DOI:
10.1016/j.midw.2020.102864
https://doi.org/10.1016/j.midw.2020.102864

Document Version
Early version, also known as pre-print

• ***Title:*** Impact of a medical mobile phone app (QUiPP) for predicting preterm birth on the anxiety and decisional conflicts faced by women in threatened preterm labour.

• ***Author names and affiliations:*** Carlisle N, Watson H A, Seed P T, Carter J, Kuhrt K, Tribe R M, Shennan A H

• ***Corresponding author:*** Naomi Carlisle naomi.h.carlisle@kcl.ac.uk

• ***Present/permanent address:*** Department of Women and Children's Health, School of Life Course Sciences, King's College London, 10th Floor North Wing, St Thomas' Hospital, Westminster Bridge Road, SE1 7EH

Introduction

In the United Kingdom (UK), 8% of babies are born preterm (before 37 weeks' gestation)(Office for National Statistics, 2019). Babies who survive preterm birth are at high risk of short and long-term health problems (Liu *et al.*, 2016; Patel, 2016), and The Department of Health is therefore aiming to reduce the preterm birth rate to 6% by 2025 (Department of Health, 2017; NHS England, 2019).

However many women who present before 37 weeks' gestation with symptoms of labour do not do on to deliver their baby early (Iams *et al.*, 2002). The difficultly in diagnosing preterm birth means that triaging women who present with symptoms of threatened preterm labour (TPTL) is a challenge (Carter *et al.*, 2018). Clinicians need to balance the risks of over-medicalising the majority of women

who will not deliver early (Iams *et al.*, 2002), with the small risk of preterm delivery for those discharged home (DeFranco, Lewis and Odibo, 2013). This challenge is compounded by current UK guidance which advises treating all women who present under 30 weeks' gestation in threatened preterm labour (TPTL) (NICE, 2015). This is despite that most women presenting in TPTL will not have delivered within seven days (Iams *et al.*, 2002), leading to psychological, economic and clinical implications of unnecessary interventions (Watson and Ridout, 2015; Watson *et al.*, 2019).

For women experiencing TPTL, alongside the potential frightening consequence of delivering their baby early, they describe conflicting emotions and feelings, which could affect the decisions they make about their care. Women's 'decisional conflict' (defined as a "state of uncertainty about a course of action" (The Ottawa Hospital Research Unit, 2003, p 2) may be increased by these contradictory emotions. For example these could include, their responsibilities (not wanted to be away from older children while needing to be an inpatient for their unborn child), the opposing advice they may receive from different clinicians, and/or feeling tense due to preterm labour symptoms while trying to stay calm for the unborn baby (Carter *et al.*, 2018). These emotions and feelings could all affect how a woman makes decisions about her care and inadvertently prevent joint decision making. To improve women's experiences and decisions around TPTL, clinicians need to be aware of their complex emotions and what issues are important to them. This would promote individualised care, shared decision-making, and reduce 'decisional conflict'. The uncertainty of TPTL and the decisions around TPTL that women face, often results in increased maternal anxiety, which is in itself a risk factor for premature delivery (Rich-Edwards and Grizzard, 2005; Latendresse, 2009; Christian, 2012).

The QUiPP app is a free, validated mobile phone application (app) that supports clinical decision-making by providing an individualised risk of delivery within clinically important time points (i.e.

within one, two and four weeks of testing, and before 30, 34 and 37 weeks' gestation (Kuhrt *et al.*, 2016; H. A. Watson *et al.*, 2017; Carter *et al.*, 2019; Watson *et al.*, 2019). The risk score is based upon the woman's risk factors for preterm birth i.e. history of preterm birth or prelabour preterm ruptured membranes (PPROM), invasive cervical surgery (e.g. LLETZ or cone biopsy) and multiple pregnancy) and her clinical test results (fetal fibronectin; fFN) and cervical length. Alongside generating a percentage risk score, the QUiPP app also provides the risk score in an infographic donut chart (see Figure 1), allowing the clinician to communicate with the woman in an easy to understand format. It is possible that informing women of their risk status using the QUIPP app [which incorporates fFN and/or cervical length, both of which have high negative predictive values (Tsoi *et al.*, 2005; Abbott *et al.*, 2013; Bruijn *et al.*, 2015)], would help to reduce anxiety in women with a lower risk of imminent delivery (Carter *et al.*, 2018).

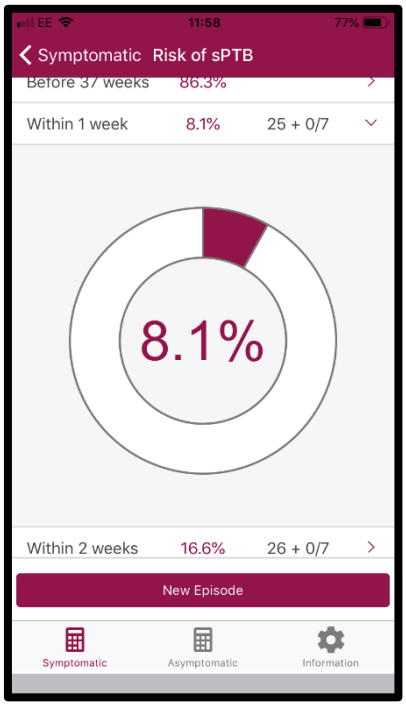

Figure 1 An image generated using the QUiPP app and knowledge of patient history, cervical length and fetal fibronectin measurement. Clicking on the numeric risk score at the top of the screen creates the associated donut risk infographic. In this example, a woman has an 8.1% risk of delivering with 1 week of the test and a 16.6% risk of delivering within 2 weeks. (Image used with consent from the QUiPP Team)

The importance of shared decision-making was highlighted by the 2015 Montgomery *v* Lanarkshire case, which emphasized the need for clinicians to present information in a meaningful way to

patients (Chan *et al.*, 2017). With this in mind, the QUiPP app was designed in collaboration with the Women's Health Academic Centre's Preterm Birth Studies Public and Patient Involvement (PPI) group and women participating in a qualitative part of the PETRA study (REC Ref.14/LO/1988), the findings of which are reported elsewhere (Carter *et al.*, 2018, 2019). The addition of the corresponding donut chart infographic in the app provides clinicians with a visual risk illustration, in addition to the % risk score, which they can use to enhance communication with women.

