ROYAL
COLLEGE
OF MIDWIVES

ISSN: 2633-8408 March 2021 Vol.19 No.1

EVIDENCE BASED MIDWIFERY

Contents

ROYAL
COLLEGE
OF MIDWIVES

Evidence Based Midwifery

The Ockenden report: personal reflections on fetal monitoring and its place in modern midwifery care

Keywords: Ockenden report, COVID-19, fetal monitoring, midwifery skills, technology and evidence based midwifery, Evidence Based Midwifery

On 10 December 2020 Donna Ockenden (midwife) published emerging findings from the Independent Review of Maternity Services at the Shrewsbury and Telford Hospital NHS Trust where allegedly avoidable maternal and infant deaths had occurred. The early report was a clear indicator of the need for immediate action and this was after a review of 250 cases from a potential of 1862 cases. It was a sobering read and the urgency of the need for immediate action was loud and clear. As I read through the recommendations I was struck by the numerous references to fetal monitoring and in particular:

'4.22 Fetal heart rate (FHR) monitoring is an essential component of the safe management of labour ... The review team found significant problems with the conduct of intermittent auscultation and in the interpretation of CTG … [Cardiotocograph]'

'4.26 A mother, admitted in labour with a breech presentation, had inappropriate use of oxytocin for her long labour with CTG concerns …'

'4.27 A woman presented in labour at 39 weeks. There were CTG abnormalities in labour, which were not escalated …'

(Ockenden 2020).

The memories of hours of observation of induced labours and the use of the continuous cardiotocography (CTG) machine flooded my thoughts. In my early research exploring the role of the CTG machine in the lifeworld of technologically supported midwifery practice, its place was securely positioned as a necessity for use in births where induction of labour was the status quo (Sinclair 1999). In 2009, 10 years later, the polarised opinions of midwives in midwife-led care settings and those in hospital settings were apparent — with the latter more favourably disposed towards the use of technology. It would be good to replicate the study in 2021.The role and model of the CTG machine has changed considerably during the last 20 years with much more sophisticated and less bulky fetal monitoring devices available off the shelf and, recently, more computerised software additions have received a favourable review (Judd et al 2020).

With the increased availability of clinical/health/hospital video consultations using downloadable desktop and mobile apps, such as PEXIP, the opportunities for remote monitoring of pregnant and labouring women during COVID-19 have escalated. Video consultation technology offers enhanced care for women and their babies and facilitates midwives to care for women while reducing the risks associated with attending maternity clinics. However, in reminiscing about the old, large, bulky CTG models, I still hold fond memories of the interviews with midwifery managers when exploring their perceptions of the arrival of the CTG machine in the 1970s. The image of one midwifery manager, in particular, is embedded in my memory as she smoked her way through the interview while she gave the most beautifully illustrated description of her memory of the CTG machine's glorious arrival to the labour ward and how it was greeted with respect and anticipation when it was proudly wheeled down the ward by the doctor shrouded in a white sheet. Mothers were delighted with the new high technology and enjoyed hearing their baby's heartbeat. The CTG machine was welcomed by midwives because it was going to be an important tool in the midwifery decision making process and it was going to enable doctors to make clinical decisions to intervene early to reduce neonatal mortality (Sinclair 1999).

The multi-professional team's expectations of the machine's capabilities were, however, beyond its capacity to deliver and the situation today has not changed. Although innovation in design has overcome some of the restrictions to a woman's need to be mobile during labour when monitoring is necessary, the CTG machine is limited by the product hardware

design and the functions of the program software. It is, therefore, impossible for any CTG machine to be 100 per cent accurate in determining abnormal fetal heart rate patterns and alerting clinical midwives and doctors to consider early/instrumental intervention. Even the more sophisticated high-tech models cannot be programmed to take into consideration every human variable likely to impact on fetal well-being, including gestation, fetal weight, fetal anomalies, maternal health and social issues, familial traits, medication and drug history. The human decision making remains the same and I see similarities in expectations of novice researchers who anticipate that by loading up their masses of interview data into a software package, such as NVIVO, the machine will magically produce a perfect data analysis output. How wrong they are!

Every doctoral midwife using software needs to be aware of the limitations of the machine to interpret results in context. The package is excellent for managing large volumes of qualitative data. You put the data in and you direct the action of the program to produce the data output and YOU are the one who must make the interpretation judgement when the software churns out patterns and nodes and clusters of data. The data has to be analysed in context and a researcher has to bear in mind the original research questions, the aim of the study and the specific objectives framing the data entries. You, as the researcher, have already imposed a framework of pre-determined questions for interviews and the data needs to be interpreted bearing this factor in mind.

When we transfer our thinking to the outputs from the CTG machine, we need to be cognisant of the fact that the software program has limitations and will be designed using data available from previously screened mothers' CTG readings. Therefore, an ability to discriminate between different clinical conditions, ethnic origins and social groupings cannot be expected at this time. The data produced is limited to each mother and baby dyad. The newer machines, designed to provide clinical alerts using a traffic light system, are also limited even though they have more sophisticated artificial intelligence. The machine cannot give a 360-degree holistic analysis of maternal and fetal well-being and it is not yet designed to discriminate for social groupings, women with medical issues or ethnic factors. Fetal heart rate is indeed a key variable but on its own, without the fetal PH and the maternal pulse, BP and temperature it is severely limited.

Data from the CTG machine is just one segment of the full picture and needs to be viewed with that lens; our expectations need to be realistic. Based on the recommendations for immediate action with regard to appropriate and effective fetal monitoring from the Ockenden (2020) report, there is a continued need for multi-professional certified competence in CTG application and interpretation with support for colleagues undertaking fetal well-being monitoring. I would go further to say that midwives and obstetricians need to become more involved in the actual design of the software to ensure the sensitivity and specificity of the future computerised programs are enhanced.

Many times I have been in discussion with academic and clinical colleagues on the role of the CTG machine as a helpful tool for midwives and a source of reassurance for mothers and fathers and a useful learning resource for student midwives. I hear, again and again, that the machine leads to increased caesarean section and instrumental delivery based on Cochrane review (Alfirevic et al 2017) but many fail to critically read the authors' conclusions:

'The question remains as to whether future randomised trials should measure efficacy (the intrinsic value of continuous CTG in trying to prevent adverse neonatal outcomes under optimal clinical conditions) or effectiveness (the effect of this technique in routine clinical practice).' (Alfirevic et al 2017:1).

The effectiveness of the technique is in our domain of practice. We are also the main interpreters of the data outputs and we are the people who initiate active interventions based on our judgement of the evidence presented.

I find myself asking the question: 'Have we lost our skill in CTG interpretation or are we so focused on normality that we are unable to conceive of fetal monitoring as an essential component of our skill set?' Furthermore: 'Are we listening to women who request continuous fetal monitoring?' I have spoken to mothers who had previous stillbirths and, for them and their partners, the CTG machine provided evidence that their baby was alive. Many women have purchased various fetal heart 'recorders' over the internet. Stories from family and friends provide anecdotal data on how some women, and some midwives, feel the value of the CTG is not just in aiding clinical decision making but provides much-needed psychological support that women need even more during COVID-19. The threat of face-to-face contact is real and, in the current pandemic, three UK midwives and five mothers have already died from COVID-19 (Knight et al 2020, Cook et al 2020). The pandemic has definitely had a huge impact on the use of the internet and online platforms that are becoming more attractive to midwives and mothers as they provide several benefits, including acting as a safety net in reducing the risk of exposure to COVID-19. I do hope 2021 will be a year in which we use the full range of technologies available to us appropriately and effectively for the good of the mothers and babies we serve — whether that be online communication via Facebook, WhatsApp, Live Chat, Skype, Zoom or through the use of supportive, surveillance

technologies, such as the CTG machine, or alternative maternal and fetal monitoring equipment.

I am certain that the future of fetal monitoring in 2021 is going to have a major overhaul and we must implement the findings from the Ockenden report (2020):

'… all maternity services must appoint a dedicated lead midwife and lead obstetrician with expertise in the field of foetal monitoring in order to improve upon practice in foetal monitoring.'

Professor Marlene Sinclair (editor)

PhD, MEd, PGDip/Ed, BSc, RM.

Professor of Midwifery Research and Head of the Centre for Maternal, Fetal and Infant Research at Ulster University, Northern Ireland.

References

Alfirevic Z, Gyte GML, Cuthbert A, Devane D (2017). Continuous cardiotocography (CTG) as a form of electronic fetal monitoring (EFM) for fetal assessment during labour. Cochrane Database of Systematic Reviews 2017, Issue 2. Art. No.: CD006066. DOI: 10.1002/14651858.CD006066.pub3 [Accessed 6 January 2021].

Cook T, Kursumovic E, Lennane S (2020). 'Exclusive: deaths of NHS staff from covid-19 analysed', *Health Services Journal*, 22 April. https://www.hsj.co.uk/exclusive-deaths-of-nhs-staff-from-covid-19-analysed/7027471.article [Accessed 15 January 2021].

Judd FA, Haran SS, Everett TR (2020). Antenatal fetal wellbeing. *Obstetrics, Gynaecology and Reproductive Medicine* 30(7):197-204.

Knight M, Bunch K, Vousden N, Morris E, Simpson N, Gale C, O'Brien P, Quigley M, Brocklehurst P, Kurinczuk JJ (2020). Characteristics and outcomes of pregnant women hospitalised with confirmed SARS-CoV-2 infection in the UK: national population based cohort study. *BMJ* 369(m2107).

Ockenden D (2020). *Review of maternity services at Shrewsbury and Telford Hospital NHS Trust: Emerging findings and recommendations from the Independent Review of maternity services at the Shrewsbury and Telford Hospital NHS Trust*. https://www.gov.uk/government/publications/ockenden-review-of-maternity-services-at-shrewsbury-and-telford-hospital-nhs-trust [Accessed 11 December 2020].

Sinclair MK (1999). Midwives' readiness to use high technology in the labour ward: implications for education and training. Unpublished PhD thesis, Queen's University, Belfast.

Sinclair M (2009). Practice: a battlefield where the natural versus the technological. *Evidence Based Midwifery* 7(2):39.

Evidence Based Midwifery

The effectiveness of online psychotherapy interventions for the treatment of perinatal mental health disorders: a systematic review

Rachel Black[1], Marlene Sinclair, Paul William Miller, Bernie Reid, Julie McCullough, Paul Slater, Mara Tesler Stein, Derek Farrell

1. Corresponding author

Date submitted: 8 October 2020. **Date accepted:** 21 January 2021. **Date published:** 1 March 2021.

ABSTRACT

Background: There is a wide spectrum of perinatal mental health (PMH) disorders including anxiety, depression, stress, compulsive disorders, post-traumatic stress disorder (PTSD), mania and postpartum psychosis. Up to 20 per cent of women are found to be suffering from anxiety and or depression during pregnancy and the first three months postnatally. Maternal mental health can have an impact on the infant both antenatally and throughout the life course, with social, emotional and cognitive consequences. Therefore, there is a critical need for evidence-based interventions which address PMH disorders. At present, cognitive behavioural therapy (CBT) and eye movement desensitisation and reprocessing therapy (EMDR) are the most commonly used therapies. They remain the only psychotherapies for PTSD in the perinatal period approved by The National Institute for Health and Care Excellence (NICE 2014).

Objective: To assess the effectiveness of online psychotherapy (O-P-T) interventions for the treatment of PMH disorders.

Methods: To address the research question 'What is the effectiveness of online psychotherapy interventions for the treatment of perinatal mental health disorders?' the researchers developed eligibility criteria using the Cochrane PICOS framework. Five electronic databases were searched: PsycInfo, MEDLINE, CINAHL Complete, ProQuest Dissertations & Theses and Scopus, along with Google Scholar and the Shapiro Library, to identify papers which investigated O-P-T interventions published before July 2020.

Interventions were included if study participants had a clinician-assessed diagnosis of a PMH illness at screening and if the study had an experimental design with clinician involvement as part of the intervention. The included studies were assessed for quality using the standard quality assessment criteria for evaluating primary research papers from a variety of fields (QualSyst) (Kmet et al 2004).

Findings: Searches yielded 2567 results from the selected databases. Five studies were included in the review (209 participants), three were postnatal depression (PND) interventions and two were antenatal depression (AD) interventions. All were CBT-based. Quality ratings of the included studies were found to be high based on QualSyst scores, however treatment fidelity was not reported. Pooled effect sizes found a small to medium effect favouring the intervention versus control on the reduction of depressive symptoms, according to their score on the selected depression measures, and in some cases remission (d=0.48, 95% CI -0.07, 1.06). The results from the pooled effect sizes on available within-group data resulted in a large treatment effect for depression, anxiety and stress outcomes (d=1.90; d=0.81; d=1.05). Attrition rates were comparable with other online psychotherapy studies for mental health.

Conclusions: This review provides evidence that O-P-T interventions for the treatment of PMH disorders are effective in improving clinical outcomes in the reduction of depression, anxiety and stress. Despite NICE (2014) recommendations for the treatment of PTSD in the perinatal period there remains a gap in the literature for online EMDR interventions. In addition, there is a lack of research for O-P-T interventions delivered via different modalities, such as videoconferencing, and a scarcity of research into O-P-T for other PMH disorders.

This is the first synthesis of research into O-P-T for women who meet diagnostic criteria for PMH disorders as classified by the International Classification of Diseases (ICD-11) (World Health Organization (WHO) 2020) or the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5) (American Psychiatric Association 2013).

Keywords: perinatal, online, systematic review, mental health, EMDR, interventions, psychotherapy, Evidence Based Midwifery

Introduction

According to the International Classification of Diseases 11th Revision (ICD-11) (WHO 2020) the perinatal period covers pregnancy and up to a year after birth. For many women, this is a time of excitement and hope for the future. However, this is not the case for every family as some of the most common morbidities associated with this period are mental health disorders (Howard et al 2014). There is a wide spectrum of perinatal mental health (PMH) disorders including anxiety, stress, depression, compulsive disorders, post-traumatic stress disorder (PTSD), mania and postpartum psychosis. Up to 20 per cent of women are found to be suffering from anxiety and or depression during pregnancy and the first three months postnatally (O'Hara & Wisner 2014).

The psychological and economic costs associated with PMH disorders are substantial and the effects extend beyond the woman (Bauer et al 2014). Maternal mental health can have an impact on the infant both antenatally and throughout the life course, with social, emotional and cognitive consequences (Goodman 2019). Research has identified links between depression in the pregnancy period and obstetric complications such as low birth weight, increased risk of caesarean section and preterm birth (Huang et al 2017). Additionally, prenatal depression and prenatal anxiety have been associated with poor breastfeeding outcomes (Fallon et al 2016). PMH disorders have also been found to be the leading cause of suicide in pregnancy and in the first year after birth, with suicide in this period accounting for between five and 20 per cent of maternal deaths in high-income countries (Khalifeh et al 2016, Knight et al 2019). Consequently, the economic cost of perinatal anxiety, depression and psychosis combined are reported to amount to £8.1 billion per each one-year cohort of births in the UK, equating to almost £10,000 for every birth (Bauer et al 2014).

