Women’s preferences for antenatal screening for Down syndrome in Northern Ireland: a choice experiment

Bernie Reid1 PhD, MSc, BSc, RGN, RM, RPHN, Marlene Sinclair2 PhD, MEd, BSc, DASE, RNT, RM, RN. Owen Barr3 PhD, SFHEA, RGN, RNLD, Frank Dobbs4 MB, BA, MD, FRCP, Grainne Crealey5 PhD, MSc, BSc.

1. Lecturer in nursing, School of Nursing, University of Ulster at Magee, Londonderry BT48 7JL Northern Ireland. Email: bb.reid@ulster.ac.uk
2. Professor of midwifery research, Institute of Nursing Research, University of Ulster, Newtownabbey BT37 0QB Northern Ireland. Email: m.sinclair1@ulster.ac.uk
3. Head of school, School of Nursing, University of Ulster at Magee, Londonderry BT48 7JL Northern Ireland. Email: o.barr@ulster.ac.uk
4. Emeritus professor of primary care, Institute of Postgraduate Medicine and Primary Care, University of Ulster at Coleraine, Londonderry BT52 1SA Northern Ireland. Email: f.dobbs@ulster.ac.uk
5. Senior health economist, Clinical Research Support Centre, Royal Belfast Hospitals, Grosvenor Road, Belfast BT12 6BA Northern Ireland. Email: grainne.crealey@csrc.s-nhs.uk

The authors would like to express their sincere thanks to the women who participated in this study. This study was funded by the Department of Employment and Learning, Northern Ireland as a PhD studentship.

Abstract

Background. Despite an increasing international trend towards the implementation of antenatal screening programmes for Down syndrome, there is currently limited information relating to women’s choices or preferences for such screening. Information about women’s preferences for and insight into the potential value they may derive from any change to screening programme design are essential components of evidence-based policy decision-making. This paper reports on a choice experiment undertaken to examine the preferences of a sample of pregnant and non-pregnant women of childbearing age with respect to antenatal screening for Down syndrome in Northern Ireland, where the offer of screening is not universal and selective abortion on the grounds of fetal abnormality is legally inaccessible.

Method. The choice experiment was conducted using an internet-based survey to obtain the preferences of a volunteer sample of 50 pregnant and 73 non-pregnant women for screening tests that varied according to six test characteristics or attributes derived from an earlier focused ethnographic study. These attributes were: source of screening information, time of test in pregnancy, accuracy of test results, cost of test as an ‘out-of-pocket’ expense, waiting time for test results, and risk of miscarrying a baby unaffected by Down syndrome as a result of subsequent diagnostic testing.

Results. Pregnant and non-pregnant women prefer screening tests for Down syndrome offering results with accuracy levels of 90% and above, and where the risk of subsequent diagnostic procedure-associated miscarriage is 2% and below. Women preferred screening tests offering more accurate results over no risk of miscarriage. Pregnant women place more value than non-pregnant women on tests carried out during the first trimester.

Conclusion. Policies for the implementation of antenatal screening programmes for Down syndrome must consider the preferences of pregnant and non-pregnant women to ensure that the needs of current and potential future service users are met. Effective pre-test counselling is essential to ensure that women understand the attributes of any screening test and its possible implications.

Key words: Choice experiment, Down syndrome, screening, women’s preferences, evidence-based midwifery

Background

Despite variations in national policies, there is an increasing international trend towards the implementation of antenatal screening programmes for Down syndrome (National Screening Unit, 2009; Boyd et al, 2008; NICE, 2008; American College of Obstetricians and Gynecologists Committee on Practice Bulletins, 2007). Ultimately, individual or population preferences determine the acceptance and participation in any screening programme, and therefore are integral to the realisation of expected population health gains from the programme (Kruisjaar et al, 2009). It is suggested that traditional decision modelling is insufficiently capable of including process effects, such as preferred source of screening information, and non-health outcomes, such as perceptions of risk associated with subsequent diagnostic interventions, and that they fail to identify the optimal screening programme (Drummond et al, 2005; Gyrd-Hansen and Søgaard, 2001). Choice experiments have been proposed as a means to improve systematic assessment of pregnant women’s preferences for, or acceptance of, screening programmes (Reid et al, 2008). In addition, choice experiments offer policy decision-makers viable alternatives and complements to existing methods of valuation and preference elicitation when planning the implementation and delivery of maternity care interventions (Lanscar and Louviere, 2008).