Aim

The aim of this study is to evaluate whether the use of the QUiPP app results in lowering levels of anxiety and decisional conflict in women experiencing threatened preterm labour.

To determine whether QUiPP reduced decisional conflict and improved shared decision making, we will investigate a subgroup of EQUIPTT trial participants' decisional conflict levels following their TPTL clinical assessment and compare findings between women where had the QUiPP app had been used and those where it had not been used. We will also investigate women's anxiety before and after TPTL clinical assessment.

Methods

Study Design

The EQUIPTT study (The Evaluation of the QUiPP app for Triage and Transfer) was a randomised cluster trial, carried out in 13 UK hospitals in London, the South East and the Midlands which aimed to evaluate the impact of the QUiPP app on inappropriate management for TPTL [REC Ref. 17/LO/1802;(Watson *et al.*, 2019)]. For the first six weeks, all 13 sites continued their normal triage practice for TPTL to allow a run-in period and collect baseline information. Seven sites were

randomised to the QUiPP app intervention (to use as a decision and communication tool) and six sites were randomised to the control (continued their normal practice) for a 9-month period. In the final 6-week period the QUiPP app intervention was established at the control sites. Ensuring the QUiPP app would be established at control sites incentivised all sites to participate in the study, regardless of how they were randomised. A subset of participants were asked to complete a questionnaire booklet which was used to evaluate decisional conflict and anxiety (supplementary file 1).

For this subset, women were enrolled through convenience sampling as they attended the maternity assessment centre with symptoms of TPTL between 26th March 2018 - 12th February 2019 (this included the 9-month study period and final 6-week period as described above). Women were recruited at all 13 sites for QUiPP app intervention as the control sites also used the app during the final 6-week period of the trial.

Data Collection

Women presenting to the maternity assessment unit with symptoms of TPTL were provided with a study patient information leaflet (supplementary file 2). If they wished to participate and complete the questionnaire booklet, written informed consent was taken by local site research midwives. Inclusion criteria required the women to be between 23^{+0} - 34^{+6} weeks' gestation with symptoms of TPTL (i.e. abdominal pain/back pain/tightenings). Women were excluded if they had a definitive diagnosis of labour (i.e. regular painful contractions with cervical change of more than 3cm on speculum or digital examination), confirmed rupture of membranes (on speculum examination) or significant vaginal bleeding.

Instrument

The 5-page questionnaire booklet comprised two sections and was designed in collaboration with the King's College London/St Thomas' Hospital Preterm Birth Studies Patient and Public Involvement (PPI) panel. The first section was completed before her assessment with a clinician and consisted of demographic and medical information (e.g. maternal age, ethnicity, gestational age, parity, previous preterm birth), symptoms, frequency of symptoms, severity of pain, and a pre-assessment anxiety score. The second section was completed following clinical assessment and included questions on what tests had been undertaken, whether the clinician had used the QUiPP app, the agreed care plan, questionnaires to determine post-assessment anxiety and decisional conflict scores and a free text comment box.

The pre and post-assessment anxiety scores utilised the Visual Analogue Scale for Anxiety (VASA) (Hornblow and Kidson, 1976), which has more recently been used in other maternity studies (Hepp *et al.*, 2016). The scale comprised of a 10 cm horizontal line with '*not at all anxious*' on the left-hand side, and '*extremely anxious'* on the right-hand side. Women were asked to mark how anxious they felt on the line. The local research midwife then measured the line with a 10 cm ruler to produce a VASA score (i.e. 2.5 cm = a score of 2.5). The scores could range from 0 (low anxiety) to 10 (high anxiety). For reliability, the questionnaire booklets were printed in advance for all sites to ensure that the visual anxiety scale printed as 10 cm exactly.

The Decisional Conflict Scale (O'Connor, 1995) involves 16 statements about the care offered to them following TPTL assessment. There are five response categories for each statement, ranging from 'strongly agree' to 'strongly disagree'. The local research midwife calculated and recorded a score to each response from 0 (strongly agree) to 4 (strongly disagree). The response scores for the 16 statements were calculated (summed, divided by 16, and multiplied by 25) as per the Decisional

Conflict Scale User Manual (The Ottawa Hospital Research Unit, 2003). The final calculated scores ranged from 0 (no decisional conflict) to 100 (extremely high decisional conflict).

Sample Size

Target recruitment for the EQUIPTT questionnaire sub-study was 300 participants (approximately 25 recruits per site), allowing for non-compliance and missing data. Based on previous research using visual analogue scale scores (Allred, Byers and Sole, 2010), a total sample size of 272 was required at standard levels of significance (alpha = 0.05 two-sided) to achieve 90% power to detect a 10% difference in mean VASA scores between sites with and without the intervention.

