518. doi: 10.1093/jssam/smu015

MedicalXpress (2015). *Using tablets to screen new, expecting moms for perinatal depression*. Available from: http://medicalxpress.com/news/2015-03-tablets-screen-moms-perinatal-depression.html [Accessed 31 March 2015]

Mohamad Yusuff, A. S., Tang, L., Binns, C. W., Lee, A. H. (2015). Prevalence and risk factors for postnatal depression in Sabah, Malaysia: A cohort study. *Women and Birth; 28:* 25 – 29

Moskowitz, D. S., Young, S. N. (2006). Ecological momentary assessment: what it is and why it is a method of the future in clinical psychopharmacology. *Journal of Psychiatry & Neuroscience; 31(1):* 12 – 20.

National Institute for Health and Care Excellence (2014). *Antenatal and postnatal mental health: clinical management and service guidance*. NICE clinical guideline 192. Available from: http://www.nice.org.uk/guidance/cg192/resources/guidance-antenatal-and-postnatal-mental-health-clinical-management-and-service-guidance-pdf [Accessed 27 March 2015]

Nielsen, J (1995). *10 Usability Heuristics for User Interface Design.* Available from: http://www.nngroup.com/articles/ten-usability-heuristics/ [Accessed 24 March 2015]

Norman, C. D., Skinner, H. A. (2006). eHEALS: The eHealth Literacy Scale. *Journal of Medical Internet Research; 8(4):* e27.

O'Connor, E., Rossom, R. C., Henninger, M., Groom, H. C., Burda, B. U. (2016). Primary Care Screening for and Treatment of Depression in Pregnant and Postpartum Women. The *Journal of the American Medical Association; 315(4):* 388 – 406.

Parker, G. B., Hegarty, B., Paterson, A., Hadzi-Pavlovic, D., Granville-Smith, I., Gokiert, A. (2015). Predictors of post-natal depression are shaped distinctly by the measure of 'depression'. *Journal of Affective Disorders; 173:* 239 - 244

Paulden, M., Palmer, S., Hewitt, C., Gilbody, S. (2009). Screening for postnatal depression in primary care: cost effectiveness analysis. *British Medical Journal; 339:* b5203. doi: http://dx.doi.org/10.1136/bmj.b5203

Royal College of Psychiatrists (2012). *Mental health in pregnancy*. Available from: http://www.rcpsych.ac.uk/healthadvice/problemsdisorders/mentalhealthinpregnancy.aspx [Accessed 10 February 2015]

Royal College of Psychiatrists (2014). *Postnatal Depression.* Available from: http://www.rcpsych.ac.uk/healthadvice/problemsdisorders/postnataldepression.aspx [Accessed 10 February 2015]

Shiffman, S., Stone A. A., Hufford, M. R. (2008). Ecological momentary assessment. *Annual review of clinical psychology; 4:* 1 – 32.

Shih, T. H., Fan, X. Comparing response rates in e-mail and paper surveys: A meta-analysis. *Educational Research Review; 4(1):* 26 – 40. doi: 10.1016/j.edurev.2008.01.003

Thomas, N., Komiti, A., Judd, F. (2014). Pilot early intervention antenatal group program for pregnant women with anxiety and depression. *Archives of Women's Mental Health; 17:* 503 – 509. doi: 10.1007/s00737-014-0447-2

Thombs, B. D., Ziegelstein, R. C. (2014). Does depression screening improve depression outcomes in primary care? *British Medical Journal; 348*: g1253. doi: 10.1136/bmj.g1253

Tourangeau, R., Rips, L., Rasinski, K. (2000). *The Psychology of Survey Response*. Cambridge University Press: USA

Tsai, A. C., Tomlinson, M., Dewing, S., le Roux, I. M., Harwood, J. M., Chopra, M., Rotheram-Borus, M. J. (2014). Antenatal depression case finding by community health workers in South Africa: feasibility of a mobile phone application. *Archives of Women's Mental Health; 17:* 423 – 431. doi: 10.1007/s00737-014-0426-7

UK National Screening Committee (NSC) (2015a). *Programme appraisal criteria: Criteria for appraising the viability, effectiveness and appropriateness of a screening programme*. Available from: http://www.screening.nhs.uk/criteria [Accessed 28 January 2015]

UK National Screening Committee (NSC) (2015b). *The UK NSC recommendation on Postnatal depression screening in pregnancy.* Available from: http://www.screening.nhs.uk/postnataldepression [Accessed 10 February 2015]

Wells, T., Bailey, J. T., Link, M. W. (2014). Comparison of Smartphone and Online Computer Survey Administration. *Social Science Computer Review; 32(2):* 238 – 255. doi: 10.1177/0894439313505829

Whooley, M. A., Avins, L. A., Miranda, J., Browner, W. S. (1997). Case-finding instruments for depression. *Journal of General Internal Medicine; 12(7):* 439 – 445.