A small number of choice experiments have described the preferences of pregnant women or current service users for antenatal screening for Down syndrome (Seror and Ville, 2009; Lewis et al, 2006; Bishop et al, 2004; Spencer and Atiken, 2004). Although incorporating pregnant women’s preferences and delivering woman-centred maternity care is the centrepiece of evidence-based policy initiatives (Department of Health, Social Services and Public Safety (DHSSPS), 2012; Department of Health (DH), 2007), there has been a tendency to overlook the preferences of non-pregnant women of childbearing age. It has been argued that the preferences of non-pregnant women of childbearing age who may potentially become engaged in screening decision-making (potential service users) together with information
on the potential value derived by women from any change to screening programme design may further assist policy decision-makers to ensure that a screening programme for Down syndrome represents and serves the needs of contemporary society and to predict the programme’s future use (Petrou and McIntosh, 2011; Bryan and Dolan, 2004). This raises the methodological issue of presenting a choice experiment to women with perhaps no experience of screening. It has nonetheless been reported that by applying a ‘think aloud’ qualitative technique to choice experiment responses, respondents are capable of expressing preferences for unfamiliar goods (Ryan et al, 2009a).

Northern Ireland provides a unique setting where the offer of antenatal screening for Down syndrome is not universal and selective abortion on the grounds of fetal abnormality is legally inaccessible. There is no clear regional policy relating to screening despite a UK-wide screening strategy recommending that all pregnant women should be offered screening as part of a screening pathway, which includes further potential choice options relating to diagnostic testing and selective abortion (UK National Screening Committee, 2009). Local variations exist not only in what screening test is offered but also in those pregnant women to whom it is offered. Despite ongoing debate and recent public consultations, the option of selective abortion on the grounds of positive fetal diagnosis of Down syndrome remains prohibited by existing abortion legislation in Northern Ireland (DHSSPS, 2009). Women who seek selective abortion must therefore travel overseas, usually to England and Wales.

The aim of this choice experiment was to describe the preferences of pregnant and non-pregnant women of childbearing age (18 to 50 years) with regard to antenatal screening for Down syndrome in Northern Ireland.

Method

Study sample

A volunteer sample of pregnant women (at any stage during pregnancy) and non-pregnant women of childbearing age (18 to 50 years) was recruited via advertisements placed in regionally available journals, radio and websites, inviting women to access the study website if they were interested in participating. The authors decided not to recruit through organisations or online communities specifically dedicated to antenatal screening and/or Down syndrome, which may potentially have attracted respondents who had a specific interest in the topic (Hewson et al, 2003).

The sample size was not pre-determined, as no specific sample size power calculator for adaptive choice experiments exists (Orme, 2006). Nevertheless, a rule of thumb for sample size power calculator for adaptive choice experiments interest in the topic (Hewson et al, 2003).

To achieve content validity, attributes must reflect as closely as possible the key drivers of women’s screening choices while also reflecting the interests of healthcare policy decision-makers and providers (Reid et al, 2008). The six key attributes used in this choice experiment consisted of five process and one outcome attributes (see Table 1) derived from an earlier focused ethnographic diary study exploring the screening decision-making processes of pregnant women (n=21) and a focus group study examining the factors influencing screening decisions from the perspectives of pregnant women (n=14), midwives (n=19) and mothers of children with Down syndrome (n=16) (Reid, 2010). This study found that women’s confusion.

Identifying the key characteristics or attributes of screening

To achieve content validity, attributes must reflect as closely as possible the key drivers of women’s screening choices while also reflecting the interests of healthcare policy decision-makers and providers (Reid et al, 2008). The six key attributes used in this choice experiment consisted of five process and one outcome attributes (see Table 1) derived from an earlier focused ethnographic diary study exploring the screening decision-making processes of pregnant women (n=21) and a focus group study examining the factors influencing screening decisions from the perspectives of pregnant women (n=14), midwives (n=19) and mothers of children with Down syndrome (n=16) (Reid, 2010). This study found that women’s confusion.