Data Analysis

Following completion of the questionnaire booklet, the local research midwives entered the anonymised data from the booklets onto a pre-prepared Microsoft Excel spreadsheet. The research midwives then cross checked the typed data with the paper copies. The spreadsheet was formatted with drop-down pre-populated answer boxes to minimise error during data transfer. The original paper questionnaire case report form was stored in the study folders (kept in a locked cupboard or room) at each local site, while the anonymous spreadsheet was sent electronically to the study coordinators. The study coordinators saved the anonymous spreadsheets on a secure King's College London network which was only accessible through an individual username and password. Once recruitment had finished at all sites, the study coordinators amalgamated the anonymous local spreadsheets to create one spreadsheet of all sites. The data were then analysed using Stata 15 Statistical Software (StataCorp, 2017).

The demographic data, symptom data, and agreed care plan were summarised. A repeated measures ANOVA (RMANOVA) undertaken for the VASA scores with time point (pre-assessment and

post-assessment) as the within subject variable. Planned independent sample t tests were used to compare change in VASA scores between sites using the QUiPP app and sites using standard care. Due to low compliance with the QUiPP app at some of the intervention sites, we also compared change in VASA scores for women aware the QUiPP app was used compared to women unaware of the QUiPP app used (independent sample t tests). This sub-analysis was required because if a woman is not involved in the decision-making using the app, it cannot influence her experience.

For analysis purposes we performed Chi-Squared tests on the proportion of women with a decisional conflict score of <25 and a score of ≥ 37.5, both between women aware of the app being used and women not aware of the app being used. These cut-offs were used as a meaningful difference between low and high levels of decisional conflict (The Ottawa Hospital Research Unit, 2003). Scores that are <25 are associated with implementing decisions, while scores that are high (≥ 37.5) are associated with decision delay or uncertainty around implementation (The Ottawa Hospital Research Unit, 2003; Kawaguchi *et al.*, 2013; Kim *et al.*, 2017).

Findings

Questionnaires were completed by 221 women from 12 of the potential 13 sites. After exclusions (Figure 2), 202 questionnaires were included in the analysis.

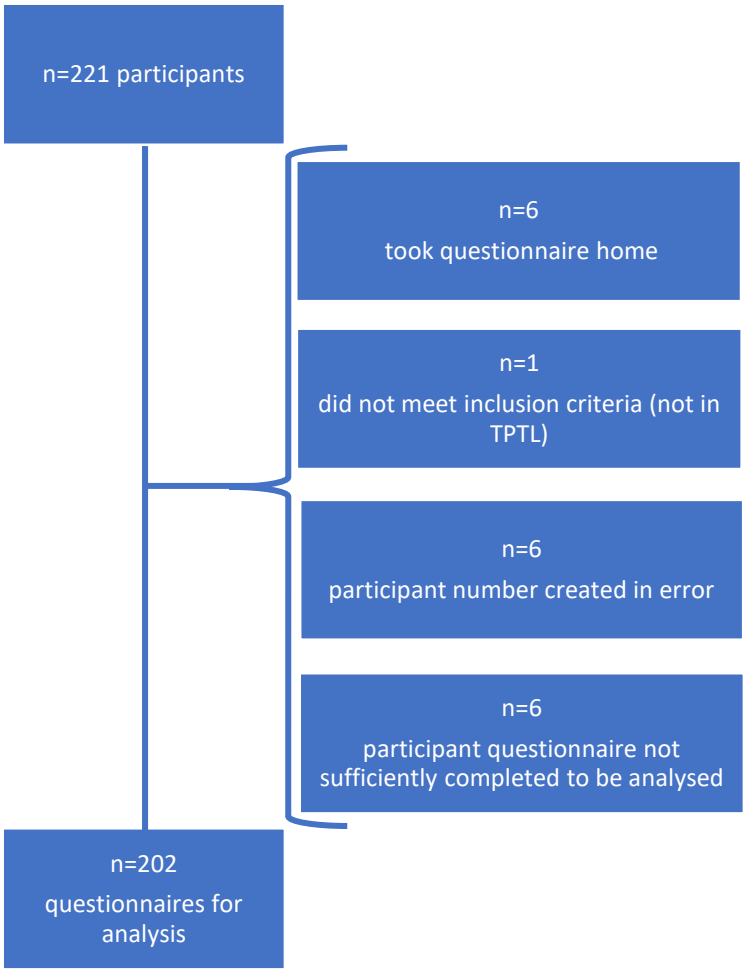

Figure 2 Flowchart of participant exclusions

Of the 202 women in threatened preterm labour completing the booklet and 81 (40%) were at sites randomised to using the app, and 121 (60%) were at sites randomised to standard care.

214

215 *Baseline characteristics*

216 Demographics, pregnancy characteristics, symptoms, test undertaken and care plan are described in

217 Table 1.

		Women at sites randomised to using QUiPP app	Women at sites randomised to using Standard care	All women
Total n (%)		81 (40)	121 (60)	202 (100)
Aware QUiPP App used in their care		48 (59)	11 (9)	59/202 (29)
Maternal age n (%)	**Under 20 years**	2 (3)	5 (4)	7/202 (4)
	21- 25 years	16 (20)	15 (12)	31/202 (15)
	26 - 30 years	24 (30)	45 (37)	69/202 (34)
	31 – 35 years	23 (28)	31 (26)	54/202 (27)
	36 – 40 years	15 (19)	20 (17)	35/202 (17)
	Over 40 years	1 (1)	5 (4)	6/202 (3)
Maternal ethnicity n (%)	**White**	41 (51)	74 (61)	115/201 (57)
	Asian	18 (23)	20 (17)	38/201 (19)
	Mixed	8 (10)	3 (2)	11/201 (5)
	Black	11 (14)	22 (18)	33/201 (16)
	Other	2 (3)	2 (2)	4/201 (2)
Primigravida n (%)		30 (37)	64 (53)	94/201 (47)
Symptoms present n (%)	**Tightenings**	63 (78)	82 (68)	145/201 (72)
	Abdominal pain	70 (86)	111 (93)	181/201(90)
	Back pain	40 (49)	73 (61)	113/200 (57)
Tests performed n (%)	**Fetal heart rate**	77 (95)	110 (91)	187/202 (93)
	Vaginal exam	60 (74)	81 (68)	141/201 (70)