Given the impact PMH disorders have, it is critical that women receive timely treatment to ensure positive outcomes for both mother and infant. Research investigating women's intentions for help-seeking has identified that they are more inclined to seek informal help from partners or family than seek professional help (Fonseca & Canavarro 2017). This could be interpreted as due to the perceived barriers to treatment that women cite, such as lack of time, lack of childcare and the costs involved (Lara et al 2014, Masood et al 2015, Millett et al 2018). An equally salient barrier to help-seeking for PMH is the perception of stigma, feelings of guilt and shame. Women have been found to express a fear of being judged as a bad mother and to fear the prospect that their children could be taken away if they are deemed unable to cope (Millett et al 2018).

Of central concern, therefore, are ways in which barriers to accessing treatment/support can be ameliorated to provide women with a safe, effective and practical way to access help in the perinatal period. Recent research into online psychotherapy (O-P-T) has found that individuals report that the ability to access therapy from home can reduce the stigma associated with attending in person, eliminate travel costs (Sunjaya et al 2020) and alleviate childcare issues (Millet et al 2018).

There is a paucity of research looking at the efficacy of O-P-T. To the authors' knowledge there are currently no systematic reviews specifically examining the efficacy of O-P-T for women who have a clinical diagnosis of a perinatal disorder. Additionally, the recent impact of the COVID-19 pandemic has emphasised the importance of individuals being able to access online therapy. With a widespread increase in acceptance of telehealth from mental health professionals and the wider community (Wind et al 2020), conducting a review to assess the efficacy of O-P-T in this population is warranted. Therefore, the aim of this review was to provide a synthesis of current evidence for effective psychotherapeutic interventions delivered via online modalities for women with a PMH diagnosis.

Methods

Search strategy

The following databases were searched systematically on 26 March 2020: PsycInfo, MEDLINE, CINAHL

Complete, Proquest Dissertations & Theses, Scopus. In addition, searches in Google Scholar and the Shapiro Library were carried out. A combination of search terms and medical subject headings (MeSH) terms related to the topic were used, such as 'mental health', 'perinatal', 'psychotherapy', and 'online' to search article titles, abstracts and keywords. The list of search terms was adapted from a previous systematic review by Ashford et al (2016) with the help of a subject matter expert librarian. The Google Scholar search strategy involved taking the first 10 pages of results from a search of the following terms: 'perinatal AND mental health diagnosis AND online AND psychotherapy'. This search strategy was based on a previous systematic review which limited the number of Google Scholar pages used for selecting papers (Sinclair et al 2018). In addition, reference lists of the included papers were searched for any other eligible articles and search alerts were set up on the databases to provide any new relevant publications.

Selection process and eligibility criteria

Following the database searches, duplicates were removed prior to title and abstract screening for eligibility. The eligibility criteria were developed using the Cochrane PICOS framework (O'Connor et al 2008) (see Table 1).

Studies were excluded if interventions were a) self-guided without therapist involvement b) had pharmacological co-interventions and c) consisted of only telephone or SMS (text) therapy. Study protocols, qualitative studies and studies in which less than half of the population met the criteria were also not eligible for review. Agreement on eligibility for inclusion involved discussion with five of the authors at both abstract and full-text screening stages.

Study selection

The database searches yielded a total of 2566 papers; one more paper was identified from one of the authors' recent reading sources. Two thousand five hundred and nine titles and abstracts were screened after duplicates were removed. A total of nine full texts were screened, with five being deemed eligible for review. Study screening and selection followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidance (Moher et al 2009) (see Figure 1).

Data extraction

A data extraction form was devised to enable the following data to be extracted: authors, year, aims/objectives, study design, participant inclusion/exclusion criteria, recruitment, sampling, randomisation, blinding, intervention (including integrity, duration, content and therapist involvement), comparator, statistical methodology, outcome measures, and results (main outcomes, attrition, power analysis, treatment compliance).

After reviewing the data and considering the studies' methodological design, intervention designs and outcome measures, it was deemed too heterogeneous to conduct a meta-analysis and therefore the review takes a narrative approach.

Findings

Study characteristics

All studies were primary research papers published in peer-reviewed journals. There was variability in the sample sizes, study origins, recruitment methods and attrition rates. An overview of study characteristics can be found in the Supplementary information.

Study design

Four of the five studies used a parallel randomised controlled (RCT) design (Milgrom et al 2016, Pugh et al 2016, Forsell et al 2017, Wozney et al 2017) with the remaining paper adopting a quasi-experimental design without control group (Kim et al 2014). Each study assessed the efficacy of the intervention through measuring outcomes at varied timepoints.

Study sample

The total number of participants across the studies was 209, with sample sizes ranging from 12 to 62. Fifty-four of the 209 participants were women in the antenatal period and 155 were postnatal women. Attrition rates ranged from five per cent to 40 per cent from the point of randomisation or time 1 (T1) to final post-intervention timepoint. All participants were over the age of 19, however, due to heterogeneity in the reporting of participant age, it was not possible to determine the range. The recruitment method adopted by the majority of studies was self-referral through advertising (Milgrom et al 2016, Pugh et al 2016, Forsell et al 2017). Two studies used a combination

Table 1. PICOS framework.

PICOS	Description
Population	Women in the perinatal period (start of pregnancy to one year postpartum) who have been diagnosed with a perinatal mental health disorder as defined by the ICD-11/DSM-5.
Interventions	Evidence-based online (computer/web-based) psychotherapies to treat mental health issues in the perinatal period.
Comparisons	Face-to-face treatment, treatment as usual (TAU), waitlist, pre-test/post-test outcome results.
Outcomes	Improved clinical outcomes, adverse effects, attrition rates.
Study design	Experimental designs e.g. randomised controlled trials, pre-post quasi-experimental design.

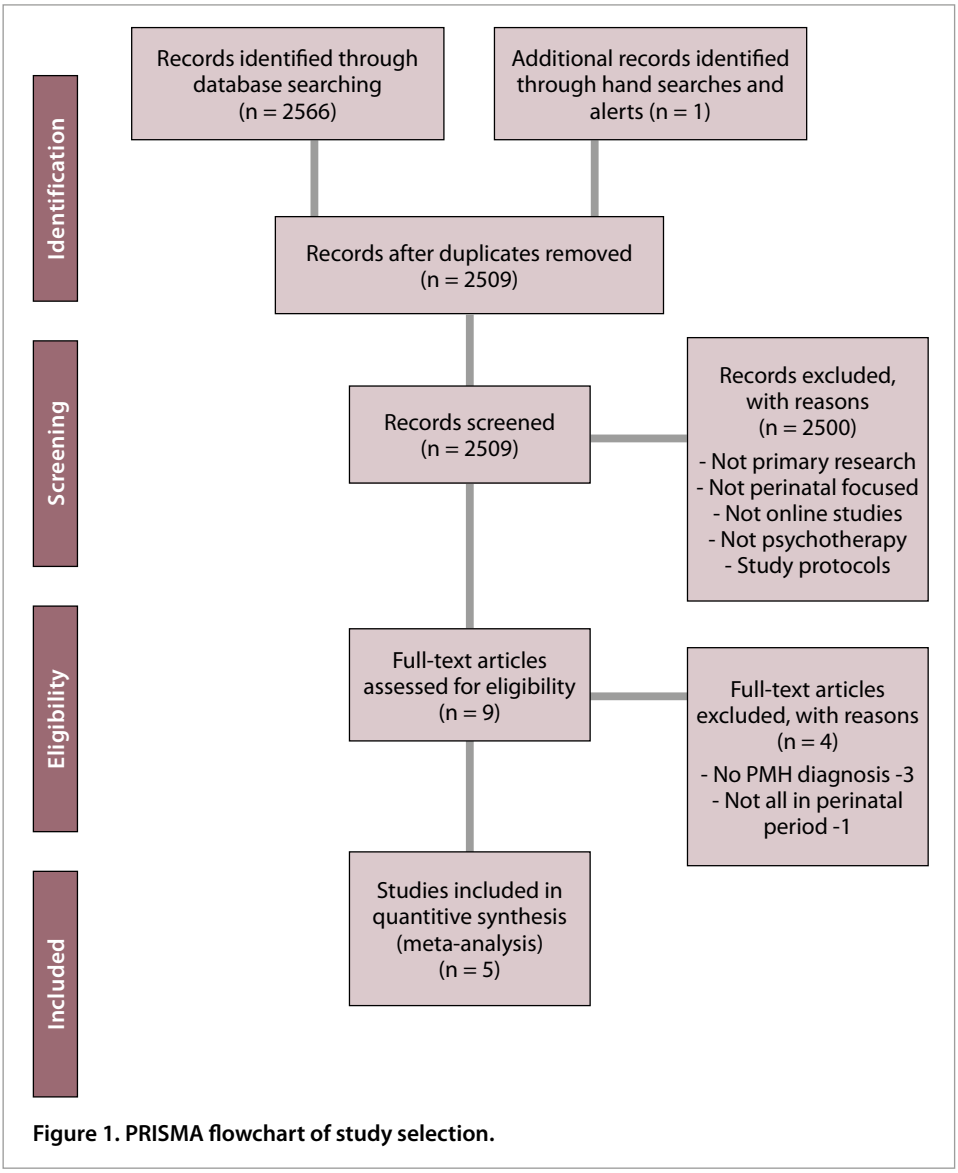


Figure 1. PRISMA flowchart of study selection.

of self-referral and referral through a health care professional (Kim et al 2014, Wozney et al 2017). Out of the participant samples, 155 were either married, co-habiting, engaged or in a committed relationship.

Methodological quality assessment

Study quality of the included papers was assessed using the standard quality assessment criteria for evaluating primary research papers from a variety of fields (QualSyst) (Kmet et al 2004). This is a 14-item checklist rated on a four-point scale: yes=2, partial=1, no=0, and N/A. With a total possible sum score out of 28 (less *2 for checklist items not applicable to the study design), summary scores were calculated by dividing the total sum score with the total possible score. This provided the authors with an assessment score of methodological quality in which a higher score represented higher methodological quality. In order to ensure rigour, quality assessment was assessed independently by five of the authors and discussed until agreement was reached.

Overall the studies were all found to score highly on the methodological rating according to QualSyst (Kmet et al 2004). Table 2 provides a breakdown of the ratings each study was allocated by the reviewers on each of the 14 criteria. The main limitations were found to be risk of response bias due to self-report scales, lack of blinding to investigators, inadequate sample size and the possible introduction of sampling bias through self-referral-only recruitment methods.

Intervention characteristics

The two PMH disorders targeted by the interventions were antenatal depression (AD) (Kim et al 2014, Forsell et al 2017) and postnatal depression (PND) (Milgrom et al 2016, Pugh et al 2016, Wozney et al 2017). All interventions used cognitive behavioural therapy (CBT) as a therapeutic approach to develop online self-guided modules with therapist support.

There was heterogeneity in the content of support and delivery method, with two interventions providing telephone support (Milgrom et al 2016, Wozney et al 2017), one intervention blending face-to-face (F2F) therapy with online modules (Kim et al 2014) and two delivering support through email communication (Pugh et al 2016, Forsell et al 2017). Duration and frequency of delivery varied from seven to 12 modules at a rate of one per week. There was substantial variability in the content of the online element of each intervention programme. More information on each intervention can be found in the Supplementary information.

Additionally, the comparison groups in each RCT study ranged from waitlist control group (WLC) (Pugh et al 2016) to treatment as usual (TAU) with information on depression (Wozney et al 2017), TAU with optional intervention following the study (Forsell 2017), TAU with no information or intervention offered (Milgrom et al 2016).

Table 2. Quality assessment of included studies.

	Kim et al 2014	Milgrom et al 2016	Pugh et al 2016	Forsell et al 2017	Wozney et al 2017
1) Question/objective sufficiently described?	2	2	2	2	2
2) Design evident/appropriate?	1	2	2	2	2
3) Participant selection described/appropriate?	2	1	1	1	2
4) Participant characteristics/input variables sufficiently described?	2	2	2	2	2
5) Random allocation described?	N/A	2	2	2	2
6) Investigator blinding reported?	2	1	1	1	2
7) Participant blinding reported?	N/A	N/A	N/A	N/A	N/A
8) Outcome measure well-defined/robust to measurement/misclassification bias?	1	1	1	1	1
9) Sample size appropriate?	0	1	1	0	1
10) Analysis described & appropriate	2	2	2	2	2
11) Estimate of variance reported for main results?	2	2	2	2	1
12) Controlled for confounding?	2	2	1	2	1
13) Results reported in sufficient detail?	2	2	2	2	2
14) Results support conclusions?	2	2	2	2	2
Total score	**18**	**22**	**21**	**21**	**22**
Sum score	**0.83**	**0.85**	**0.81**	**0.81**	**0.85**

Yes=2; Partial=1; No=0

Country of origin

The studies were undertaken in four different countries: two in Canada (Pugh et al 2016, Wozney et al 2017); one in Sweden (Forsell et al 2016); one in the United States of America (Kim et al 2014) and one in Australia (Milgrom et al 2016).

Mental health outcomes

Where possible, continuous measures were converted to a standardised effect size (Cohen's d) (Cohen 1988) using standardised mean difference; dichotomous data were converted from chi-squared results. This enabled the production of forest plots to report between-group (intervention versus control) effect sizes for mental health outcomes (see Figure 2) and within-group effects for mental health outcomes in the intervention group (see Figure 3). Where data were not available, study authors were contacted and if data were not provided the study outcomes were omitted from the forest plot.

As demonstrated in Figure 2, the results from the pooled effect sizes show that the majority of mental health outcomes favoured the intervention with the exception of DASS-Stress and DASS-Depression in the study by Pugh et al (2016). Reduction of depression symptoms and diagnosis was approaching a medium overall effect size (d=0.48 CI -0.07, 1.06), anxiety symptom reduction had a small overall effect (d=0.35 CI -0.36, 0.81) and total stress reduction saw a negative small effect size (d=-0.23 CI -0.08, 1.14).

The results from the pooled effect sizes on available within-group data show a large treatment effect for depression, anxiety and stress (d=1.90; d=0.81; d=1.05)

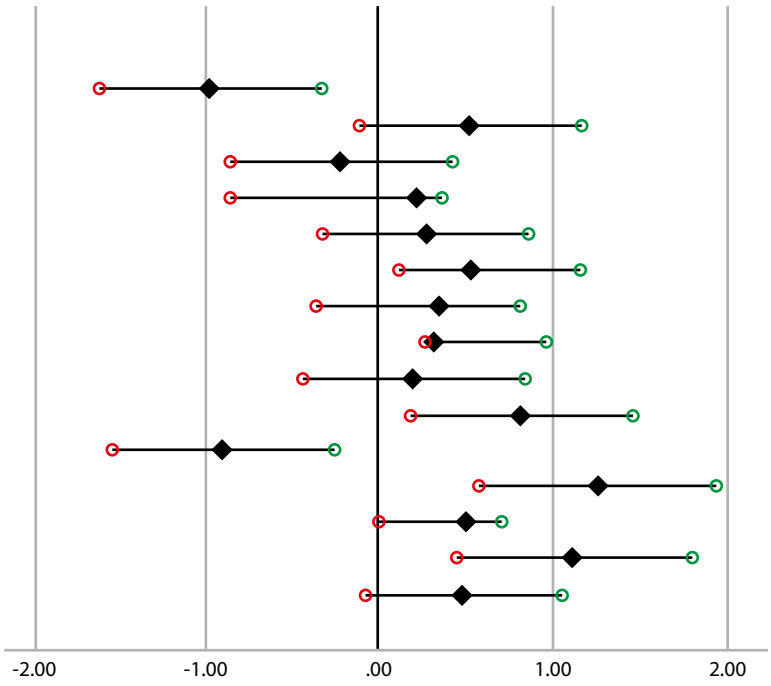
The primary outcome in all studies was depression, with each study using a number of different depression measures. Three out of five studies assessed clinical diagnostic status of depression at baseline using the SCID-I (Kim et al 2014, Forsell et al 2017, Wozney et al 2017) and one used the SCID-IV (Milgrom et al 2016). Pugh et al (2016) assessed diagnostic status at baseline using the MINI International Neuropsychiatric Interview.