Table 1. Key attributes of screening for Down syndrome identified from diary and focus group studies

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Source of screening information</td>
<td>Accurately screening test results (<em>accuracy</em> used by women to denote sensitivity)</td>
</tr>
<tr>
<td>Timing of screening test in pregnancy</td>
<td>Cost of screening as an ‘out-of-pocket expense’</td>
</tr>
<tr>
<td>Accuracy of screening test results</td>
<td>Waiting time for screening test results</td>
</tr>
<tr>
<td>Risk of miscarrying a baby unaffected by Down syndrome as a result of subsequent diagnostic testing</td>
<td>Only included/involved in sub-experiment 2</td>
</tr>
</tbody>
</table>

© 2014 The Royal College of Midwives. Evidence Based Midwifery 12(4): 112-120
with respect to the difference between screening and diagnostic testing was reflected in their associating the risk of miscarriage with screening rather than with diagnostic testing, often declining screening for this reason. The researchers therefore included ‘risk of miscarrying a baby unaffected by Down syndrome as a result of subsequent diagnostic testing’ as an outcome attribute due to concerns that women would express a very strong preference for 0% ‘risk of miscarriage’, and therefore be unwilling to trade among other attributes (Ryan et al, 2009b; Lloyd, 2003). Such a dominant preference would render the calculation of willingness to ‘trade’ or marginal rates of substitution meaningless. Accordingly, the authors decided to generate a choice experiment consisting of two sub-experiments. The first sub-experiment included ‘risk of miscarriage’ as an attribute while the second sub-experiment included it. In this way, it was possible to ascertain the impact of ‘risk of miscarriage’ upon women’s screening decisions.

Assigning levels to attributes
Assigning levels to the attributes involved specifying for each attribute a number of levels or values that were numerical or ordinal. The levels assigned to each attribute had to be plausible, reflecting what is relevant to current screening experience (Ryan and Farrar, 2000). Most importantly, the attribute levels had to be capable of being traded, in that the range of levels would provide enough variation while simultaneously being narrow enough to create competitive choices for women (Ryan, 1999). For this choice experiment, attribute levels were derived from the findings of diary and focus group studies and from existing research literature (Mujezinovic and Alfirevic, 2007; Heyman et al, 2006; Lewis et al, 2006; Bishop et al, 2004).

Data collection
Preference data were collected by means of an internet-based survey developed using Sawtooth Software CBC version 6.4 and hosted by a university-based website. Women were firstly asked to provide demographic details. They were then presented with a series of hypothetical choice scenarios created by combining different levels of each attribute. Finally, women’s views on screening methods and abortion were ascertained. By asking women to consider specific screening alternatives, scenarios made trade-offs between competing alternatives explicit (Fraenkel et al, 2004). As recommended (Mangham et al, 2008), a fraction of the full set of scenarios was used with 16 scenarios being presented to women (each sub-experiment consisting of eight scenarios). Each scenario contained two test alternatives together with a ‘No, I don’t wish to have either of these two tests’, as in practice women may choose to decline screening. The scenarios were generated using the same Sawtooth Software. This programme facilitated approximate orthogonality, that is, attribute levels were chosen independently of other attribute levels so that each attribute level’s utility could be measured independently of all other effects (Street et al, 2008). To enhance reliability, the software allowed for the viewing of only one choice set at a time (initially those choice sets with this attribute) and thus prevented women from returning to already completed scenarios and making choice changes in light of decisions made in subsequent scenarios.

The survey was piloted with seven women to determine ease of understanding and completion. Following this pilot, changes were made to the wording in the survey describing the attributes, but not the overall design of the choice experiment. Data collected during the pilot study were not included in the main study.

Statistical analysis
Preference data derived from the choice experiment were imported into Sawtooth Software SMRT and merged with the participant demographic data set. Data from each sub-experiment were analysed separately and then merged to ascertain the impact of introducing the outcome attribute, risk of miscarriage, upon women’s screening preferences. Sawtooth Software SMRT facilitated the analysis of data using three methods. Firstly, estimations for main effects and for two-way interaction effects were calculated using counts or proportions. A main effect is the direct independent effect that each attribute has on choice (Hensher, 2005). A two-way interaction effect is the effect upon choice of combining two attributes with such an effect not necessarily having been observed had each of the attributes been estimated separately (Hensher, 2005). This method was also useful for understanding variations in patterns of demand for specific attribute levels. Secondly, a multinomial logit (MNL) model was used to estimate the utilities that represent the contribution of each attribute level to the overall choice outcome. Such utilities were analogous to the relative desirability or value of each attribute level studied. Attribute levels with higher utility values were considered more desirable or beneficial by participant women. The computation started with estimates of zero for all effects or utilities. Pertinent information was found in the relative differences between the utilities.