	Fetal Fibronectin	66 (81)	65 (55)	131/199 (66)
	Cervical length	12 (15)	19 (17)	31/195 (16)
	Urinalysis	70 (86)	108 (90)	178/201 (89)
	Blood test	20 (25)	14 (13)	34/193 (18)
	Don't know	0 (0)	2 (2)	2 (1)
Agreed care plan n (%)	**Discharged home**	60 (75)	88 (74)	148 (74)
	Not discharged home	14 (18)	17 (14)	31 (16)
	Unsure	6 (8)	14 (12)	20 (10)

Table 1 Baseline characteristics of participants completing the questionnaire

The majority of participants were of White ethnicity (n=115, 57%) and within the 26-30 years age category (n=69, 34%). Women presented with a range TPTL symptoms and most had multiple symptoms. The majority (n=145/201, 72%) of women reported tightenings, usually described as irregular (n=76/201, 38%). Nearly all women (n=181/201, 90%) women reported having lower abdominal pain, with the majority (n = 95/201, 47%) describing this pain as moderate. More than half of all women (n=113/200, 57%) suffered with back pain. Baseline characteristics and symptoms were similar between sites that were randomised to using the app, and sites randomised to standard care.

Clinical tests performed

Nearly all women (n=187/202, 93%) reported that the fetal heart rate was checked. The majority of women reported that they had a vaginal examination (n=141/201, 70%), a cervico-vaginal swab for an fFN test (either qualitative or quantitative) (n=131/199, 66%) and urinalysis (n=178/201, 89%). Only 16% had a cervical length measurement, 18% a blood test and 1% of women said they did not

know what tests had been undertaken on them. Women in sites randomised to using the QUiPP app were more likely to have an fFN test (81% vs 55%) and a blood test undertaken (25% vs 13%).

Agreed care plan

Following TPTL assessment, the majority (74%, 148/202) of women knew the plan was to be discharged home. However, despite being seen by a clinician, 10% (20/202) reported that they were unsure what the agreed care plan was.

Aware QUiPP app was being used in their care.

Of the 202 women who completed the booklet, 59/202 (29%) were aware the QUiPP App had been used in their care. At sites randomised to using the app, 48 women [48/81 (59%)] were aware the QUiPP app had been used in their care. At sites randomised to standard care, 11 women [11/121 (9%)] said that they were aware the clinicians had used the QUiPP app in their care. This may be because clinicians at the standard care sites were aware of the app from working in other hospitals, despite the fact that they were currently working at a sites where using the app was not in their protocol.

To meet the aim of this study the research team analysed anxiety outcomes and decision conflict outcomes based on the randomisation of the site the women presented at (site randomised to using the app vs. site randomised to standard care) and on whether women were aware the app had been used in her care. This was decided because if a woman is unaware the QUiPP app was used in her care, it is unlikely to influence a woman's anxiety and/or decisional conflict.

Outcomes for women's perceptions of triage experience

Anxiety outcomes: VASA Scores

Four women did not provide a complete set of pre- and post-assessment VASA scores, leaving 198 participants available for analysis of anxiety levels. The outcomes are presented in Table 2.

A one-way RMANOVA showed a significant reduction in difference between VASA scores before and after assessment. $F(1,197) = 114.6$ $p = 0.000$. An independent t-test was run to assess differences in anxiety reduction between sites using and not using the QUiPP app. The results showed a greater reduction in VASA scores for women in QUiPP protocol sites (2.39 ± 0.36) compared to those receiving standard care (1.93 ± 0.23) but this was not statistically significant $t(196) = 1.13$, $p = 0.26$). We also used an independent t-test to compare the post-assessment VASA scores of women aware of QUiPP app use with to those unaware of its use. This showed a reduction that approached borderline statistical significance (2.75 ± 0.49 vs. 1.91 ± 0.21; $t(196) = 1.84$, $p = 0.07$).

Decisional Conflict Outcomes

Of the original 202 participants, 200 completed the decisional conflict outcomes questions, the results of which are shown in Figure 3.