Only two studies followed up by assessing post-intervention clinical diagnosis using the same clinician-administered instrument (Milgrom et al 2016, Wozney et al 2017). Remission rates in the RCT treatment groups post-intervention and follow up ranged from 33 per cent to just over 90 per cent (Milgrom et al 2016, Pugh et al 2016, Forsell et al 2017, Wozney et al 2017). In the quasi-experimental study by Kim et al (2014), 60 per cent of participants were deemed in remission.

Kim et al (2014) reported the Hamilton Depression Rating Scale (HDRS) as their primary outcome measure and found a large significant effect size (d=2.38, 95% CI 1.30 – 3.31) at eight weeks post-treatment. Forsell et al (2016) assessed depression using the Montgomery-Åsberg Depression Rating Scale-Self report version (MADRS-S) and found large significant between-group post-treatment effect sizes (g=1.21, 95% CI, 0.50- 2.92).

The Beck Depression Inventory (BDI and BDI-II) was reported in three studies (Kim et al 2014, Milgrom et al 2016, Wozney et al 2017). Kim et al (2014) reported a significant reduction in BDI scores at eight weeks with a large effect size (d=2.47, 95% CI 1.37- 3.42). Similarly, Milgrom et al (2016) found a significant reduction in between group BDI-II scores

Figure 2. Between group forest plot for mental health outcomes.



Study	Outcome measure	Effect size	CI lower	CI upper
Pugh et al (2016)	DASS-S	-0.98	-1.62	-0.33
Milgrom et al (2016	DASS-S	0.53	-0.08	1.14
	Total stress	-0.23	-0.85	0.41
Pugh et al (2016)	DASS-A	0.23	-0.85	0.38
Milgrom et al (2016)	DASS-A	0.27	-0.33	0.87
Forsell et al (2017)	GAD-7	0.54	0.1	1.18
	Total anxiety	0.35	-0.36	0.81
Forsell et al (2017)	EPDS	0.33	0.3	0.96
Pugh et al (2016)	EPDS	0.19	-0.43	0.81
Milgrom et al (2016)	BDI	0.83	0.2	1.45
Pugh et al (2016)	DASS-D	-0.9	-1.54	-0.26
Forsell et al (2017)	MADRS-S	1.26	0.57	1.95
Wozney et al (2017)	SCID	0.51	0	0.72
Milgrom et al (2016)	SCID	1.12	0.44	1.81
	Total depression	0.48	-0.07	1.06

Abbreviations

BDI: Beck Depression Inventory; DASS-S: Depression Anxiety Scale Short-form (stress); DASS-A: Depression Anxiety Scale Short-form (anxiety) DASS-D: Depression Anxiety Scale Short-form (depression); EPDS: Edinburgh Postnatal Depression Scale; GAD-7: Generalised Anxiety Disorder scale; MADRS-S: Montgomery-Åsberg Depression Rating Scale- Self report version; SCID: The Structured Clinical Interview for DSM-IV/ DSM-V.

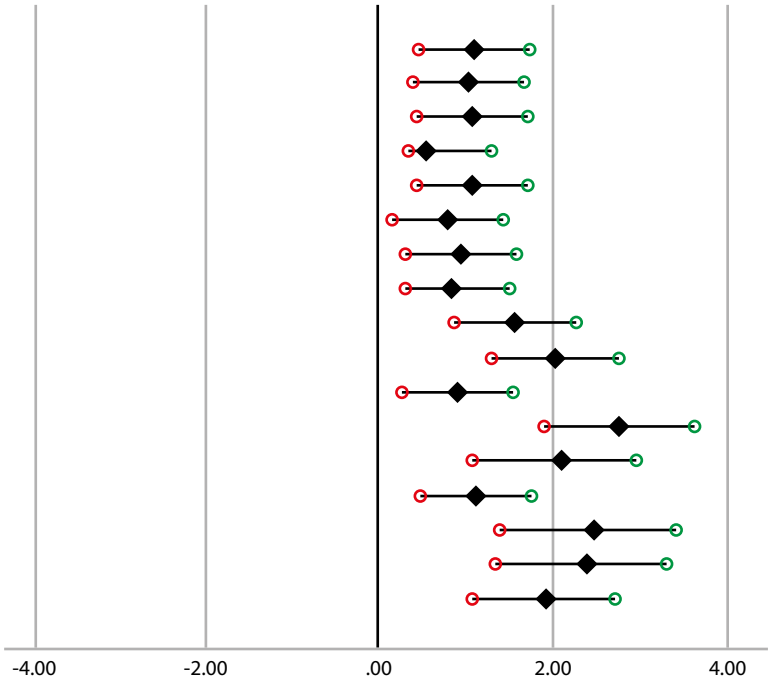
at 12 weeks post-treatment: *d*=0.83 (CI 95% 0.20 to 1.45). Wozney et al (2017) conducted a mixed effects regression on BDI score and reported a time by treatment interaction at six months (*B=- 0.99, t=1.885, p=0.064*) but the effect was not as large at 12 months (*B*=- 0.44, *t*=1.52, *p*=0.133) suggesting that a reduction in BDI-II scores could be temporary and there is a potential risk of depressive symptom relapse.

The Edinburgh Postnatal Depression Scale (EPDS) was administered in four studies with one study (Kim et al 2014) reporting a large significant effect size (*d*=2.08 95% CI 1.05-2.97), another (Forsell et al 2017) a non-significant medium effect (*g*=0.52, 95% CI, - 1.08 – 2.12) and Pugh et al (2016) finding no effect for condition on EPDS scores. However, in the study by Pugh et al (2016) a chi-squared analysis approached significance (*p*=0.08) on participants' classification of improved or recovered following the intervention based on reduction of EPDS scores. Participants in the intervention group were found to be 20 per cent improved and 62 per cent recovered in the interventions group. In the WLC group, 38 per cent of participants were classed as recovered and over 50 per cent exhibited no change. A regression analysis on time by interaction for EPDS scores resulted in non-significant reductions.

Depression ratings from the Depression Stress and Anxiety Scale Short-form (DASS21) were reported by Pugh et al (2016) using a multiple regression analysis which found the intervention lowered depression levels *B*=-0.25, *d*=-0.90 (-1.54, -0.26).

In addition to depression, four out of the five studies assessed anxiety as a secondary outcome (Kim et al 2014, Milgrom et al 2016, Pugh et al 2016, Forsell et al 2017). Three different anxiety scales were reported: Generalised Anxiety Disorder Scale (GAD-7), the Beck Anxiety Inventory (BAI) and the DASS anxiety. Three of the studies reported a reduction in anxiety symptoms ranging from a small to large effect, however, results were predominately non-significant (3/4); Kim et al (2014) reported a significant reduction in BAI score (*d*=0.5 CI-0.32, 1.28); Milgrom et al (2016) reported a non-significant reduction in between group scores on the DASS anxiety at 12 weeks (*d*=0.27, -0.33, 0.87); Forsell et al (2017) assessed anxiety using the GAD-7 and reported non-significant between group reductions (*g*=0.63 CI -0.84, 2.10) but significant within group reductions (*g*=0.93 CI 0.05, 1.81). Pugh et al (2016) conducted regression analysis on the DASS anxiety and stress scales between groups and found non-significant reduction in anxiety (*B*=-20) and a significant reduction in stress (*B*=-0.41). Milgrom et al (2016) also reported on the reduction of stress using the DASS and found a significant medium

Figure 3. Within group forest plot for mental health outcomes.



Study	Outcome measure	Effect size	CI lower	CI upper
Milgrom et al (2016)	DASS-S	1.07	0.43	1.72
Pugh et al (2016)	DASS-S	1.02	0.36	1.68
	Total stress	1.05	0.40	1.70
Kim et al (2014)	BAI	0.50	0.32	1.28
Forsell et al (2017)	GAD-7	1.02	0.39	1.65
Milgrom et al (2016)	DASS-A	0.76	0.14	1.39
Pugh et al (2016)	DASS-A	0.94	0.28	1.58
	Total anxiety	0.81	0.28	1.48
Pugh et al (2016)	DASS-D	1.54	0.83	2.26
Forsell et al (2017)	MADRS-S	2.01	1.28	2.75
Forsell et al (2017)	EPDS	0.88	0.26	1.51
Pugh et al (2016)	EPDS	2.76	1.88	3.64
Kim et al (2014)	EPDS	2.08	1.05	2.97
Milgrom et al (2016)	BDI	1.11	0.46	1.76
Kim et al (2014)	BDI	2.47	1.37	3.42
Kim et al (2014)	HRDS	2.38	1.30	3.31
	Total depression	1.90	1.05	2.70

Abbreviations

BAI: Beck Anxiety Inventory; BDI: Beck Depression Inventory; DASS-S: Depression Anxiety Scale Short-form (stress); DASS-A: Depression Anxiety Scale Short-form (anxiety) DASS-D: Depression Anxiety Scale Short-form (depression); EPDS: Edinburgh Postnatal Depression Scale; GAD-7: Generalised Anxiety Disorder scale; HDRS: Hamilton Depression rating scale; MADRS-S: Montgomery-Åsberg Depression Rating Scale - Self report version.

effect between groups at 12 weeks (*d*=0.53 95% CI, - 0.08 – 1.14).

Treatment engagement and adherence

Participant engagement and adherence varied due to differences in study design, however, overall the studies reported moderate to high engagement with the interventions.

Kim et al (2014) reported participants spent a mean average of 215.8 minutes on the 'Good Days Ahead' computer programme over the eight sessions and attended a mean of 7.2 in-person sessions.

Milgrom et al (2016) reported that 86 per cent of participants completed all sessions; the total time spent online averaged 370 minutes with participants completing a mean of 4.3 coaching calls out of a possible six.

Pugh et al (2016) reported 60 per cent completion of all seven possible modules in the intervention with the mean number of modules completed recorded as 5.92. Additionally, they reported high engagement with the website with the mean number of visits reported as 26.88, mean number of client emails sent reported as 5.4 and mean number of client emails received was 10.52 throughout the seven modules.

In the study by Forsell et al (2017) the researchers considered those participants that completed six out of the 10 modules as completers. They suggest this is because the first six modules contain the main components of CBT for depression. Therefore, they reported 82 per cent of participants as having completed six or more modules.

Finally, the study by Wozney et al (2017) reported that 22 intervention participants completed the full 12 coaching sessions and the mean average number of sessions over all the intervention participants was 9.13 sessions. The mean duration of coaching calls in the intervention was 36.74 minutes.

Discussion

The focus of this systematic review was to investigate the effectiveness of O-P-T for women with a clinical diagnosis of a PMH disorder. To the authors' knowledge it is the first to investigate online interventions specifically in this population. Five studies were included from a total of 2567 identified through systematic searches and included a total sample size of 209.

The interventions identified were found to have high heterogeneity in their design and varied in their focus

on women in the antenatal and postnatal periods. The studies explored the efficacy of the intervention with a number of mental health outcomes; however, the primary outcome of each study was depression. Despite the heterogeneity in intervention design, the results of this review suggest that O-P-T for the treatment of women classified as having a PMH disorder is clinically effective.

Evidence for this can be found in the overview of results as they demonstrate that, for the reduction of depressive symptomatology, and in some cases remission of depressive diagnosis, psychotherapy delivered online provided positive between-group treatment effects. It can be seen from the included forest plots that each of the diagnostic mental health scales were found to favour the intervention.

This is important as it builds upon the findings of the previous systematic review by Ashford et al (2016) who found depressive symptomatology to be reduced. However, they noted that the use of self-report measures made it unclear as to whether similar results would be found in women with diagnosed mental health issues. This review provides evidence for the effective reduction of clinical diagnostic symptomatology in women with perinatal depression.

In addition to depression, the review also highlighted the potential of O-P-T to be effective in the treatment of other PMH symptoms, such as anxiety and stress. The results were mixed in terms of statistical significance; however, overall treatment effects favoured the intervention, with only the study by Pugh et al (2016) found to elicit between-group outcome measures favouring the control.

Within-group data also provided a strong case for the efficacy of O-P-T with all outcomes favouring the intervention. Unfortunately, as the authors were unable to get the required data to complete within-group analysis on the results from the SCID outcome measures in Milgrom et al (2017) and Wozney et al (2017) they were not included in the forest plot. It is, however, posited that they would have elicited a favourable response toward the intervention based on the reported remission statistics.

Study quality

Quality assessment found the studies to have high quality ratings, however, there were notable drawbacks in a number of areas. Three studies (Milgrom et al 2016, Pugh et al 2016, Forsell et al 2017) were found to be inadequate in their sampling methods as they relied on self-referrals only which was deemed high risk for self-referral bias. In contrast, two of the studies (Kim et al 2014, Wozney et al 2017) included professional referrals though health care providers, thereby reducing the risk of bias.

Of the five studies, three reported that blinding of investigators was not possible due to the nature of psychotherapy (Milgrom et al 2016, Pugh et al 2016, Forsell et al 2017). However, the two studies conducted by Kim et al (2014) and Wozney et al (2017) reduced the risk of bias through the blinding of evaluators/independent assessment assistants.

All studies were found to have inadequate sample sizes which impacts on the ability to generalise the findings. This is of particular relevance to those studies which completed regression analysis. It is suggested that regression analysis requires a sample size of n=50 + (8 x number of variables) (Tabachnick & Fidell 2013) therefore, as both Pugh et al (2016) and Wozney et al (2017) had sample sizes below n=50, the results of their regression cannot be considered generalisable.

An equally significant aspect was the use of validated self-report scales in all the studies; although their use is commonplace in research it can leave the findings open to self-report bias and an under- or over-estimation of symptoms (Gorber & Tremblay 2016). Despite this criticism, two of the studies (Milgrom et al 2016, Wozney et al 2017) also measured clinical diagnosis outcome pre- and post-intervention using the clinician-administered SCID which provides a more robust form of measurement; therefore findings can be more readily accepted.

Study design

Of particular salience is the heterogeneity of study design. Three out of four of the RCT studies used a variation of treatment as usual, however, Pugh et al (2016) utilised a waitlist control design as a comparator. It has been suggested that the use of a waitlist control can overinflate the estimate of the intervention effect. Cunningham et al (2013) postulate that stalling participants in a 'non-action' state, whereby they do not attempt behaviour change or seek alternative help while waiting for the intervention, could elicit less improvement than if a control group had been used. Interestingly, the opposite can be seen in the study by Pugh et al (2016), who noted that WLC participants reported seeking regular contact with a family physician in almost 58 per cent of cases; 32 per cent sought psychotherapy and 25 per cent used psychotropic medication.