Thirdly, ‘what if’ screening tests were then built to assess the impact of varying screening attribute levels on women’s preferences. In other words, it was possible to generate different combinations of attribute levels and sum their corresponding utilities in order to estimate the overall utility that women would derive from different screening tests. It was also possible to predict the percentage of women projected to choose each test. In this way, marginal rates of substitution were calculated by varying the levels of one attribute within a particular ‘what if’ screening test, the levels of all other attributes being kept constant. In this way, it was possible to compute the incremental increase or decrease in utility for each level of an attribute.

Ethical approval
Ethical approval was obtained from the Office for Research Ethics Committees Northern Ireland (reference 07/NIR01/40).

Results
Respondent characteristics
Of the 142 women who accessed the survey, 127 (89.4%)
completed the first sub-experiment and 123 (86.6%) completed both sub-experiments. A total of 50 women (40.7%), who completed both sub-experiments, reported an ongoing current pregnancy and 23 (46%) reported being offered screening for Down syndrome, with 13 (56.5%) accepting the test.

When asked about method of screening, 79 (64.2%) of the 123 women who completed both sub-experiments stated that they would prefer a combined blood (biochemical) and ultrasound scan screening test. A further 36 of these women (29.3%) expressed a preference for screening to be carried out by ultrasound scan. Only eight women (6.5%) expressed a preference for blood (biochemical) testing.

In the event of a positive antenatal diagnosis of fetal Down syndrome, 76 (61.8%) of the women who completed both sub-experiments reported that they would continue with the pregnancy, 14 (11.4%) reported that they would not continue with the pregnancy and 33 (26.8%) did not know what they would do. Women’s attitudes were further examined by age (less than or greater than 35 years, p=0.4475), religious affiliation (Roman Catholic or other values obtained for some categories did not allow for Protestant or other, p=0.7159), and educational attainment (degree level or above/other, p=0.4432). These factors were not found to be statistically significant.

Estimation of main effects and two-way interactions
Main effects were calculated for all attributes studied with timing of the screening test during pregnancy and the accuracy of screening test results being found to be statistically significant when making choices about screening for Down syndrome (p<.01). When making choices about screening in conjunction with future diagnostic testing, women considered these two attributes together with the risk of procedure-associated miscarriage (p<.01) to be the most important of those attributes studied. Two-way interaction effects were calculated for all attributes but only one such effect, namely the impact of timing of the screening test on accuracy of screening test results, was found to be statistically significant for deciding about screening but, somewhat surprisingly, not significant when deciding about screening in conjunction with future diagnostic testing.

Pregnant and non-pregnant women’s utilities
Figure 1 charts the respondent utilities for each attribute and attribute level studied. The bars of the left-hand side represent negative utilities and those on the right-hand side represent positive utility.

For both sub-experiments, women derived greatest utility or value from screening tests that offered 100% accurate test results. Women’s utility for such test results increased by 7.8% when the prospect of future diagnostic testing became a decision-making factor. The greatest gains in utility were achieved for tests that offered results with accuracy levels of 90% and above. By contrast, women had an aversion to tests that offered results with accuracy levels of 80% and below. When deciding about a screening test only, women did not have a preference for, or against, a second trimester test.
Nevertheless, women derived some minimal utility from tests carried out during the second trimester when deciding about screening and diagnostic testing. Yet, even for these women approximately 2.5 times greater utility was derived from tests carried out during the first trimester as compared to the second trimester. With respect to source of screening information, women in both sub-experiments expressed positive preferences for face-to-face contact with midwives or hospital doctors. Midwives were women’s preferred source of screening information. The utility women derived from information provided by midwives increased by 11.2% when the potential future risk of diagnostic procedure-associated miscarriage was introduced. Women were indifferent to screening information provided by CD or DVDs, but tended to avoid information provided by non-healthcare professionals, hospital websites and information leaflets. Women in both sub-experiments expressed positive preferences for screening tests that provided results within 24 hours or two days. Interestingly, women derived greater utility from waiting times of two days as compared to waiting times of 24 hours. This may indicate that women perceived two days to be a more realistic waiting time for test results. Women were found to be more responsive to cost when deciding about a screening test only. Positive utility was derived from a nominal cost of £10 with comparable negative utility being derived from a cost of £120. By contrast, women were indifferent to cost when deciding about screening in conjunction with future risk of diagnostic procedure-associated miscarriage.