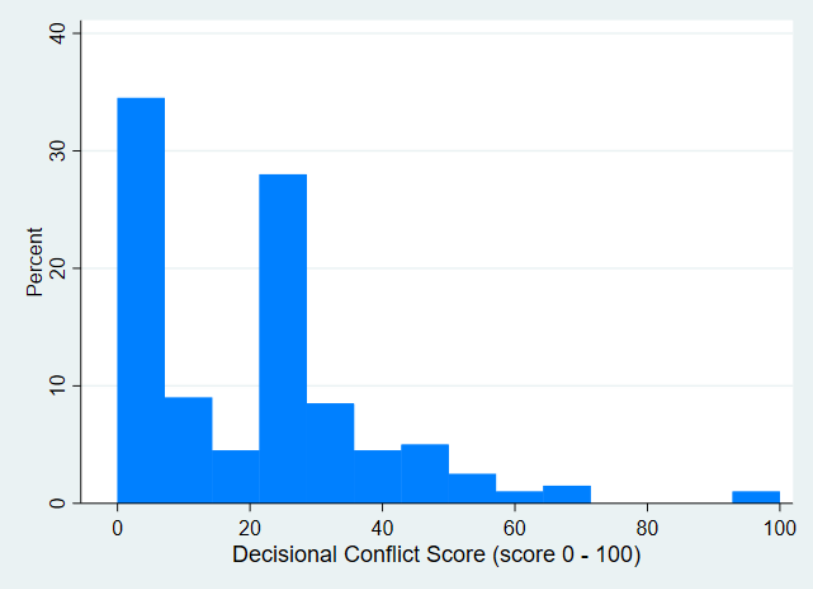

Figure 3 Histogram of women's decisional conflict scores (n=200). Decisional conflict score can be between 0-100, with a higher score indicating more decisional conflict.

Our cohort had low decisional conflict which was not normally distributed (Figure 3). Most women in our cohort had low decisional conflict (the median decisional conflict score in did not go above 23), which reinforced our analysis plan to divide decisional conflict scores into those <25 and those ≥ 37.5.

For women who were aware of the QUiPP app, 23% had high decisional conflict (a decisional conflict score ≥ 37.5) compared to 77% high decisional conflict in women who were not aware of the QUiPP app being used. A Chi-squared test found that this was not statistically significant. Full results are presented in Table 3.

	Women at EQUIPTT study site			**Women aware of QUiPP use during their visit**			**Total**	
	Site randomised to using QUiPP	**Site randomised to using standard care**	**P value**	**Yes**	**No**	**P value**	**Mean VASA score ***	**n**
Number of women with VASA scores (n)	78	120	-	47	151	-		198

Pre-assessment anxiety*	5.39 (4.82-5.95)	5.47 (5.05-5.90)	0.818	5.69 (4.90 - 6.48)	5.36 (4.98 - 5.74)	0.161	5.44 (5.10 - 5.78)	
Post-assessment anxiety*	3.00 (2.39 - 3.61)	3.54 (2.39-3.61)	0.161	2.94 (2.11 - 3.76)	3.45 (3.04-3.87)	0.244	3.33 (2.96 - 3.70)	
Difference in anxiety*	2.39 (1.67 - 3.10)	1.93 (1.48 - 2.38)	t(196) = 1.13 , *p* = 0.26)	2.75 (1.76 - 3.74)	1.91 (1.50 - 2.32)	t(196) = 1.84, *p* = 0.07)	2.11 (1.72 - 2.50)	

Table 2 Women's perception of outcomes: Visual Analogue Scale for Anxiety

**Data are presented as mean with 95% confidence intervals in brackets*

Difference in anxiety P values reported as: the observed t-value, (degrees of freedom), and significance level

	Women at EQUIPTT study site			**Women aware of QUiPP use during their visit**			**Median decisional conflict score for all women**	**Total**
	Site randomised to using QUiPP	**Site randomised to using standard care**	**P value**	**Yes**	**No**	**P value**		**n**
Total number of women with decisional conflict scores (n)	79	121	-	58	142	-	-	200
Median decisional conflict score	16	23	-	11	23	-	23	-
Decisional conflict scores < 25 n (%)	47 (43)	62 (57)	0.252	37 (34)	72 (66)	0.092	-	109
Decisional conflict scores ≥ 37.5 n (%)	8 (27)	22 (73)	0.119	7 (23)	23 (77)	0.458	-	30

Table 3 Women's perception of outcomes: Decisional Conflict Scores

Free text comments

Of the 202 questionnaires, 47 women provided written comments in a free text box. Full thematic analysis including coding of the data was not undertaken as most of the written comprised a few words or one sentence. Two researchers read the comments and assigned them into appropriate 'theme' categories. The final theme categories were agreed by consensus. Of the 47 comments, the majority (43/47) could be assigned into one of the five final categories which are described below:

1. Positive comments describing reassurance, good care and friendly staff (n=30).
2. Complaints about waiting times and its contribution to anxiety (n=3). One participant wrote that "the wait made me more anxious than I was before I arrived."
3. How having a history of preterm birth increases their anxiety when in threatened preterm labour (n=1).
4. Clinicians did not always involve women in the decision-making, so women did not feel they had options about their care (n=4).
5. Unclear communication and care plans (n=5). One woman was not sure why she was being admitted and another said "I was anxious when the neonatologist spoke to me as they said I was going into labour, and I started crying. There was a lack of communication, I haven't seen a senior doctor". Another respondent said that she "came here to find answers about the pain I am experiencing. No answers were given but thankfully I'm not in labour".

Of the remaining comments (4/47) that were not assigned into categories, 3 women used the comment box to write down their clinical diagnosis, and 1 woman used it to advocate for plastic speculums.

Discussion

This study assessed anxiety and decisional conflict in a large sample of women across 12 different hospital sites. Despite the fact our cohort is slightly younger than the UK average maternal age (Office for National Statistics, 2019), and is more ethnically diverse (Office for National Statistics, 2018), our findings are likely to be generalisable to the South of England, if not a wider UK pregnant population.

For all women presenting in threatened preterm labour in this study, anxiety was significantly reduced following clinical assessment compared to how they felt when they first arrived at the maternity assessment unit. This demonstrates that seeing a clinician, regardless of how they assess the woman, has a positive effect on reducing maternal anxiety. As most women in this study underwent fFN testing, our findings suggest that increased maternal anxiety is not associated with fFN use. Although this contrasts with previous reports (Shennan *et al.*, 2005), our findings align with a previous Canadian study where fFN testing was associated with maternal reassurance, even when women received a 'positive' fFN result (Peterson *et al.*, 2014).