Research looking at help-seeking in the perinatal period has often found women to be less likely to seek help due to barriers such as perceived stigma or the lack of identifying PMH problems (Button et al 2017), therefore it is postulated that the screening procedure which involved a diagnostic outcome facilitated the removal of such barriers and provided opportunity for treatment. The impact of TAU group participants' engagement of outside sources of help is also of salience. Milgrom et al (2016) noted that 81 per cent of participants in the TAU group reported using one or more sources of outside support. Only one intervention group participant (5%) reported the use of antidepressant medication whilst involved in the trial compared to four (19%) in the TAU. Similarly, Forsell et al (2017) reported that eight

women in each study arm had begun counselling or psychological treatment during the course of the study. Consequently, this may have limited the detection of true treatment effects and should be considered when interpreting results.

Another important aspect to consider is the length of follow up; only one study (Wozney et al 2017) analysed treatment effects beyond three months. Their results found a significant treatment effect with treatment group participants 5.2 times more likely to experience diagnostic remission at 12 months compared to 1.2 times at three months. This suggests that designs which include a longer follow up period could identify larger treatment effects. This supports previous research which has found that the probability of recovering from PND increases by more than 50 per cent between six and 12 months (Torres et al 2019).

When considering the timescale for remission it is reasonable to look at other biological factors which may impact on treatment response in the perinatal period. Due to low oestrogen levels it has been suggested that there may be resistance to treatment for PND (Robakis & Williams 2013). Additionally, there is emerging evidence of the association between maternal nutrient depletion and increased risk of PMH issues. The use of prebiotics as a mediator to enhance the maternal microbiome is suggested to be a way to support maternal mental health by way of the hypothalamic-pituitary-adrenal (HPA) axis (Rackers et al 2018). When added to the associated demands of motherhood such as extreme fatigue, possible recovery following caesarean section and the care of existing children it is not unreasonable to expect a delay in recovery. In the study by Pugh et al (2016) between-group analysis was only assessed at four weeks post-intervention, therefore the suggestion of a recovery delay may account for the impact on their results.

Attrition rates

A notable finding was that attrition rates ranged from low to moderate (five per cent to 40 per cent), depending on the length of follow up and the design of the intervention. This variation suggests that, although the interventions are deemed effective in terms of clinical outcomes, the perceived acceptability for participants may vary between the intervention design with the lowest attrition found in the studies which adopted an RCT design and TAU control (Milgrom et al 2016, Forsell et al 2017, Wozney et al 2017). Previous systematic reviews which investigate various online modalities in other mental health populations have also found varied attrition rates ranging from 0 per cent–92 per cent (Bee et al 2008, Berryhill et al 2019, Bennett et al 2020). This emphasises the importance of intervention design and suggests that a PMH population has a similar treatment engagement with O-P-T as other mental health populations.

Treatment fidelity and demographics

An important aspect to consider when looking at psychological interventions is treatment fidelity. As all the online interventions were computer-based modules, the fidelity of the CBT treatment was more transparent for four out of five of the studies which provided detailed information on the module content and clinician involvement (Kim et al 2014, Milgrom et al 2016, Forsell et al 2017, Wozney et al 2017). However, it is noted that treatment fidelity was not assessed and reported in the studies and this can have a negative impact (Waltman et al 2017).

A pertinent point to note regarding the demographic information included in the studies was the lack of reporting on breastfeeding status, with just one study reporting it in between-group participant characteristics (Pugh et al 2016). Breastfeeding has been found to reduce the severity of PND and anxiety as increased prolactin levels during breastfeeding regulate the stress response and modulate anxiety and depressive-like behaviours (Kendall-Tackett 2007). Conversely, poor breastfeeding experiences and short breastfeeding duration are associated with higher depression scores (Brown et al 2016). This is not only pertinent for studies which look at postnatal disorders but, as some women breastfeed previous children throughout pregnancy, it is relevant for studies focusing on AD.

Implications for future research

This review highlights a gap in the literature for research which explores the use of a wide range of therapeutic approaches using online modalities. Despite a plethora of research exploring the use of eye movement desensitisation and reprocessing therapy (EMDR) for depression there were no studies which addressed this area. Current NICE guidelines (CG192, section 1.9.5) state that CBT and EMDR are the only approved therapies for PTSD in the perinatal period (NICE 2014).

Additionally, with the impact of the restrictions on F2F therapy due to the COVID-19 pandemic it is pertinent to consider research into different online modalities, such as videoconferencing, to enable women to gain access to a range of treatments from home. Furthermore, it was evident from the small number of included studies, that there is a need for further research into O-P-T for other diagnosed PMH disorders beyond depression. Indeed, there is a lack of studies which look at online interventions for women diagnosed with PTSD following childbirth, perinatal OCD, perinatal psychosis and tokophobia.

Strengths and limitations

There are a number of notable strengths of the current review. It adopted a clear search strategy and inclusion criteria which followed PRISMA guidelines and included dissertations and theses. Study eligibility and inclusion were assessed by five of the authors to ensure methodological rigour and minimise bias.

Moreover, study quality was assessed individually and subsequently discussed as a group to ensure agreement.

It is, however, important to note the limitations of this review. As the number of included studies was small, and high in heterogeneity of intervention design, it was not possible to complete a full meta-analysis with findings hard to synthesise. The small number of included studies and their small sample sizes limited the strength of evidence of this review.

In addition, it is important to consider the impact of publication bias. It is often the case that studies which produce non-significant results are less likely to be published leading to an over-inflation of the review's positive outcomes (Lefebvre et al 2019). Additionally, as resources were not available to translate non-English papers the review could be subject to language bias and the possibility of studies not included, which may reduce the ability to generalise the results (Lefebvre et al 2019). However, evidence has suggested that for the majority of systematic reviews the exclusion of non-English language studies has not affected the review conclusions (Morrison et al 2012, Hartling et al 2017).

Conclusion

This systematic review has demonstrated that O-P-T interventions for the treatment of PMH disorders are effective in the reduction of mental health symptomatology and in some cases may result in the remission of depressive diagnosis. Findings also suggest perinatal populations are likely to engage well with online treatment modalities. More research is required for interventions that not only assess different therapeutic approaches such as EMDR but also different online modalities for a broader range of PMH disorders.

Authors

Rachel Black MSc, PhD researcher, Ulster University (corresponding author).

Email: Black-R14@ulster.ac.uk

Prof Marlene Sinclair, Professor of Midwifery, Head of Maternal, Fetal and Infant Research Centre, Ulster University.

Prof Paul William Miller, Consultant Psychiatrist, EDMR Consultant Trainer, visiting professor within the School of Nursing at Ulster University.

Dr Bernie Reid, Lecturer in Nursing, School of Nursing, Ulster University.

Dr Julie McCullough, Research Associate, School of Nursing, Ulster University.

Dr Paul Slater, Lecturer in Statistics, School of Nursing, Ulster University.

Dr Mara Tesler Stein, Clinical Psychologist, The Touchstone Institute, Lincolnwood, IL, USA.

Dr Derek Farrell MBE, Clinical Psychologist, Principal Lecturer and Course Leader, MSc EMDR Therapy, University of Worcester.

References

American Psychiatric Association (2013). *Diagnostic and statistical manual of mental disorders (DSM-5)*. 5th ed. Philadelphia, PA: American Psychiatric Association.

Ashford MT, Olander EK, Ayers S (2016). Computer- or web-based interventions for perinatal mental health: a systematic review. *Journal of Affective Disorders* 197:134-46.

Bauer A, Parsonage M, Knapp M, Iemmi V, Adelaja B (2014). *The costs of perinatal mental health problems*. London: Centre for Mental Health and London School of Economics. https://www.centreformentalhealth.org.uk/sites/default/files/2018-09/costsofperinatal.pdf [Accessed 19 January 2020].

Bee PE, Bower P, Lovell K, Gilbody S, Richards D, Gask L, Roach P (2008). Psychotherapy mediated by remote communication technologies: a meta-analytic review. *BMC Psychiatry* 8(60). https://doi.org/10.1186/1471-244X-8-60 [Accessed 13 November 2020].

Bennett CB, Ruggero CJ, Sever AC, Yanouri L (2020). eHealth to redress psychotherapy access barriers both new and old: a review of reviews and meta-analyses. *Journal of Psychotherapy Integration*, 30(2):188-207.

Berryhill MB, Culmer N, Williams N, Halli-Tierney A, Betancourt A, Roberts H, King M (2019). Videoconferencing psychotherapy and depression: a systematic review. *Telemedicine Journal and e-Health* 25(6):435-46.

Brown A, Rance J, Bennett P (2016). Understanding the relationship between breastfeeding and postnatal depression: the role of pain and physical difficulties. *Journal of Advanced Nursing* 72(2):273-82.

Button S, Thornton A, Lee S, Shakespeare J, Ayers S (2017). Seeking help for perinatal psychological distress: a meta-synthesis of women's experiences. *British Journal of General Practice* 67(663):e692-9.

Cohen J (1988). *Statistical power analysis for the behavioral sciences*. 2nd ed. Hillsdale, NJ: Lawrence Erlbaum Associates.

Cunningham JA, Kypri K, McCambridge J (2013). Exploratory randomized controlled trial evaluating the impact of a waiting list control design. *BMC Medical Research Methodology* 13(150).

Fallon V, Groves R, Halford JCG, Bennett KM, Harrold JA (2016). Postpartum anxiety and infant-feeding outcomes: a systematic review. *Journal of Human Lactation* 32(4):740-58.

Fonseca A, Canavarro MC (2017). Women's intentions of informal and formal help-seeking for mental health problems during the perinatal period: the role of perceived encouragement from the partner. *Midwifery* 50:78-85.

Forsell E, Bendix M, Holländare F, von Schultz BS, Nasiell J, Blomdahl-Wetterholm M, Eriksson C, Kvarned S, van der Linden JL, Söderberg E, Jokinen J, Wide K, Kaldo V (2017). Internet delivered cognitive behavior therapy for antenatal depression: a randomised controlled trial. *Journal of Affective Disorders* 221:56-64.

Goodman JH (2019). Perinatal depression and infant mental health. *Archives of Psychiatric Nursing* 33(3):217-24.

Gorber SC, Tremblay MS (2016). Self-report and direct measures of health: bias and implications. In: Shephard R, Tudor-Locke C *eds. The objective monitoring of physical activity: contributions of accelerometery to epidemiology, exercise science and rehabilitation*. Cham: Springer: 369-76. http://doi-org-443.webvpn.fjmu.edu.cn/10.1007/978-3-319-29577-0_14 [Accessed July 2020].

Hartling L, Featherstone R, Nuspl M, Shave K, Dryden DM, Vandermeer B (2017). Grey literature in systematic reviews: a cross-sectional study of the contribution of non-English reports, unpublished studies and dissertations to the results of meta-analyses in child-relevant reviews. *BMC Medical Research Methodology* 17(64).

Howard LM, Molyneaux E, Dennis C-L, Rochat T, Stein A, Milgrom J (2014). Non-psychotic mental disorders in the perinatal period. *Lancet* 384(9956):1775-88.

Huang H-C, Sung F-C, Chen P-C, Chang CY-Y, Muo C-H, Shiue H-S, Huang J-P, Li T-C, Tzeng Y-L, Wu S-I (2017). Obstetric outcomes in pregnant women with and without depression: population-based comparison. *Scientific Reports* 7(13937).

Kendall-Tackett K (2007). A new paradigm for depression in new mothers: the central role of inflammation and how breastfeeding and anti-inflammatory treatments protect maternal mental health. *International Breastfeeding Journal* 2:6.

Khalifeh H, Hunt IM, Appleby L, Howard LM (2016). Suicide in perinatal and non-perinatal women in contact with psychiatric services: 15 year findings from a UK national inquiry. *Lancet Psychiatry* 3(3):233-42.

Kim DR, Hantsoo L, Thase ME, Sammel M, Epperson CN (2014). Computer-assisted cognitive behavioral therapy for pregnant women with major depressive disorder. *Journal of Women's Health* 23(10):842-8.

Kmet LM, Lee RC, Cook LS (2004). *Standard quality assessment criteria for evaluating primary research papers from a variety of fields*. Calgary, AB. Alberta Heritage Foundation for Medical Research, University of Calgary.

Knight M, Bunch K, Tuffnell D, Shakespeare J, Kotnis R, Kenyon S, Kurinczuk JJ *eds*, on behalf of MBRRACE-UK (2019). *Saving lives, improving mothers' care - lessons learned to inform maternity care from the UK and Ireland Confidential Enquiries into Maternal Deaths and Morbidity 2015-17*. Oxford: National Perinatal Epidemiology Unit, University of Oxford. https://www.npeu.ox.ac.uk/downloads/files/mbrrace-uk/reports/MBRRACE-UK%20Maternal%20Report%202019%20-%20WEB%20VERSION.pdf [Accessed 19 August 2020].

Lara MA, Navarrete L, Nieto L, Berenzon S (2014). Acceptability and barriers to treatment for perinatal depression: an exploratory study in Mexican women. *Salud Mental* 37(4):293-301.

Lefebvre C, Glanville J, Briscoe S, Littlewood A, Marshall C, Metzendorf M-I, Noel-Storr A, Rader T, Shokraneh F, Thomas J, Wieland LS (2019). Chapter 4: Searching for and selecting studies. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA eds. *Cochrane handbook for systematic reviews of interventions*. https://training.cochrane.org/handbook/current/chapter-04 [Accessed 4 July 2019].

Masood Y, Lovell K, Lunat F, Atif N, Waheed W, Rahman A, Mossabir R, Chaudhry N, Husain N. (2015). Group psychological intervention for postnatal depression: a nested qualitative study with British South Asian women. *BMC Women's Health* 15(109).

Milgrom J, Danaher BG, Gemmill AW, Holt C, Holt CJ, Seeley JR, Tyler MS, Ross J, Ericksen J (2016). Internet cognitive behavioral therapy for women with postnatal depression: a randomized controlled trial of MumMoodBooster. *Journal of Medical Internet Research 18*(3):e54.

Millett L, Taylor BL, Howard LM, Bick D, Stanley N, Johnson S (2018). Experiences of improving access to psychological therapy services for perinatal mental health difficulties: a qualitative study of women's and therapists' views. *Behavioural and Cognitive Psychotherapy* 46(4):421-36.

Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLOS Medicine* 6(7):e1000097. https://doi.org/10.1371/journal.pmed.1000097 [Accessed 13 November 2020].

Morrison A, Polisena J, Husereau D, Moulton K, Clark M, Fiander M, Mierzwinski-Urban M, Clifford T, Hutton B, Rabb D (2012). The effect of English-language restriction on systematic review-based meta-analyses: a systematic review of empirical studies. *International Journal of Technology Assessment in Health Care* 28(2):138-44.

National Institute for Health and Care Excellence (2014). *Antenatal and postnatal mental health: clinical management and service guidance*. Clinical guideline CG192 section 1.9.2. Updated 11 February 2020. https://www.nice.org.uk/guidance/cg192/chapter/1-Recommendations#treating-specific-mental-health-problems-in-pregnancy-and-the-postnatal-period. [Accessed 12 June 2020].