Woolhouse, H., Gartland, D., Mensah, F., Brown, S. J. (2014). Maternal depression from early pregnancy to 4 years postpartum in a prospective pregnancy cohort study: implications for primary health care. *BJOG: An International Journal of Obstetrics and Gynaecology; 122:* 312 – 321. doi: 10.1111/1471-0528.12837

Yazici, E., Kirkan, T. S., Aslan, P. A., Aydin, N., Yazici, A. B. (2015). Untreated depression in the first trimester of pregnancy leads to postpartum depression: high rates from a natural follow-up study. *Neuropsychiatric Disease and Treatment; 11:* 405 – 411

Appendix 1. Personal information and demographic survey

1. What is your age?
    a. 18 to 22 years
    b. 23 to 27 years
    c. 28 to 32 years
    d. 33 to 37 years
    e. 38 years or older
2. How would you describe your race/ethnicity?
    a. White
        i. British
        ii. Irish
        iii. Other
    b. Mixed
        i. White and Black Caribbean
        ii. White and Black African
        iii. White and Asian
        iv. Other
    c. Asian or Asian British
        i. Indian
        ii. Pakistani
        iii. Bangladeshi
        iv. Other
    d. Black or Black British
        i. Caribbean
        ii. African
        iii. Other
    e. Other
        i. Chinese
        ii. Other
    f. Not stated
3. What is your marital status?
    a. Single
    b. Married/In civil partnership
    c. Divorced/Civil partnership that has been dissolved
    d. Widowed
    e. Separated
    f. Not stated
4. What is your employment status?
    a. Employed, full-time
    b. Employed, part-time
    c. Self-employed
    d. Not employed, looking for work

   e. Not employed, not looking for work
   f. Disability/Not able to work
5. What is your highest level of education?
   a. University or college degree
   b. University or college qualification below degree level
   c. A Levels
   d. GCSE
   e. None of these
6. Do you own a smartphone (e.g., iPhone, Samsung, Blackberry, HTC, Sony)?
   a. Yes
   b. No
7. Do you own a tablet computer (e.g., iPad, iPad mini, Samsung tablet)?
   a. Yes
   b. No
8. Is this your first pregnancy?
   a. Yes
   b. No
9. (If answer to previous question is *No*) How many children have you given birth to?
10. When is your current baby due?
    a. Please indicate an approximate date
11. Have you ever been diagnosed with depression?
    a. Yes
    b. No

Appendix 2. EMA assessments

1. How do you feel?
    a. Excellent
    b. Good
    c. OK
    d. Not good
    e. Terrible
2. How rested do you feel?
    a. Very rested
    b. Somewhat rested
    c. Neither rested nor tired
    d. Somewhat tired
    e. Very tired
3. How worried do you feel?
    a. Extremely
    b. Very
    c. Moderately
    d. Slightly
    e. Not at all
4. How much are you enjoying yourself?
    a. A lot
    b. Quite a bit
    c. Somewhat
    d. A little bit
    e. Not at all
5. How energetic do you feel?
    a. Extremely
    b. Very
    c. Moderately
    d. Slightly
    e. Not at all

Appendix 3. Contextual Information

1. Where are you?
    a. At home
    b. At work
    c. At the shops
    d. At the park
    e. At a Café
    f. Other
        i. Please specify
2. What are you doing?
    a. Relaxing
    b. Working
    c. Shopping
    d. Walking
    e. Eating/Drinking
    f. Travelling
    g. Other
        i. Please specify

Appendix 4. Post-study acceptance survey

Please rate your agreement with the following statements with respect to this system:

1. This app is easy to use
    a. Strongly agree – Strongly disagree (7-point Likert Scale)
2. I learned to use this app quickly
    a. Strongly agree – Strongly disagree (7-point Likert Scale)
3. I would recommend this app to a friend
    a. Strongly agree – Strongly disagree (7-point Likert Scale)
4. I would repeat the experience of using this app
    a. Strongly agree – Strongly disagree (7-point Likert Scale)
5. I found using this app an engaging experience
    a. Strongly agree – Strongly disagree (7-point Likert Scale)
6. I found the assessments useful
    a. Strongly agree – Strongly disagree (7-point Likert Scale)
7. I was asked to provide reports:
    a. Much too infrequently
    b. Too infrequently
    c. The right amount
    d. Too frequently
    e. Much too frequently
8. The experience of using this app met my needs:
    a. Strongly agree – Strongly disagree (7-point Likert Scale)
9. What were your motivations in using this app?
10. What did you like the most about the experience of using this app?
11. What did you like least about the experience of using this app?
12. How would you improve the experience of using this app?
13. Do you have any other comments that you would like to make?