For women with abdominal pain in pregnancy, anxiety is not just linked to her risk of preterm birth. While the QUiPP app provides reassurance to women who are less likely to be in preterm labour, if a woman is not provided with an acceptable alternative reason for her pain her anxiety may not be relieved. This represents a divide between a clinician's priorities (e.g. excluding the most severe pathology) and women's motivations for seeking help and need to 'make sense of the symptoms' (Carter *et al.*, 2018). This emphasises that the QUiPP app should be used as a decision support tool, helping clinicians and women to decide on the nature and timing of the most appropriate management, given the risk, rather than a simple rule in/rule out tool. Clinician's still need to communicate effectively with the woman regardless of her QUiPP score.

While the majority (30/47, 65%) of written comments spoke positively about their care, the remaining comments highlighted negative aspects of care which may have contributed to anxiety levels and decisional uncertainty. In some cases, women were unclear as to what occurred during their assessment either reporting that they were unsure of what the agreed care plan was or unsure what tests had been undertaken. This, again, highlights the divide between a clinician's priorities and women's motivations for seeking help. Clinicians need to ensure that they communicate effectively with women, understanding the importance of effective clinical explanation and counselling, including using clear language and checking that they have been understood (Ali, 2017).

While we used the recommended cut-offs when analysing decisional conflict scores (The Ottawa Hospital Research Unit, 2003) other researchers have used alternative thresholds (e.g. ≥ 25) (Thompson-Leduc *et al.*, 2016). As a score of 25 can be gained by answering 'neither agree nor disagree' for all questions, we felt that using the higher cut-off of ≥ 37.5 was more appropriate as an accurate measure of women with meaningful decisional conflict.

Limitations

Many additional factors (not just concerns about TPTL or awareness of QUiPP app use) may have affected a woman's anxiety which were not revealed by the questionnaire. This could include events that occurred prior to seeing a clinician, such as longer waiting times highlighted in our respondents written comments. Indeed, this is a factor known to increase anxiety (Thu *et al.*, 2015). Unfortunately, a limitation is that waiting times were not recorded as part of the study and would have been informative. Maternity services could explore waiting room interventions that have

shown promise elsewhere in reducing anxiety, such as music and visual art (Holm and Fitzmaurice, 2008) (Nanda *et al.*, 2012).

While there were reductions in anxiety and decisional conflict for women who were aware of the QUiPP app use, this failed to reach statistical significance. Clinicians at all sites were trained in the use of the QUiPP app and a training plan was developed to reinforce the guidance (Watson *et al.*, 2019). However, a limitation is that some women at the intervention sites were not aware of the QUiPP app being used. For example, of the 30 women in this study with high decisional conflict scores (≥ 37.5) only 23% (7/30) were aware of the QUiPP had been used, compared to 77% (23/30) who were not aware. The proportion of women with high decisional conflict was reduced in women who were aware the QUiPP app was used, compared to those unaware (p=0.458). Although this finding was not statistically significant it is possible that clinicians who discussed QUiPP scores with the women were more effective at communicating and involving women in their care. Our power calculation did not adequately allow for poor compliance with QUiPP app use, so a larger sample size may have resulted in findings reaching statistical significance

Enabling patients to understand risks can be a challenge but is crucial for clinicians when making shared decisions (Say, Robson and Thomson, 2003). Given working pressures, staff changeover and unfamiliarity with the app, it is possible that some staff were not confident in using the app as a communication tool, or using the donut chart infographic to visually represent risk (Figure 1). Greater reductions in women's anxiety or decisional conflict may be possible as clinicians' familiarity and confidence in the reliability of the prediction algorithms increases, along with enhanced dissemination and future roll-out of the QUiPP app (such as producing a QUiPP Implementation Toolkit for hospital sites to utilise). . While QUiPP may have potential to reduce anxiety and

decisional conflict for women experiencing TPTL, further research is needed to support generalised use of the tool.

While not all women were aware the QUiPP app was used in the sites randomised to using the QUiPP app, another limitation is that some women were aware of the QUiPP app being used despite presenting at sites randomised to standard care. Any potential affect on the results was minimised by analysing anxiety outcomes and decisional conflict outcomes by both site randomisation and women's awareness (Table 2 and Table 3).

Although the DCS has been used in several studies of decision making in pregnancy (Say, Robson and Thomson, 2003, 2011), we are not aware that it has been validated for research with pregnant women. While the DCS is sensitive to change and reliable (O'Connor, 1995), some argue its clinical applicability is limited as it does not incorporate the concept of pairing the choice of a patient with their values and preferences (Say, Robson and Thomson, 2011).

Whilst 221 women entered data into the questionnaire booklets, after exclusions only 202 were analysed. Of the included 202 booklets, not all were complete meaning overall numbers were less that the target recruitment of 300 questionnaire booklets that specified pre-study power for primary outcomes. Recruitment was challenging due to several factors. Local research midwives often work office hours when women in TPTL can arrive at any time. The study design meant women were required to complete the first section of the booklet before TPTL assessment, and the second section afterwards. Often by the time the research midwives arrived to explain the study and undertake consent, the women had already been assessed. Some women also took the questionnaire booklet home with them in error. On reflection we believe that recruitment could

have been improved in numerous ways including identifying a local study clinical champion who is known to the maternity staff. Additionally, recruitment could have been increased if more clinical staff were trained to take consent rather than relying solely on research midwives who are often committed to several projects at the same time. Both refinements could have resulted in women in TPTL arriving during the night or at weekends being consented to the study.