O'Connor D, Green S, Higgins JPT *eds* (2008). Chapter 5: Defining the review question and developing criteria for including studies. In: Higgins JPT, Green S *eds*. *Cochrane handbook of systematic reviews of intervention*. Version 5.0.1. https://training.cochrane.org/handbook/archive/v5.0.0/ [Accessed 13 November 2020].

O'Hara MW, Wisner KL (2014). Perinatal mental illness: definition, description and aetiology. *Best Practice and Research: Clinical Obstetrics and Gynaecology* 28(1):3-12.

Pugh NE, Hadjistavropoulos HD, Dirkse D (2016). A randomised controlled trial of therapist-assisted, internet-delivered cognitive behavior therapy for women with maternal depression. *PLoS ONE* 11(3):e0149186.

Rackers HS, Thomas S, Williamson K, Posey R, Kimmel MC (2018). Emerging literature in the microbiota-brain axis and perinatal mood and anxiety disorders. *Psychoneuroendocrinology* 95:86-96.

Robakis TK, Williams KE (2013). Biologically based treatment approaches to the patient with resistant perinatal depression. *Archives of Women's Mental Health* 16(5):343-51.

Sinclair M, Latos-Bieleńska A, McCullough J, Elliott E (2018). ConnectEpeople: a review of the online support organizations across the EU for parents with children who have Down syndrome, spina bifida, cleft lip and severe heart disease requiring surgery. *European Journal of Medical Genetics* 61(9):570-1.

Sunjaya AP, Chris A, Novianti D (2020). Efficacy, patient-doctor relationship, costs and benefits of utilizing telepsychiatry for the management of post-traumatic stress disorder (PTSD): a systematic review. *Trends in Psychiatry and Psychotherapy* 42(1):102-10.

Tabachnick BG, Fidell LS (2013). *Using multivariate statistics*. 6th ed. Boston, MA: Pearson Education.

Torres A, Gelabert E, Roca A, Navarro P, Plaza A, Subirà S, Martin-Santos R, Ascaso C, Garcia-Esteve L (2019). Course of a major postpartum depressive episode: a prospective 2 years naturalistic follow-up study. *Journal of Affective Disorders* 245:965-70.

Waltman SH, Sokol L, Beck AT (2017). Cognitive behavior therapy treatment fidelity in clinical trials: review of recommendations. *Current Psychiatry Reviews* 13(4):311-5.

Wind TR, Rijkeboer M, Andersson G, Riper H (2020). The COVID-19 pandemic: the 'black swan' for mental health care and a turning point for e-health. *Internet Interventions* 20:100317.

World Health Organization (WHO) (2020). *ICD-11 International classification of diseases 11th revision*. https://icd.who.int/ [Accessed 13 November 2020].

Wozney L, Olthuis J, Lingley-Pottie P, McGrath PJ, Chaplin W, Elgar F, Cheney B, Huguet A, Turner K, Kennedy J, Lingley-Pottie P, McGrath PJ, Chaplin W, Elgar F, Cheney B, Huguet A, Turner K, Kennedy J (2017). Strongest Families™ Managing Our Mood (MOM): a randomized controlled trial of a distance intervention for women with postpartum depression. *Archives of Women's Mental Health*, 20(4):525-37.

How to cite this paper:

Black R, Sinclair M, Miller PW, Reid B, McCullough J, Slater P, Stein MT, Farrell D (2021). The effectiveness of online psychotherapy interventions for the treatment of perinatal mental health disorders: a systematic review. *Evidence Based Midwifery* 19(1):6-18.

Supplementary information

Table 3. Study characteristics.

Study	Origin	Study design	Target problem	Therapeutic approach	Sample/recruitment	Intervention/duration	Comparator	Mental health outcome measures/times	Clinical outcomes/attrition rates
Kim et al (2014)	USA	Pre/post quasi experimental	AD	CBT	Pregnant women aged 18–49 (10–32 weeks' gestation) with MDD assessed by SCID HDRS score ≥ 14 Self- referrals through advertising/single site referral N=12	CCBT: 'Good Days Ahead' software alongside F2F CBT. 8 sessions over 6 weeks, 3.75 hours F2F therapy	N/A	*Primary outcome:* HDRS *Secondary outcomes:* BDI; BAI; EPDS *Outcomes measured at:* 1) baseline 2) after session 4 3) after session 8 4) three months follow up.	Significant improvement HDRS, EPDS, BAI, BDI p's = < 0.043* After session 8, 80% showed treatment response and 60% in remission Attrition=38%
Milgrom et al (2016)	AUS	Parallel RCT	PND	CBT	Australian resident women over 18 years, less than 1 year postnatal with diagnosis of PND assessed by SCID EPDS score of 11–23 & < 3 on #10 Self-referrals & advertising through maternal health centres N=43	MMB supported by low intensity telephone coaching & included access to online forum of participants 6 interactive sessions at 1 session per week Coaching max 30 mins per week	TAU – varied at the discretion of each participant's nominated health professional and was expected to include a heterogeneous mix of interventions/ supports	*Primary outcomes:* Remission of depressive episode & severity of symptoms SCID-IV at 12 weeks post treatment *Secondary outcomes:* BDI-ii DASS-21 anxiety; DASS-21 Stress; ATQ *Outcomes measured at:* 1) baseline 2) 9 weeks 3) 12 weeks	Primary outcome: MMB 79% (15/19) of women no longer met the DSM-IV criteria at 12 weeks vs 18% (4/22) in TAU p=0.001 Attrition=5%
Pugh et al (2016)	Canada	Parallel RCT	PND	CBT	Postnatal women over 18 years, living in Saskatchewan EPDS score ≥ 10 MINI interview to confirm sub-clinical and clinical symptoms Self-referral through advertising/ notifying healthcare professionals N=50	Maternal Depression Online: TA-ICBT comprising of questions at the beginning of each module and homework at the end. Therapist support through weekly emails 7 modules: suggested 1 per week	WLC, participants received leaflet on postnatal depression	*Primary outcome:* EPDS *Secondary outcomes:* DASS-21, PSI-SF *Outcomes measured at:* 1) baseline 2) post-treatment 3) 4 week follow-up	EPDS Between group: p=0.56 EPDS change over time: p=0.001* EPDS Condition by time: p=0.02* Approx. 20% of TA-ICBT participants were classified as improved, while over 62% were recovered. 38% classified recovered in WLC Attrition WLC=16% Intervention=20% Intervention T1 – follow up=40%

Study	Origin	Study design	Target problem	Therapeutic approach	Sample/recruitment	Intervention/duration	Comparator	Mental health outcome measures/times	Clinical outcomes/attrition rates
Forsell et al (2017)	Sweden	Parallel RCT	AD	CBT	Swedish women (over 18 years) between 10–28 weeks' gestation with MDD according to SCID-I MADRS-S between 15 and 35. Low risk of suicide indicated by score of 4 or less on MADRS-S. Self-referral: through various advertising outlets/ clinics N=42	ICBT: ICBT for depression used by the Internet Psychiatry Clinic Stockholm, adapted to include antenatal depression in pregnancy. Comprised of reading materials, assessments, homework and worksheets. CBT therapist feedback/ support through email. 10 modules over 10 weeks	TAU: continuation of their antenatal care for 10 weeks, followed by optional ICBT or ICBT given immediately as an add on to maternity care	*Primary outcome:* MADRS-S – remission defined as score below 13 *Secondary outcomes:* SCID-I, EPDS, GAD-7, *Outcomes measured at:* 1) baseline 2) post-treatment	MADRS-S between group post-treatment p=0.001** Remission observed in 33% ICBT participants vs 11% in TAU (not significant) Change of 8 points or more: 71% ICBT 22% TAU p=0.004* Attrition=7%
Wozney et al (2017)	Canada	Parallel RCT	PND	CBT	Women aged 19–45 with PPD living in Nova Scotia between 2006–2009 MDD with peripartum onset according to DSM-IV Self-referrals through advertising/referred through public health nurses & mental health professionals N=62	MOM: developed as part of 'strongest families distance system of services' consisting of handbook, video stories of PPD & corresponding exercises. Weekly telephone coaching involving manualised script 12 sessions – 1 weekly Booster session offered 1-month post-treatment	Control group: standard community care. Information on PND. Participants provided with information on PND via two columns from newspaper advice Dr 'Ask Dr Pat' and an information brochure on PND describing illness, causes and treatments	*Primary outcome:* PND diagnosis measured by SCID-I *Secondary outcomes:* BDI-II EPDS *Outcomes measured at:* 1) baseline 2) 3 months 3) 6 months 4) 12 months	SCID-I between group 3 months p=0.73 6 months p=0.4 12 months p=0.053* Attrition=19%

Abbreviations

AD: antenatal depression; AUS: Australia; BAI: Beck Anxiety Inventory; BDI/BDI-II: Beck Depression Inventory; CBT: Cognitive Behavioural Therapy; CCBT: Computerised Cognitive Behavioural Therapy; EPDS: Edinburgh Postnatal Depression Scale; F2F: face-to-face; GAD-7: Generalised Anxiety Disorder scale; HDRS: Hamilton Depression rating scale; ICBT: Internet delivered Cognitive Behavioural Therapy; MADRS-S: Montgomery-Åsberg Depression Rating Scale – Self report version; MDD: Major Depressive Disorder; MMB: MumMoodBooster programme; MOM: Managing our Mood programme; PND: postnatal depression; PSI-SF: parenting stress index-short form; RCT: randomised controlled trial; SCID-I/SCID-IV: The Structured Clinical Interview for DSM-IV/DSM-V; TAU: treatment as usual; TA-ICBT: Therapist Assisted Internet Delivered CBT; USA: United States of America; WLC: Waitlist control.

** *Difference is significant p < 0.01*

* *Difference is significant p < 0.05*

How can Big Data be used to answer public health research questions?

Hafi Saad[1], *Brendan P Bunting, Julie EM McCullough*

1. corresponding author

Date submitted: 5 June 2020. **Date accepted:** 21 January 2021. **Date published:** 1 March 2021.

ABSTRACT

Background: The use of existing clinical/administrative Big Data (BD) when conducting research has become common. In the absence of a single comprehensive source of data, data linkage of multiple sources has proved to be of value in midwifery and public health research.

Aim: The purpose of this paper is to describe the process of accessing BD, linking sources and validating the outcomes to answer specific public health questions.

Method: This article describes and discusses the experience and challenges of using BD whilst undertaking a study to answer the research question: 'What is the prevalence and risk factors of congenital heart defect (CHD) in Northern Ireland (NI)?'

Data from the Northern Ireland Maternity System (NIMATS) (2010–2014), Enhanced Prescribing Database (EPD) (2010–2014), and the HeartSuite database (HSD) (2005–2014) were accessed, linked and analysed using Stata version 13.

Ethical approval: Ethical approval was obtained from the National Health Service (NHS) Research Ethics Committee (REC) (No. 17/SC/0103, 27 February 2017). Permission to access data was obtained from the Honest Broker Service (HBS) acting as a safe haven for the conduct of BD research. Training in the use of the HBS was undertaken to ensure confidentiality, anonymity and data protection. There was a fee to be paid to HBS to cover data preparation cost.

Findings: Successfully linking routinely collected BD sets from multiple sources (HSD, NIMATS, EPD) facilitates the creation of a robust and more complete data set, from which conclusions can be drawn, than may previously have been possible. The experience of working in the HBS was a huge learning curve in acquiring an in-depth understanding of the data extraction process, data verification and data analysis.

Conclusions and implications: Data anonymisation and controlling access to the data are important as they enable ethical access of clinical/administrative BD. The quality of the data was variable in the databases but the linkage process enhanced the value of the outputs. Research of this type has value for midwives, other healthcare practitioners and policy makers in public health settings as it provides factual data that can impact public health. However, as in all types of research, using and linking clinical/administrative BD for public health research has its strengths and limitations which should always be acknowledged.

Keywords: big data, clinical data, administrative data, public health research, Evidence Based Midwifery, Northern Ireland Maternity System, Enhanced Prescribing Database, HeartSuite database, Honest Broker Service

Introduction

There is growing awareness of the significant potential for researchers to maximise the use of clinical and administrative data. Habl et al (2016) provided recommendations about how to improve population well-being by utilising the opportunities provided by Big Data (BD) for public health. Recommendations included urgent need for raising awareness of the importance and possibilities of using BD in public health quality improvement (Habl et al 2016).

BD can enable the establishment of a population surveillance system, which can assist policy makers to identify health policy issues, improvements and impact (Kemm et al 2010). BD is used to discover

risk factors and the association between risk factors and outcome (Riley et al 2013). It can also be used in predicting health outcomes (Maragatham & Devi 2019), rare disease presentation (Moore et al 2014), and adverse events after surgery (Humes et al 2013). BD can be used to identify the sensitivity of diagnostic methods (Borgquist et al 2014) and the underlying cause of death (Hassanzadeh et al 2017).

The study data refer to the clinical and administrative records of all pregnant women in NI, known as NIMATS, the general practitioner (GP) prescribing data, known as EPD, and the HSD which holds records of all diagnosed cases of congenital heart defect (CHD) in NI.

The concept of clinical, administrative databases and health-related BD

Data used in health research can be clinical or administrative or both. Administrative data include diagnostic codes or International Statistical Classification of Diseases and Related Health Problems (ICD) codes, mortality, claims, and demographic data (Ambroggio & Shah 2013). Data such as vital signs, medical or surgical history are regarded as clinical data (Raghupathi & Raghupathi 2014). However, data can also be a mix of clinical and administrative data, as was the case in this study.

A database is defined as *'a collection of organised, related data, especially one in electronic form that can be accessed and manipulated by specialized computer software'* (Random House Kernerman Webster's College Dictionary 2010). The term BD in health has been defined as *'large routinely or automatically collected datasets, which are electronically captured and stored. It is reusable in the sense of multipurpose data and comprises the fusion and connection of existing databases for the purpose of improving health and health system performance. It does not refer to data collected for a specific study'* (Habl et al 2016).

Process of making data accessible in Honest Broker Service

Honest Broker Service (HBS), which was founded by Business Services Organisation (http://www.hscbusiness.hscni.net/index.htm), enables non-identifiable data to be safely shared and accessed in order to expand data utilisation and the related health service benefits which can be attained by using these data.

Early communication

For any study that requires clinical/administrative BD from the data custodian (HBS in this study), early communication is necessary between the researcher and data custodian in order to explain the project, discuss the feasibility of the study protocol, and agree subsequent steps. Ongoing clarification of issues related to the variables was necessary in this study as there was no available standard for variable format, definition, or quality and completeness.

A database transfer plan was needed as the HSD was not available in the regional data warehouse. HBS agreed to load a temporary file from HSD into their data warehouse, pending Data Access Agreements, NHS REC approval, and Honest Broker Governance Board (HBGB) approval. These issues had to be taken into consideration before applying for ethical and governance approval to begin the study.

Ethical and governance consideration

To adequately address ethical and governance issues, Data Access Agreements with Belfast Health and Social Care Trust were signed by HBS and Ulster University (UU) as required by the Department of Health and in accordance with the Health and Social Care (HSC) protocol for sharing service information for secondary purposes.