To be answered by those participants allocated to the retrospective plus momentary assessment strategy:

14. I found it useful to compare reports made *right now* with those made *over the past 7 days:*
    a. Strongly agree – Strongly disagree (7-point Likert Scale)

To be answered by those participants who are receiving feedback from a designated member of the clinical team:

15. I felt connected to my supporter:
    a. Strongly agree – Strongly disagree (7-point Likert Scale)
16. I found the feedback provided useful:
    a. Strongly agree – Strongly disagree (7-point Likert Scale)

Appendix 5. Whooley Questions

1. Over the past month, have you been bothered by feeling down, depressed or hopeless?
    a. Yes/No
2. Over the past month, have you been bothered by having little interest or pleasure in doing things?
    a. Yes/No

Appendix 6. Edinburgh Postnatal Depression Scale

As you are pregnant or have recently had a baby, we would like to know how you are feeling. Please check the answer that comes closest to how you have felt IN THE PAST 7 DAYS, not just how you feel today

In the past 7 days:

1. I have been able to laugh and see the funny side of things
    - As much as I always could
    - Not quite so much now
    - Definitely not so much now
    - Not at all
2. I have looked forward with enjoyment to things
    - As much as I ever did
    - Rather less than I used to
    - Definitely less than I used to
    - Hardly at all
3. I have blamed myself unnecessarily when things went wrong
    - Yes, most of the time
    - Yes, some of the time
    - Not very often
    - No, never
4. I have been anxious or worried for no good reason
    - No, not at all
    - Hardly ever
    - Yes, sometimes
    - Yes, very often
5. I have felt scared or panicky for no very good reason
    - Yes, quite a lot
    - Yes, sometimes
    - No, not much
    - No, not at all
6. Things have been getting on top of me
    - Yes, most of the time I haven't been able to cope at all
    - Yes, sometimes I haven't been coping as well as usual
    - No, most of the time I have coped quite well
    - No, I have been coping as well as ever
7. I have been so unhappy that I have had difficulty sleeping
    - Yes, most of the time
    - Yes, sometimes
    - Not very often
    - No, not at all
8. I have felt sad or miserable
    - Yes, most of the time
    - Yes, quite often
    - Not very often
    - No, not at all

9. I have been so unhappy that I have been crying
    - Yes, most of the time
    - Yes, quite often
    - Only occasionally
    - No, never
10. The thought of harming myself has occurred to me
    - Yes, quite often
    - Sometimes
    - Hardly ever
    - Never

Appendix 7. Post-study supporter survey

1. This system is easy to use:
    a. Strongly agree – Strongly disagree (7-point Likert Scale)
2. I learned to use this system quickly:
    a. Strongly agree – Strongly disagree (7-point Likert Scale)
3. I would recommend this system to a colleague:
    a. Strongly agree – Strongly disagree (7-point Likert Scale)
4. I would repeat the experience of using this system:
    a. Strongly agree – Strongly disagree (7-point Likert Scale)
5. I found using this system an engaging experience:
    a. Strongly agree – Strongly disagree (7-point Likert Scale)
6. I found the data provided by participant self-report useful:
    a. Strongly agree – Strongly disagree (7-point Likert Scale)
7. I found it useful to compare reports asking the participant how they feel right now with those asking how they felt over the past 7 days:
    a. Strongly agree – Strongly disagree (7-point Likert Scale)
8. I was asked to provide feedback:
    a. Much too infrequently
    b. Too infrequently
    c. The right amount
    d. Too frequently
    e. Much too frequently
9. I felt connected to participants:
    a. Strongly agree – Strongly disagree (7-point Likert Scale)
10. I felt that the feedback I provided was useful:
    a. Strongly agree – Strongly disagree (7-point Likert Scale)
11. The experience of using this system met my needs:
    a. Strongly agree – Strongly disagree (7-point Likert Scale)
12. What were your motivations in using this system?
13. What did you like most about the experience of using this system?
14. What did you like least about the experience of using this system?
15. How would you improve the experience of using this system?
16. Do you have any other comments you would like to make?