Conclusions

The QUiPP app has potential to reduce anxiety and decisional conflict in women who are aware that it is being used in their care, however further research is needed to support this. Clinicians could consider use of QUiPP as both a decision support and communication tool to appropriately target care where it is needed and to reassure women whose risk of preterm birth is low. Additional work is required to ensure clinicians are aware of the QUiPP app and optimise using it as a communication tool when counselling women.

447

448 **References**

449 Abbott, D. S. *et al.* (2013) 'Evaluation of a quantitative fetal fibronectin test for spontaneous preterm
450 birth in symptomatic women', *American journal of obstetrics and gynecology*. Elsevier, 208(2), pp.
451 122-e1. doi: 10.1016/j.ajog.2012.10.890.

452 Ali, M. (2017) 'Communication skills 1: benefits of effective communication for patients', *Nursing*
453 *Times*, 113(12), pp. 18–19.

454 Allred, K. D., Byers, J. F. and Sole, M. Lou (2010) 'The Effect of Music on Postoperative Pain and
455 Anxiety', *Pain Management Nursing*. W.B. Saunders, 11(1), pp. 15–25. doi:
456 10.1016/J.PMN.2008.12.002.

457 Bruijn, M. M. C. *et al.* (2015) 'Quantitative fetal fibronectin testing in combination with cervical
458 length measurement in the prediction of spontaneous preterm delivery in symptomatic women',
459 *BJOG: An International Journal of Obstetrics & Gynaecology*. Wiley Online Library.

460 Carter, J. *et al.* (2018) 'Threatened preterm labour: Women's experiences of risk and care
461 management: A qualitative study', *Midwifery*. doi: 10.1016/j.midw.2018.06.001.

462 Carter, J. *et al.* (2019) 'Development and validation of prediction models for the QUiPP App v.2: a
463 tool for predicting preterm birth in women with symptoms of threatened preterm labor', *Ultrasound*
464 *in Obstetrics & Gynecology*. John Wiley & Sons, Ltd, 0(ja). doi: 10.1002/uog.20422.

465 Chan, S. W. *et al.* (2017) 'Montgomery and informed consent: where are we now?', *BMJ*, 357, p.
466 j2224. doi: 10.1136/bmj.j2224.

467 Christian, L. M. (2012) 'Psychoneuroimmunology in pregnancy: Immune pathways linking stress with
468 maternal health, adverse birth outcomes, and fetal development', *Neuroscience and Biobehavioral*
469 *Reviews*. doi: 10.1016/j.neubiorev.2011.07.005.

DeFranco, E. A., Lewis, D. F. and Odibo, A. O. (2013) 'Improving the screening accuracy for preterm labor: is the combination of fetal fibronectin and cervical length in symptomatic patients a useful predictor of preterm birth? A systematic review', *American Journal of Obstetrics and Gynecology*. Mosby, 208(3), p. 233.e1-233.e6. doi: 10.1016/J.AJOG.2012.12.015.

Department of Health (2017) *Safer Maternity Care - The National Maternity Safety Strategy - Progress and Next Steps*. Available at: http://www.euro.who.int/__data/assets/pdf_file/0004/277735/Born-too-soon_preterm-birth-in-Europe-trends,-causes-and-prevention.pdf?ua=1.

Hepp, P. *et al.* (2016) 'Measuring the course of anxiety in women giving birth by caesarean section: a prospective study', *BMC Pregnancy and Childbirth*, 16(1), p. 113. doi: 10.1186/s12884-016-0906-z.

Holm, L. and Fitzmaurice, L. (2008) 'Emergency department waiting room stress: Can music or aromatherapy improve anxiety scores?', *Pediatric Emergency Care*. doi: 10.1097/PEC.0b013e31818ea04c.

Hornblow, A. R. and Kidson, M. A. (1976) 'The Visual Analogue Scale for Anxiety: A Validation Study', *Australian & New Zealand Journal of Psychiatry*. SAGE PublicationsSage UK: London, England, 10(4), pp. 339–341. doi: 10.3109/00048677609159523.

Iams, J. D. *et al.* (2002) 'Frequency of Uterine Contractions and the Risk of Spontaneous Preterm Delivery', *New England Journal of Medicine*. Massachusetts Medical Society, 346(4), pp. 250–255. doi: 10.1056/NEJMoa002868.

Kawaguchi, T. *et al.* (2013) 'Development and validation of the Japanese version of the Decisional Conflict Scale to investigate the value of pharmacists' information: a before and after study', *BMC Medical Informatics and Decision Making*, 13(1), p. 50. doi: 10.1186/1472-6947-13-50.

Kim, J. *et al.* (2017) 'Validation of the Decisional Conflict Scale for Evaluating Advance Care Decision Conflict in Community-dwelling Older Adults', *Asian Nursing Research*. Elsevier, 11(4), pp. 297–303.

494 doi: 10.1016/J.ANR.2017.11.004.

495 Kuhrt, K. *et al.* (2016) 'Development and validation of a tool incorporating cervical length and
496 quantitative fetal fibronectin to predict spontaneous preterm birth in asymptomatic high-risk
497 women', *Ultrasound in Obstetrics and Gynecology*, 47(1), pp. 104–109. doi: 10.1002/uog.14865.