Analysis was undertaken in a safe setting in HBS, a research access agreement and a disclosure policy agreement were signed, obliging the researcher not to attempt to identify individuals. The agreement also stated the principles to be followed when disclosing and using service user information, encapsulated in the Code of Practice and Good Practice principles, Data Protection Act 2018 (DPA 2018) principles and Caldicott principles. The researcher completed HBS Safe Researcher training. To protect patient confidentiality, HBS applied statistical disclosure control to all outputs before providing them in an anonymised format to the researcher.

Databases used in this study and preparations requested to make this study feasible

The databases which were used in this study are discussed in the following section, detailing the data preparation by HBS and the variables requested for the purpose of the study. Prevalence and risk factors of CHD in NI was the main research question, which required accessing three databases, and the discussion of those three databases and the variables related to them were based on the main research question.

HeartSuite database (2005–2014)

HeartSuite (HSD) is mainly a clinical database which holds records of all diagnosed cases of CHD in NI. It is operated and maintained by the Paediatric Cardiology Department of Royal Belfast Hospital for Sick Children. HSD is a cohesive patient management system specifically designed for cardiothoracic surgery and paediatric cardiology.

In this complicated field, HSD is an adaptable and developing system which provides an online electronic medical record while also recording and generating information in real time on all phases of a patient's care. The majority of the UK specialist paediatric cardiac centres use HSD, which enables

comprehensive clinical audit information to be gathered (HeartSuite 2015).

The variables from HSD that were requested, and details of their completeness, are described in Table 1 (see Supplementary information). The quality of any data is determined by the completeness and accuracy of the data (Olsen et al 2010). Most of the variables requested attained more than 90 per cent completeness. The accuracy of the diagnostic variables in HSD is expected to be high, due to the high percentage of completeness. When a cardiologist makes a diagnosis, that information can be automatically inserted into HSD by selecting the code corresponding to the disease from a drop-down menu, thereby streamlining the data capture process. The ease of this process is important, because the accuracy of certain information in clinical data sets diminishes as the actions required to transfer that data into the data set become more complex (van Walraven & Austin 2012).

Four tables (echo results, diagnoses and comorbidity, cause of death and procedures) were extracted from HSD, transferred to HBS, and merged together to create a single table with one row per case, using the patient's unique health and care number (HCN), case note, or date of birth.

Data preparation by HBS for HSD

Notes for file preparation were given to HBS explaining how to merge the data into a single file for analysis. The notes stated that the echo results table (n=34,092 records), diagnoses and comorbidity table (n=16,422 records), the cause of death table (n=189 records), and the procedures table (n=893 records) should be formatted so that one row is given per case.

The echo results table was used to deal with duplicates (where values of variables differ) and to identify the date of first diagnosis (n=34,092 records; as there were multiple rows per case and for each CHD case, the first entry date was taken for inclusion in the study file and all other entries deleted). For the echo results table, taking the date of the earliest echo for each identity document (ID) reduced the number of rows from 34,092 to 11,425. Of 11,413 rows, 8378 (73 per cent) in the diagnosis and comorbidity table linked successfully to the echo table.

As the date of the first echo is used with date of birth to calculate the age at diagnosis, the fact that 27 per cent of patients in the diagnosis and comorbidity table did not have an echo date was problematic. Rather than have approximately 3000 null ages, the date of first echo was used where available, and earliest date of diagnosis, from the diagnosis and comorbidity table, was used when there was no echo date. Consequently, all patients have a value for age at diagnosis.

For the diagnoses and comorbidity table, the duplicate diagnoses were appended as additional variables to the relevant rows, that is, European Paediatric Cardiac Code (EPC)1 2, EPC2 2, ICD9 2 and ICD10 2 relate to that patient's second diagnosis entry. Each patient could have up to nine diagnosis entries, although most had only one.

For the cause of death table, of 188 rows, 145 rows (77 per cent) were joined to the diagnosis and comorbidity table and the date of death was blank in 12 of these rows.

Once the four tables were combined, non-NI postcode cases were excluded on the assumption that those cases were not NI residents. The final output merged file contains 11,410 rows and had one patient per row. Postcodes were available for 8638 of the 11,410 patients (76 per cent). This allowed urban/rural status and deprivation quintile to be assigned using the latest urban rural classification (2015) (Northern Ireland Statistics & Research Agency 2016). For a small number of patients (<10), a postcode was available but was invalid.

Northern Ireland Maternity System (2010–2014)

Northern Ireland Maternity System (NIMATS) is an electronic clinical database used by HSC Trusts to record details of women seeking maternity care in NI. It includes details of current and any past pregnancies and records whether a woman gave birth to a live-born or stillborn baby. Midwives are the first point of contact for pregnant women and maternity care is led by midwives for all non-complex cases.

The variables from NIMATS that were requested for the study are described in Table 2 (see Supplementary information), together with the completeness of each variable. Only records of resident pregnant women were extracted for use in the current study. Some variables showed a very low level of completeness and thus were not used in the analysis, for example MOTHER_FOLIC_ACID (10 per cent completeness). However, this variable was incomplete as it was only available for part of the time period covered by this study. Fully complete folic acid variables are available from 2014/15 onwards.

The purpose of NIMATS is to follow the woman during the antenatal period, hence it is expected to have a high degree of accuracy especially with regard to variables such as the woman's age, body mass index (BMI), previous family history of chromosomal abnormalities and CHD. Midwives input the required details into the system during interviews and most of the variables are for special concerns regarding case management and advice provided.

The variables requested in this study were not available in a single table, therefore, NIMATS data were extracted by HBS into eight tables. These tables were then joined together by HBS to create one table containing all the variables. The tables were joined using the HCN of the pregnant woman (MOTHER_

HCN) and the booking interview date (DATE_OF_BOOKING_INTERVIEW). Together, these two variables uniquely identified a pregnancy.

The data contained 119,165 rows. The NIMATS tables often contained more than one row for a pregnancy when that pregnancy had different values for a given variable. For example, if the woman had more than one medication, each medication was recorded in a separate row. In order to provide a single table with one row per pregnancy where possible, HBS recoded values into additional variables, for example, MOTHER_PRES_MEDICATION_1 to MOTHER_PRES_MEDICATION_8, containing the (up to) eight different values present for each pregnancy in the Present Medications table.

Data preparation by HBS for NIMATS

HBS agreed to take specific steps before making the file ready for analysis. Each row related to a different birth. For multiple pregnancies only the baby with CHD was included; only the most recent pregnancy was included except where a woman had more than one pregnancy with a baby with CHD, in which case all pregnancies were included, and except in the case of multiple births, where all were included. The STUDY_ID variable (number of all births) was a case number from 1 to 96,233, while the MOTHER_ID variable (number of mothers) was a case number from 1 to 94,545. Multiple births were distinguished by the BIRTH_NUMBER variable. After removing multiple pregnancies not involving more than one case of CHD, 96,321 rows remained. Removing 88 rows due to non-NI postcodes produced the final NIMATS dataset containing 96,233 rows, including 1204 with blank postcodes. A valid urban/rural status was assigned to 94,930 rows (99 per cent of all rows) and a valid deprivation status was assigned to 95,000 rows (99 per cent of all rows).

Enhanced Prescribing Database (2010–2014)

Enhanced Prescribing Database (EPD) includes information on all drugs that have been prescribed by GPs and redeemed by patients in NI. However, patient information is only available for prescriptions that can be scanned electronically, which may reduce the number of records available. EPD was created to facilitate payment to pharmacies. Although designed for budget management, payment databases have been widely used in health services research (Marrie et al 2017), and EPD has been used in previous studies (Wemakor et al 2014).

The list of variables requested from EPD for the current study are listed in Table 3 (see Supplementary information). Some of the variables were used to create the specific variable which included the specific medications needed to perform this study. Based on British National Formulary (BNF) classifications, the medications were grouped into four main groups: mental health medications (MHM), folic acid and vitamins, folic acid antagonist medications and medications for case exclusion.

In the EPD file there were 182 different types of MHM; this number included different doses and different brand names for the same medication. These were collapsed into 49 specific medications by grouping brand names and different doses of the same medications.

General Data Protection Regulation and Data Protection Act

The implementation, in May 2018, of the General Data Protection Regulation (GDPR) (Regulation EU 2016/679 of the European Parliament and of the Council 2016) in the UK was based on the DPA (Department for Digital, Culture, Media and Sport and the Home Office 2018). The GDPR and the DPA have an impact on how researchers deal with BD in research.

The changes introduced by the GDPR make it clear that anonymised data are no longer regarded as personal data (GDPR 2018, Information Compliance Team 2018, The British Psychological Society 2018). The GDPR also highlighted that: '*The principles of data protection should therefore not apply to anonymous information*' (GDPR Recitals 2018:26).

These changes highlighted that the utmost effort is required to keep data anonymous. To this end, regulations included in the DPA (2018) state that: '*it is a criminal offence in cases where anyone uses anonymised data knowingly or recklessly to re-identify information that is de-identified personal data*' without data custodian consent (DPA 2018:99).

The researcher and data custodian should make every effort to keep data anonymous. In this study, this involved training courses for researchers which outlined these concepts and highlighted the importance of maintaining anonymity of data. Data should not be kept longer than is necessary, this was specified in the HBS application form before obtaining access to the data. Data were accessed only in the HBS safe haven. Where less than 10 cases for a particular variable were retrieved, or where there was any risk of identification of an individual, HBS withheld these cases in accordance with their disclosure control procedure.

GDPR clarifies the legal basis for processing data, which should be because doing so is 'in the public interest'. This means that studies must clearly set out the public interest which will be served through accessing required data. The purpose of future studies is to attempt to guide public health action for prevention and service planning, therefore accessing data is legal for studies in the public interest (Information Compliance Team 2018).

Data validation and verification

Some of the key issues that arose in relation to using available clinical/administrative BD were related to validity of the data, and this has two aspects: accuracy and completeness. The accuracy of data entry, their selection and quality, and collection methods were not under the control of the researcher and were sometimes impossible to validate. In databases used in this study, standard definitions for each variable were not available, and data entry practices used by health providers were not standardised for some variables. Use of non-standardised texts made some variables difficult to understand and to analyse, as the text used depended on the personal preference of the data inputter.

For HSD, values were acceptable for most of the variables, however there were a small number of variables which could not be analysed. For example, values for the age at diagnosis (AGE_AT_DIAGNOSIS_YEARS) turned out to be negative in 13 cases and meant that those cases were not included in analysis for that specific variable. Another example was the fetal diagnosis (FETAL_DIAGNOSIS) variable which took values of 1, N, U, Y, or blank which did not give a clear idea whether a child had a fetal diagnosis or not (see Table 1).

In this study, and in HSD, the CHD diagnosis was the main outcome and measured by using ICD and EPC codes. Both codes were presumed to be reliable and accurate as they are used to facilitate individual clinical management. It is important to note, however, that HSD used EPCC definitions in addition to ICD codes. ICD codes, produced by the World Health Organization (WHO 2016), have been used in health research for studying disease pattern, patterns of care and outcomes of disease. ICD9 and ICD10 codes, although used all over the world, have limitations and are increasingly viewed by paediatric cardiologists as insufficient for depicting CHD (Houyel et al 2011). Hence, additional clinical information related to CHD which may be useful for more clinically oriented studies can be obtained by using EPCC to cover that gap.

In an effort to verify some of the data in this study, a comparison of mortality data from the HSD file with results from external reports was performed. Mortality data from HSD using the variable age at death (AGE_AT_DEATH_IN_DAYS) showed that, for the period 2005–2014, there were 63 deaths among children who were live born with CHD. This is less than the figure for infant mortality due to CHD for the same period (78 deaths) sourced from the Northern Ireland Statistics and Research Agency (NISRA) website, which provides statistics on vital events registered by the General Registrar Office for NI. The information from NISRA is considered to be of high quality and regarded as a source of research data in NI (https://www.nisra.gov.uk/), therefore suggesting that this variable is not complete in HSD.

Other examples of lower quality variables in HSD include the cause of death variable (CAUSE_OF_DEATH) which was written as free text and very sparsely populated, presence of comorbidity (COMORBIDITY_PRESENT) and rating of the catheterisation (CATHETERISATION_RATING) which also showed blank values for all the data. There is, therefore, room for data quality improvement.

For NIMATS, the values were acceptable for most of the variables, however, again there were a small number of variables which had issues related to the range of acceptable values. For example, values for the infant gender (INFANT_GENDER) variable had a small number of cases with values of 1 and 2 but did not specify which was male/female. The variables showing whether or not a woman was taking folic acid (MOTHER_FOLIC_ACID_1 to MOTHER_FOLIC_ACID_4) were low quality, with 10 per cent or less value completeness (see Table 2). Again, there is room for data quality improvement.

In another effort to verify some of the data in this study, checks were conducted on the data comparing some descriptive results from the NIMATS file with results from the Children's Health in Northern Ireland report (Public Health Agency 2016). Six of the eight variables from the report showed almost the same results from the current data; two variables showing a recognisable difference were number of live births and stillbirths for years 2010 and 2011. This can most likely be attributed to the request in the project application for only the most recent delivery for women with more than one pregnancy in the time period 2010–2014.

It is important to note that, as demonstrated in Tables 1 and 2, the completeness (coverage) of most of the variables in both databases was reasonably good. A missing rate of five per cent or less is considered to be insignificant (Schafer 1999), however, Bennett (2001) stated that when more than 10 per cent of data are missing the statistical analysis is likely to be biased. For this reason, variables with missing values of more than 10 per cent were not used in the study, and sensitivity analysis was based on that figure.

Benefits and limitations of using clinical/administrative BD and linking BD for public health research

Using clinical/administrative BD conferred several benefits, one of the most important being a very low drop-out rate as this type of database covers the whole population (Harron et al 2014). When recruiting participants for prospective studies, bias may result as people with higher educational attainment, higher socio-economic status and better health are more likely to be recruited than the wider population (Nishiwaki et al 2005). When recruiting for prospective studies, it is also acknowledged that

less healthy individuals are more likely to drop out during follow-up (Murphy et al 2011). This can lead to biased estimates of association between risk factors and outcome, a situation which can be avoided by using routinely collected data samples.

The clinical/administrative BD in this study was used in a longitudinal cohort study which is considered to be of good quality in the hierarchy of evidence (Hagger-Johnson 2014). Using the availability of existing databases is less time-consuming and more cost-effective in comparison to research based on primary data.

As with other BD studies, not all required variables were available, but understanding of variables in the database was crucial for the researchers to be able to create derived variables. For example, age at diagnosis was important to measure the prevalence of CHD precisely, as prevalence changes based on age at diagnosis. Age at diagnosis was not available in HSD as a distinctive variable hence the date of the first echo was used, along with date of birth, in order to calculate the age at diagnosis.

The use of clinical/administrative BD involves large samples which is very helpful, especially when studying rare diseases such as CHD. Using big health data may produce results with significant association. However, statistically significant results should be discussed cautiously to highlight whether their significance is clinical or statistical. To distinguish between statistical and clinical significance, absolute and relative differences and confidence intervals should be used when comparing two different populations (van Walraven & Austin 2012).

Linking BD for public health research

Linking clinical/administrative BD has been acknowledged to be beneficial for public health research and data linkage is a cost-efficient way of obtaining information rapidly to guide policy (Hagger-Johnson 2016). In this study, which involved data covering up to 10 years, it took less than six months from the HBGB approval of the study to data being accessed. The main research question of the study regarding prevalence and risk factors of CHD in NI was answered rapidly using linked data for the purpose of guiding public health action for prevention and service planning, without the need to spend years gathering new data.

Linking data to obtain a final data file may improve the quality, integrity, and strength of data (Kemm et al 2010). For example, a variable might be available in a data set with poor quality and linking this data with another database, which includes the same variable with a better quality, improves the final data file. This was the case in this study, where there was eight per cent missing data on the postcode variable in HSD (the postcode was crucial to assign deprivation and urban/rural status). However, when using the linked file and the variable from NIMATS (covering the period 2010–2014), missing data from the postcode variable decreased to one per cent and data quality improved dramatically. Additionally, as there was a unique identifier for each case, data linkage allowed identification of duplicates.

Data linkage studies may promote data privacy and confidentiality. In this study, the privacy and confidentiality of individuals was protected by ensuring that personal identifiers, such as names, were not available in the linked dataset analysed by the researchers. Data access took place in the safe haven, the researcher and HBS worked together to apply statistical disclosure control for all results and numbers less than 10 were not presented in the results.

Conducting a data linkage study encourages cooperation between different professionals such as researchers, clinicians and data providers. Communication within the multidisciplinary team improved understanding and identification of current issues related both to the study and to future opportunities (Holman et al 2008). In this study professionals agreed that the research questions could not be clearly answered without data linkage, and that data linkage would make the study more feasible.

Funders from within and outside NI may be more likely to finance research which uses linked data from NI and therefore has commercial and competitive benefits for NI. NI Chest, Heart and Stroke (https://nichs.org.uk) provided partial funding to cover the costs incurred by HBS in creating a linked, anonymised dataset, use of the safe setting in which to analyse the data, and the time taken for disclosure control checks prior to outputs leaving the safe setting.

Difficulties and limitations

Linking data has some difficulties and limitations. Data linkage studies can be time-consuming (Flowers & Ferguson 2010) and a laborious process (Lawrence & Bradley 2018). This was the case in this study, not in terms of performing the technical linkage, but to fully identify and understand all the issues related to data sets and the high number of variables (especially where no metadata is available for each variable required).

In this study different files and variables were linked together for the first time. The possibility of failure to obtain correct and good quality variables might be frustrating to researchers. Dealing with different organisations that may have different procedures and regulations to gain approval for the research requires personal skills such as perseverance, leadership, and inter-sectoral cooperation.

Although maintenance of privacy and confidentiality of the data has been recognised as a strength of data linkage studies, in this study and the literature (Holman et al 2008), some risk of disclosure of individual information and other confidentiality issues remains (Flowers & Ferguson 2010). In this

study, disclosure control was applied by HBS and the researcher as explained earlier. It is important to emphasise that when undertaking research with human subjects, tension can arise between the need for data protection on one hand and data sharing on the other. Corti et al (2014) emphasised that even sensitive and confidential research data can be shared ethically and legally, as long as certain criteria are met. These criteria include anonymising the data and controlling access to the data if necessary.

Conclusions

In conclusion, routinely collected clinical/administrative BD from HSD, NIMATS, and EPD that can be linked together are relatively reliable and provide important variables that can be used in public health research. The variables from different databases used in this study were generally of good quality, covering the whole population of NI. Understanding the variables in databases is mandatory to create necessary derived variables. The usage of BD allows the study of rare diseases and permits sub-groups analysis. It is important for midwives and other health care professionals to understand the value of their role in BD research studies, including the need for clear and accurate record taking and data inputting.

Reporting the percentage and completeness of variables used in databases has given an idea about data quality for potential future research covering the period of 2005–2014 and 2010–2014. Clinical/administrative BD should be checked regularly for quality, and quality improvement should be targeted continuously to ensure a high standard of data that are fit for the purpose of research. Standard clarification of what each variable in NIMATS means is required to facilitate standard completion of patients' information in the system and to enhance the data available for research. Metadata should also be established for all clinical data in NI. It is important to note that, while there was a lack of available metadata for the NIMATS database during the period of this research, NIMATS metadata is now available.

The research questions posed by this study, which were the prevalence of CHD and the association between specified risk factors and CHD in NI, were answered. The study demonstrated that data linkage studies, using different files from various sources through the HBS, are feasible to answer public health research questions. Data anonymisation and access control were important as these enabled ethical access of clinical/administrative BD.

To promote transparency in discussions and results, strengths and limitations of using and linking clinical/administrative BD for public health research should always be acknowledged. Researchers should be prepared to overcome challenges in dealing with these databases in order to set up agreed standards for variable definitions and qualities and to benefit from and share their experiences.

Funding

The study was funded by Vice-Chancellor's Research Scholarship Ulster University. Northern Ireland Chest Heart and Stroke covered the costs of data preparation carried out by HBS.

Acknowledgments

The author would like to acknowledge the support of Professor Marlene Sinclair in preparing this manuscript for publication and to acknowledge the original PhD research supervisory team at Ulster University: Professor Helen Dolk, Dr Karen Casson, Dr Maria Loane, Dr Paul Slater and Dr Nichola McCullough.

The author would also like to acknowledge Professor Frank Casey, Mrs Rita Butler from Royal Belfast Hospital for Sick Children, and Mr Neil Marsden, Mr Gary Ewing, and Mr Scott Mathieson from HBS within the Business Services Organisation Northern Ireland.

Authors

Hafi Saad, MD Community Medicine, MSc Molecular Medicine, MPH, and PhD in Public Health and Epidemiology, former PhD researcher in Maternal Fetal and Infant Research Centre, Ulster University, Jordanstown, UK (corresponding author)
Email: Saad-h@ulster.ac.uk.

Brendan P Bunting, Professor of Psychology, School of Psychology, Ulster University, Coleraine, UK.

Julie EM McCullough, BSc (Hons), PGDip, PhD Research Associate, Institute of Nursing and Health Research, Ulster

References

Ambroggio LV, Shah SS (2013). Administrative data: expanding the infrastructure for pediatric research. *The Journal of Pediatrics* 162(4):681-4.

Bennett DA (2001). How can I deal with missing data in my study? *Australian and New Zealand Journal of Public Health* 25(5):464-9.

Borgquist R, Haugaa KH, Gilljam T, Bundgaard H, Hansen J, Eschen O, Jensen HK, Holst AG, Edvardsen T, Svendsen JH. Platonov PG (2014). The diagnostic performance of imaging methods in ARVC using the 2010 Task Force criteria. *European Heart Journal – Cardiovascular Imaging* 15(11):1219-25.

British Psychological Society (2018). *Data protection regulation: guidance for researchers*. https://www.bps.org.uk/sites/bps.org.uk/files/Policy/Policy%20-%20Files/Data%20Protection%20Regulation%20-%20Guidance%20for%20Researchers.pdf [Accessed 3 April 2019].

Corti L, Van den Eynden V, Bishop L, Woollard M (2014). *Managing and sharing research data: a guide to good practice*. London: Sage Publications Ltd.

Data Protection Act 2018 (c12). Part 3, Law enforcement processing, Chapter 2, Principles. http://www.legislation.gov.uk/ukpga/2018/12/pdfs/ukpga_20180012_en.pdf [Accessed 4 March 2019].

Department for Digital, Culture, Media and Sport and the Home Office (2018). *Data Protection Act 2018 explanatory notes*. http://www.legislation.gov.uk/ukpga/2018/12/pdfs/ukpgaen_20180012_en.pdf [Accessed 4 March 2019].

Flowers J., Ferguson B (2010). The future of health intelligence: challenges and opportunities. *Public Health* 124(5):274-7.

General Data Protection Regulation Recitals (2018). *Recital 26*. https://www.privacy-regulation.eu/en/r26.htm [Accessed 22 December 2018].

Habl C, Renner A-T, Bobek J, Laschkolnig A (2016). *Study on Big Data in public health, telemedicine and healthcare*. Brussels: European Commission. https://ec.europa.eu/health/sites/health/files/ehealth/docs/ev_20161121_co13_en.pdf [Accessed 17 March 2019].

Hagger-Johnson G (2014). *Introduction to research methods and data analysis in the health sciences*. Abingdon: Routledge.

Hagger-Johnson G (2016). Opportunities for longitudinal data linkage in Scotland. *Scottish Medical Journal* 61(3):136-45.

Harron K, Wade A, Gilbert R, Muller-Pebody B and Goldstein H (2014). Evaluating bias due to data linkage error in electronic healthcare records. *BMC Medical Research Methodology* 14(36).

Hassanzadeh HR, Sha Y, Wang MD (2017). Deep death: learning to predict the underlying cause of death with Big Data. 39th *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. Seogwipo, South Korea, 11-15 July 3373-6.

HeartSuite (2015). *HeartSuite*. http://www.heartsuite.com/5.html [Accessed 5 March 2018].

Holman CD'Arcy, Bass AJ, Rosman DL, Smith MB, Semmens JB, Glasson EJ, Brook EL, Trutwein B, Rouse IL, Watson CR, de Klerk NH, Stanley FJ (2008). A decade of data linkage in Western Australia: strategic design, applications and benefits of the WA data linkage system. *Australian Health Review: a publication of the Australian Hospital Association* 32(4):766-77.

Houyel L, Khoshnood B, Anderson RH, Lelong N, Thieulin A-C, Goffinet F, Bonnet D (2011). Population-based evaluation of a suggested anatomic and clinical classification of congenital heart defects based on the International Paediatric and Congenital Cardiac Code. *Orphanet Journal of Rare Diseases* 6(64).

Humes DJ, Radcliffe RS, Camm C, West J (2013). Population-based study of presentation and adverse outcomes after femoral hernia surgery. *British Journal of Surgery* 100(13):1827-32.

Information Compliance Team (2018). *Data protection and research, version 1, 26 March 2018*. Oxford: University of Oxford. https://researchsupport.admin.ox.ac.uk/sites/default/files/researchsupport/documents/media/data_protection_and_research.pdf [Accessed 3 April 2019].

Kemm JR, Robinson J, Verne J (2010). Social care data in England: what they tell us and what they do not tell us. *Public Health* 124(5):265-8.

Lawrence NR, Bradley SH (2018). Big Data and the NHS – we have the technology, but we need patient and professional engagement. *Future Healthcare Journal* 5(3):229-30.

Maragatham G, Devi S (2019). LSTM model for prediction of heart failure in Big Data. *Journal of Medical Systems* 43(5):111.

Marrie RA, Walld R, Bolton JM, Sareen J, Walker JR, Patten SB, Singer A, Lix LM, Hitchon CA, El-Gabalawy R, Katz A, Fisk JD, Bernstein CN, for the CIHR Team in Defining the Burden and Managing the Effects of Psychiatric Comorbidity in Chronic Immunoinflammatory Disease (2017). Estimating annual prevalence of depression and anxiety disorder in multiple sclerosis using administrative data. *BMC Research Notes* 10(619).

Moore A, Harnden A, Mayon-White R (2014). Recognising Kawasaki disease in UK primary care: a descriptive study using the Clinical Practice Research Datalink. *British Journal of General Practice* 64(625):e477-83.

Murphy TE, Han L, Allore HG, Peduzzi PN, Gill TM, Lin H (2011). Treatment of death in the analysis of longitudinal studies of gerontological outcomes. *The Journals of Gerontology: Series A, biological sciences and medical sciences*. 66(1):109-14.

Nishiwaki Y, Clark H, Morton SM, Leon DA (2005). Early life factors, childhood cognition and postal questionnaire response rate in middle age: the Aberdeen *Children of the 1950s study. BMC Medical Research Methodology* 5(16).

Northern Ireland Statistics & Research Agency (NISRA) (2016). *Technical guidance on production of official statistics for settlements and urban-rural classification*. https://www.nisra.gov.uk/sites/nisra.gov.uk/files/publications/Settlement15-guidance.pdf [Accessed 12 June 2016].

Olsen J, Saracci R, Trichopoulos D eds (2010). *Teaching epidemiology: a guide for teachers in epidemiology, public health and clinical medicine*. Oxford: Oxford University Press.

Public Health Agency (2016). *Children's health in Northern Ireland*. https://www.publichealth.hscni.net/sites/default/files/RUAG%20report%202015-16%20-%20Childrens%20Health%20in%20NI%20-%20FINAL%20REPORT%20-%20May%202016.pdf [Accessed 6 March 2018].

Raghupathi W, Raghupathi V (2014). Big Data analytics in healthcare: promise and potential. *Health Information Science and Systems* 2:3.

Random House Kernerman Webster's College Dictionary (2010). da'ta•base or da'ta base. https://www.thefreedictionary.com/database [Accessed 12 May 2018].

Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016. Regulations. https://gdprinfo.eu/ [Accessed 13 December 2018].

Riley RD, Hayden JA, Steyerberg EW, Moons KGM, Abrams K, Kyzas PA, Malats N, Briggs A, Schroter S, Altman DG, Hemingway H (2013). Prognosis research strategy (PROGRESS) 2: Prognostic Factor Research. *PLoS Medicine* 10(2):e1001380. https://doi.org/10.1371/journal.pmed.1001380 [Accessed 14 March 2016].

Schafer JL (1999). Multiple imputation: a primer. *Statistical Methods in Medical Research* 8(1):3-15.

van Walraven C, Austin P (2012). Administrative database research has unique characteristics that can risk biased results. *Journal of Clinical Epidemiology* 65(2):126-31.

Wemakor A, Casson K, Dolk H (2014). Prevalence and sociodemographic patterns of antidepressant use among women of reproductive age: a prescription database study. *Journal of Affective Disorders*16:299-305.

World Health Organization (2016). *International classification of diseases*. http://www.who.int/classifications/icd/en/ [Accessed 15 June 2018].

How to cite this paper:

Saad H, Bunting BP, McCullough JEM (2021). How can Big Data be used to answer public health research questions?. *Evidence Based Midwifery* 19(1):19-31.

Supplementary information

Table 1. Description and completeness of variables from HSD used in the study.