498 Latendresse, G. (2009) 'The Interaction Between Chronic Stress and Pregnancy: Preterm Birth from A
499 Biobehavioral Perspective', *Journal of Midwifery and Women's Health*. doi:
500 10.1016/j.jmwh.2008.08.001.

501 Liu, L. *et al.* (2016) 'Global, regional, and national causes of under-5 mortality in 2000–15: an
502 updated systematic analysis with implications for the Sustainable Development Goals', *The Lancet*.
503 doi: 10.1016/S0140-6736(16)31593-8.

504 Nanda, U. *et al.* (2012) 'Impact of visual art on patient behavior in the emergency department
505 waiting room', *Journal of Emergency Medicine*. doi: 10.1016/j.jemermed.2011.06.138.

506 NHS England (2019) *z NHS Long Term Plan » Overview and summary*, *NHS England*. Available at:
507 https://www.longtermplan.nhs.uk/online-version/overview-and-summary/.

508 NICE (2015) 'Preterm Labour and Birth', *National Institute for Health and Care Excellence*. London:
509 National Institute for Health and Care Excellence (UK) (National Institute for Health and Care
510 Excellence: Clinical Guidelines), (November).

511 O'Connor, A. M. (1995) 'Validation of a Decisional Conflict Scale', *Medical Decision Making*. Sage
512 PublicationsSage CA: Thousand Oaks, CA, 15(1), pp. 25–30. doi: 10.1177/0272989X9501500105.

513 Office for National Statistics (2018) 'Regional ethnic diversity'. Available at: https://www.ethnicity-
514 facts-figures.service.gov.uk/uk-population-by-ethnicity/national-and-regional-populations/regional-
515 ethnic-diversity/latest.

516 Office for National Statistics (2019) *Birth characteristics in England and Wales: 2018*. Available at:

https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/livebirths/datasets/birthcharacteristicsinenglandandwales.

Patel, R. M. (2016) 'Short- and Long-Term Outcomes for Extremely Preterm Infants.', *American journal of perinatology*. United States, 33(3), pp. 318–328. doi: 10.1055/s-0035-1571202.

Peterson, W. E. *et al.* (2014) 'Women's perspectives of the fetal fibronectin testing process: a qualitative descriptive study', *BMC Pregnancy and Childbirth*, 14(1), p. 190. doi: 10.1186/1471-2393-14-190.

Rich-Edwards, J. W. and Grizzard, T. A. (2005) 'Psychosocial stress and neuroendocrine mechanisms in preterm delivery', in *American Journal of Obstetrics and Gynecology*. doi: 10.1016/j.ajog.2005.01.072.

Say, R. E., Robson, S. and Thomson, R. (2003) 'The importance of patient preferences in treatment decisions—challenges for doctors', *BMJ : British Medical Journal*, 327. doi: https://doi.org/10.1136/bmj.327.7414.542.

Say, R., Robson, S. and Thomson, R. (2011) 'Helping pregnant women make better decisions: a systematic review of the benefits of patient decision aids in obstetrics.', *BMJ open*. British Medical Journal Publishing Group, 1(2), p. e000261. doi: 10.1136/bmjopen-2011-000261.

Shennan, A. *et al.* (2005) 'Fetal fibronectin test predicts delivery before 30 weeks of gestation in high risk women, but increases anxiety', *BJOG: An International Journal of Obstetrics & Gynaecology*, 112(3), pp. 293–298. doi: 10.1111/j.1471-0528.2004.00420.x.

StataCorp (2017) 'Stata Statistical Software: Release 15'. College Station, TX: StataCorp LLC.

The Ottawa Hospital Research Unit (2003) *Decisional Conflict Scale User Manual*. Available at: https://decisionaid.ohri.ca/docs/develop/User_Manuals/UM_Decisional_Conflict.pdf.

Thompson-Leduc, P. *et al.* (2016) 'Prevalence of clinically significant decisional conflict: an analysis of

five studies on decision-making in primary care.', *BMJ open*. British Medical Journal Publishing Group, 6(6), p. e011490. doi: 10.1136/bmjopen-2016-011490.

Thu, H. *et al.* (2015) 'Factors Associated With Increased Anxiety in the MRI Waiting Room', *Journal of Radiology Nursing*. doi: 10.1016/j.jradnu.2015.04.009.

Tsoi, E. *et al.* (2005) 'Sonographic measurement of cervical length in threatened preterm labor in singleton pregnancies with intact membranes', *Ultrasound in Obstetrics & Gynecology*. John Wiley & Sons, Ltd., 25(4), pp. 353–356. doi: 10.1002/uog.1809.

Watson, H. . *et al.* (2017) 'Pregnancy Outcome Poster Abstracts', *BJOG: An International Journal of Obstetrics & Gynaecology*, 124, pp. 122–154. doi: 10.1111/1471-0528.14589.

Watson, H. . and Ridout, A. (2015) 'Re: Preterm labour: summary of NICE guidance', *BMJ : British Medical Journal*. Available at: https://www.bmj.com/content/351/bmj.h6283/rr.

Watson, H. A. *et al.* (2017) 'The QUIPP app: a safe alternative to a treat-all strategy for threatened preterm labour', *Ultrasound in Obstetrics & Gynecology*. Wiley Online Library, 50(3), pp. 342–346. doi: 10.1002/uog.17499.

Watson, H. A. *et al.* (2019) 'EQUIPTT: The Evaluation of the QUiPP app for Triage and Transfer protocol for a cluster randomised trial to evaluate the impact of the QUiPP app on inappropriate management for threatened preterm labour', *BMC Pregnancy and Childbirth*, 19(1), p. 68. doi: 10.1186/s12884-019-2210-1.