Variable	Description	Completeness of data/comments
1. PREVALENCE_STUDY_ID	A case number numbered from 1 to 11,410.	100%
2. YEAR_OF_BIRTH	Based on date of birth.	100%
3. AGE_AT_DEATH_YEARS	Calculated from date of birth and date of death, where available.	1%
4. AGE_AT_DIAGNOSIS_YEARS	Calculated from date of birth and date of earliest echo. Where date of earliest echo was not available, earliest date of diagnosis (from the diagnosis table) was used instead.	100%
5. AGE_AT_PROCEDURE_YEARS [1-5]	Calculated from date of birth and procedure date. There are up to five separate variables for patients who have had multiple procedures.	AGE_AT_PROCEDURE_YEARS1 5% AGE_AT_PROCEDURE_YEARS2 1% AGE_AT_PROCEDURE_YEARS3 <1% AGE_AT_PROCEDURE_YEARS4 <1% AGE_AT_PROCEDURE_YEARS5 <1%
6. YEAR_OF_DEATH	Based on year of death (in HSD).	1%
7. AGE_AT_DEATH_IN_DAYS	Calculated from date of birth and date of death, where available. Values include '0-6 days', '7-27 days', '28-364 days' and 'not within first year'.	1%
8. CAUSE_OF_DEATH	Free text. Very sparsely populated.	<1%
9. SEX	Takes values 'M', 'F', 'U' or blank.	99.90%
10. FETAL_DIAGNOSIS	Takes values '1', 'N', 'U', 'Y' or blank.	20%
11. EPC1 [1-9]	European Paediatric Cardiac (EPC) code. Each diagnosis record had two EPC variables: EPC1 and EPC2. Duplicate diagnoses were recoded as additional variables and indicated by the second digit in the variable name. For example, EPC1 3 indicates that a patient has three separate entries in the diagnosis table. Up to nine rows are available for each patient in the HeartSuite diagnosis table.	EPC1 1 100% EPC1 2 29% EPC1 3 8% EPC1 4 3% EPC1 5 1% EPC1 6 <1% EPC1 7 <1% EPC1 8 <1% EPC1 9 <1%
12. EPC2 [1-9]	These are the second EPC codes for each diagnosis record. For example, EPC2 3 gives the second EPC code value when a single patient has three separate entries in the diagnosis table.	EPC2 113% EPC2 2 6% EPC2 3 2% EPC2 4 1% EPC2 5 <1% EPC2 6 <1% EPC2 7 <1% EPC2 8 <1% EPC2 9 <1%
13. ICD9 [1-9]	International Classification of Diseases and Related Health Problems (ICD), edition 9. See EPC1 for explanation of '[1-9]'.	ICD9 1 74% ICD9 2 21% ICD9 3 6% ICD9 4 3% ICD9 5 1% ICD9 6 <1% ICD9 7 <1% ICD9 8 <1% ICD9 9 <1%

Variable	Description	Completeness of data/comments
14. ICD10 [1-9]	International Classification of Diseases and Related Health Problems (ICD), edition 10. See EPC1 for explanation of '[1-9]'.	ICD10 1 80% ICD10 2 27% ICD10 3 8% ICD10 4 3% ICD10 5 1% ICD10 6 <1% ICD10 7 <1% ICD10 8 <1% ICD10 9 <1%
15. DIAGNOSIS [1-5]	Diagnosis variable from the HeartSuite procedures table. Originally coded using 25 separate variables, linked together here to prevent an excessive number of variables. Free text, including codes. Up to five separate variables are assigned to patients who have had multiple procedures.	DIAGNOSIS1 5% DIAGNOSIS2 1% DIAGNOSIS3 <1% DIAGNOSIS4 <1% DIAGNOSIS5 <1%
16. COMORBIDITY_PRESENT	Blank values for all cases.	0%
17. COMORBID_CONDITIONS [1-5]	Comorbid conditions variable from the HeartSuite procedures table. Free text, including codes. Up to five separate variables are assigned to patients who have had multiple procedures.	COMORBID_CONDITIONS 1 1% COMORBID_CONDITIONS 2 <1% COMORBID_CONDITIONS 3 <1% COMORBID_CONDITIONS 4 <1% COMORBID_CONDITIONS 5 <1%
18. PROC_DATE [1-5]	Procedure date variable from the HeartSuite procedures table. Up to five separate variables are assigned to patients who have had multiple procedures. The first PROC_DATE variable contains the latest procedure.	PROC_DATE1 5% PROC_DATE2 1% PROC_DATE3 <1% PROC_DATE4 <1% PROC_DATE5 <1%
19. PROC_TYPE [1-5]	Procedure type variable from the HeartSuite procedures table. Up to five separate variables are assigned to patients who have had multiple procedures. The first PROC_TYPE variable contains the latest procedure.	PROC_TYPE1 5% PROC_TYPE2 1% PROC_TYPE3 <1% PROC_TYPE4 <1% PROC_TYPE5 <1%
20. CATHETERISATION_RATING	Blank values for all cases.	0%
21. URBAN_RURAL_2015	Urban rural classification based on NISRA's Review of the Statistical Classification and Delineation of Settlements (March 2015). Assigned from patient postcode using the November 2016 Central Postcode Directory.	76%
22. NIMDM_2010_QUINTILE	Deprivation quintile, from 1 (most deprived) to 5 (least deprived). Calculated from the NI Multiple Deprivation Measure 2010 Super Output Area rank. Super Output Areas were assigned from patient postcodes using the November 2016 Central Postcode Directory.	76%

Table 2. Completeness and notes of variables in NIMATS (2010–2014).

Variable	Notes	Completeness of data/comments
1. STUDY_ID	A case number numbered from 1 to 96,233. Allows data linkage to EPD data.	100%
2. MOTHER_ID	A case number numbered from 1 to 94,545. Can be used to identify small number of multiple pregnancies where both births involved CHD.	100%
3. PARITY	Number of times a woman has given birth.	>99%
4. GEST_AT_BOOKING_WEEKS	Gestational age at booking, measured in weeks.	>99%
5. AGE_AT_BOOKING	Mother's age at booking, measured in years and grouped as follows: '<20', '20-24', '25-29', '30-34', '35-39', '40-44', '45+'.	>99%
6. PLANNED_PREGNANCY	Yes/No. 2,524 blanks.	97%
7. FAMILY_CONGEN_ABNORM_01_ CARDIAC	Yes/No for family history of congenital anomalies of type '01 - Cardiac'.	100%
8. FAMILY_CONGEN_ABNORM_02_ CHROMOSOMAL	Yes/No for family history of congenital anomalies of type '02 - Chromosomal'.	100%
9. MOTHERS_BMI	Mother's body mass index: 3,596 missing cases due to mother's height and/or weight missing. Height and weight were requested in the hope that additional BMI values could be calculated, but on inspection this variable contained values whenever both height and weight were recorded.	96%
10. SMOKING	Cigarettes per day.	>99%
11. PARTNER_SMOKING	Cigarettes per day.	96%
12. ALCOHOL	Units per week.	>99%
13. MOTHER_FOLIC_ACID_1	Present medications – FA.	10%
14. MOTHER_FOLIC_ACID_2	Patients were observed to have up to four values recorded against a single pregnancy.	<1%
15. MOTHER_FOLIC_ACID_3	Patients were observed to have up to four values recorded against a single pregnancy.	<1%
16. MOTHER_FOLIC_ACID_4	Patients were observed to have up to four values recorded against a single pregnancy.	<1%
17. MOTHER_PRES_MEDICATION_1	Present medications – all medications used including FA.	97%
18. MOTHER_PRES_MEDICATION_2	Patients were observed to have up to eight values recorded against a single pregnancy.	24%
19. MOTHER_PRES_MEDICATION_3	Patients were observed to have up to eight values recorded against a single pregnancy.	5%
20. MOTHER_PRES_MEDICATION_4	Patients were observed to have up to eight values recorded against a single pregnancy.	1%
21. MOTHER_PRES_MEDICATION_5	Patients were observed to have up to eight values recorded against a single pregnancy.	<1%
22. MOTHER_PRES_MEDICATION_6	Patients were observed to have up to eight values recorded against a single pregnancy.	<1%
23. MOTHER_PRES_MEDICATION_7	Patients were observed to have up to eight values recorded against a single pregnancy.	<1%
24. MOTHER_PRES_MEDICATION_8	Patients were observed to have up to eight values recorded against a single pregnancy.	<1%
25. MENTAL_HEALTH_CODE_AND_ DESC_1	Mental health disorders from previous medical history. Patients were observed to have up to five values recorded against a single pregnancy.	95%
26. MENTAL_HEALTH_YEAR_1	Patients were observed to have up to five values recorded against a single pregnancy. The year variables are not clean data.	16%
27. MENTAL_HEALTH_CODE_AND_ DESC_2	Mental health disorders from previous medical history. Patients were observed to have up to five values recorded against a single pregnancy.	1%
28. MENTAL_HEALTH_YEAR_2	Patients were observed to have up to five values recorded against a single pregnancy. The year variables are not clean data.	1%
29. MENTAL_HEALTH_CODE_AND_ DESC_3	Mental health disorders from previous medical history. Patients were observed to have up to five values recorded against a single pregnancy.	<1%

Variable	Notes	Completeness of data/comments
30. MENTAL_HEALTH_YEAR_3	Patients were observed to have up to five values recorded against a single pregnancy. The year variables are not clean data.	<1%
31. MENTAL_HEALTH_CODE_AND_DESC_4	Mental health disorders from previous medical history. Patients were observed to have up to five values recorded against a single pregnancy.	<1%
32. MENTAL_HEALTH_YEAR_4	Patients were observed to have up to five values recorded against a single pregnancy. The year variables are not clean data.	<1%
33. MENTAL_HEALTH_CODE_AND_DESC_5	Mental health disorders from previous medical history. Patients were observed to have up to five values recorded against a single pregnancy.	<1%
34. MENTAL_HEALTH_YEAR_5	Patients were observed to have up to five values recorded against a single pregnancy. The year variables are not clean data.	<1%
35. CARDIO_CODE_AND_DESC_1	Cardiovascular condition from previous medical history, including Code 02: Congenital Heart Disease.	99%
36. CARDIO_CODE_AND_DESC_2	Patients were observed to have up to three values recorded against a single pregnancy.	<1%
37. CARDIO_CODE_AND_DESC_3	Patients were observed to have up to three values recorded against a single pregnancy.	<1%
38. STATUS_AT_BIRTH	Live or still. 16 blanks.	>99%
39. INFANT_BIRTH_WEIGHT	Measured in grams. 116 blanks.	>99%
40. INFANT_GENDER	F/M. 11 cases of I, two cases of U, nine cases missing.	>99%
41. PRESENT_PREG_PROB_CODE_DESCRIPTION_1	Antenatal summary information – present pregnancy problems code and description. Patients were observed to have up to nine values recorded against a single pregnancy.	99%
42. PRESENT_PREG_PROB_CODE_DESCRIPTION_2	Antenatal summary information – present pregnancy problems code and description. Patients were observed to have up to nine values recorded against a single pregnancy.	12%
43. PRESENT_PREG_PROB_CODE_DESCRIPTION_3	Antenatal summary information – present pregnancy problems code and description. Patients were observed to have up to nine values recorded against a single pregnancy.	3%
44. PRESENT_PREG_PROB_CODE_DESCRIPTION_4	Antenatal summary information – present pregnancy problems code and description. Patients were observed to have up to nine values recorded against a single pregnancy.	1%
45. PRESENT_PREG_PROB_CODE_DESCRIPTION_5	Antenatal summary information – present pregnancy problems code and description. Patients were observed to have up to nine values recorded against a single pregnancy.	<1%
46. PRESENT_PREG_PROB_CODE_DESCRIPTION_6	Antenatal summary information – present pregnancy problems code and description. Patients were observed to have up to nine values recorded against a single pregnancy.	<1%
47. PRESENT_PREG_PROB_CODE_DESCRIPTION_7	Antenatal summary information – present pregnancy problems code and description. Patients were observed to have up to nine values recorded against a single pregnancy.	<1%
48. PRESENT_PREG_PROB_CODE_DESCRIPTION_8	Antenatal summary information – present pregnancy problems code and description. Patients were observed to have up to nine values recorded against a single pregnancy.	<1%

Variable	Notes	Completeness of data/comments
49. PRESENT_PREG_PROB_CODE_DESCRIPTION_9	Antenatal summary information – present pregnancy problems code and description. Patients were observed to have up to nine values recorded against a single pregnancy.	<1%
50. CONGEN_ABNORMALITY	Flag variable for infant congenital abnormality (at birth).	1%
51. GEST_AT_DELIVERY_WEEKS	Gestational age at delivery, measured in weeks.	100%
52. ADMITTED_TO_NEO_NATAL_UNIT	Y/N.	100%
53. OTHER_OP_PROCEDURES_CODE_DESCRIPTION_1	Other operative procedures code and description. Patients were observed to have up to three values recorded against a single pregnancy.	97%
54. OTHER_OP_PROCEDURES_CODE_DESCRIPTION_2	Other operative procedures code and description. Patients were observed to have up to three values recorded against a single pregnancy.	<1%
55. OTHER_OP_PROCEDURES_CODE_DESCRIPTION_3	Other operative procedures code and description. Patients were observed to have up to three values recorded against a single pregnancy.	<1%
56. BIRTHS_THIS_PREGNANCY	Number of births this pregnancy: 1, 2 or 3.	100%
57. BIRTH_NUMBER	Birth number: generally, 1, 2 or 3. Two blanks. One unusual value of '6'.	>99%
58. INFANT_YEAR_OF_BIRTH	Based on infant date of birth.	100%
59. MDM_QUINTILE	Deprivation quintile. 1 represents the most deprived quintile and 5 the least deprived. Calculated from the NI Multiple Deprivation Measure 2010 Super Output Area rank. Super Output Areas were assigned from patient postcodes using the November 2016 Central Postcode Directory.	99%
60. URBAN_RURAL	Urban rural classification based on NISRA's Review of the Statistical Classification and Delineation of Settlements (March 2015). Assigned from patient postcode using the November 2016 Central Postcode Directory.	99%
61. LAST_MENSTRUAL_PERIOD_DATE	Used to identify exposure within the exposure window.	100%

Table 3. Variables and notes from EPD used in the study (2010–2014).

Variable	Notes
1. STUDY_ID (by using mother's HCN number)	A case number from 1 to 96,233. Allows linkage to NIMATS data.
2. NAME_FORM_STRENGTH	Generic/brand name. Contains the name, form and strength of the medicinal product.
3. FORM_DMD	Medicinal product form/presentation, e.g. tablet, capsule.
4. STRENGTH_DMD	Medicinal product strength. The amount of the active ingredient that is present in each dosage.
5. DISPENSED_QUANTITY_VALUE	The quantity of tablets, capsules, liquid etc. dispensed.
6. DDD	Defined daily dose.
7. DDD_PER_UNIT	Derived daily dose per quantity, calculated by the Business Services Organisation. For use in calculation: (DDD_PER_UNIT * QUANTITY)/number of days.
8. BNF_CHAPTER	British National Formulary Chapter
9. BNF_SECTION	British National Formulary Section
10. BNF_PARAGRAPH	British National Formulary Paragraph
11. PRESCRIPTION_CATEGORY_IND	Indicator of type of prescription. M=multiple dispensed, N=normal, R=repeat dispensed, S=substitute.
12. ISSUE_DATE	The date the prescription was dispensed by the pharmacy. For HBS use only to calculate Exposure Window (EW).
13. SPECIFIED_MEDICATION	Flag variable 0/1, indicating whether the prescribed medication was one of the medications specified in Section C2 of the project application form. Other medications were extracted to allow for consideration of co-medications.
14. EXPOSURE_TO_COMEDS_IN_EW	Y/N. Exposure to co-medications during the exposure window.
15. MULTIPLE_BNF_CODES_Y_N	Y/N. Flag variable that identifies where a medication has more than one BNF code.
16. MULTIPLE_BNF_CODES	If MULTIPLE_BNF_CODES=Y, the BNF codes are listed here, separated by a semicolon.

ROYAL
COLLEGE
OF MIDWIVES