When the relative differences between accuracy and risk of diagnostic procedure-associated miscarriage were examined for the second sub-experiment, women were found to derive greater utility from 100% accurate test results as compared to a 0% risk of miscarriage. Greater negative utility was derived from 55% accurate test results than a 10% risk of miscarriage. Women valued screening tests where the risk of future diagnostic procedure-associated miscarriage was 2% and below. Women felt that a 0% risk of miscarriage was almost six times more valuable than a 2% risk. Negative preferences were expressed for tests when the risk of miscarriage was greater than 3% as compared to the positive utility derived from tests where the risk of miscarriage ranged from 0% to 2%. As anticipated, women appeared to derive greater utility from the ‘No test’ alternative when the risk of procedure-associated miscarriage was included as a decision-making factor.

**Marginal rates of substitution**

Figure 2 represents current Down syndrome screening practice in Northern Ireland in terms of attribute levels. Women’s utilities for each of these levels are those already presented in Figure 1. Cost was included as pregnant women attending some maternity units until recent years were required to pay approximately £30 for a screening test. The negative cumulative utility derived from a screening test only was of a much greater magnitude than that derived from screening and diagnostic testing. A 1% risk of miscarriage increased the cumulative utility derived from testing by 93%. While the risk of diagnostic procedure-associated miscarriage may generally be perceived negatively, the authors suggest that this finding supports the view that women valued additional information about risk of miscarriage provided the potential for such risk actually occurring remained low.

Women’s utilities for current practice were then used as comparators to assess how controllable and uncontrollable changes or ‘best’ case scenarios in attribute levels impacted on women’s utilities for screening test alternatives. For example, realistic controllable changes incorporated the inclusion of a first trimester screening test with 85% accurate test results thereby reflecting the timing and detection rate of the existing ‘combined’ screening test (biochemical and nuchal translucency scanning) (List et al, 2006). Figure 3 charts the magnitude of difference between the cumulative utilities associated with such simulated tests and the current practice or base case scenarios (denoted as zero on the x-axis). Importantly, all simulated tests improved upon the cumulative utility of current practice.

The findings indicate that a reduction in the cost of screening tests to a nominal £10, while keeping all other attribute levels constant, would cause an increase in women’s cumulative utility from that of current practice for a screening test only. When again compared to current practice, the cumulative utility derived from reducing cost would be much less when screening is considered in conjunction with diagnostic testing. As compared to current practice which offers second trimester screening, it is predicted that offering screening tests only during the first trimester of pregnancy would cause an increase in women’s cumulative utility. Reducing the waiting time for screening test results to two days in conjunction with first
trimester screening would result in women deriving positive cumulative utility from tests regardless of whether risk of miscarriage is a decision-making factor. Indeed the greatest leverage in controllable cumulative utility for screening tests only is provided by first trimester screening offering results within two days. By contrast, it was found that a moderate increase in the accuracy of screening test results from 75% to 85% would provide the greatest leverage in controllable cumulative utility when screening was considered in conjunction with diagnostic testing. When faced with potential decision-making that includes risk associated with future diagnostic testing, efforts to improve the accuracy of screening tests are valued more by women than efforts directed towards providing less expensive, earlier and more expedient screening test results.

Although such realistic or controllable changes to current practice may be desirable to women, they fall well short of the optimal cumulative utilities as described by the maximum achievable or ‘best case’ scenarios for both sub-experiments.

**Sub-group analysis for pregnant and non-pregnant women**

Demographic characteristics of age (less than or greater than 35 years, p=0.2853), marital status (married or other, p=0.5167), and educational attainment (degree level or above/other, p=0.3331) were not statistically significantly different for pregnant and non-pregnant women who had completed both experiments.

With respect to the method of screening, a higher proportion of pregnant women (34%) expressed a preference for screening using ultrasound scanning only as compared to non-pregnant women (26%). By contrast, a higher proportion of non-pregnant women (68.5%) preferred combined ultrasound and maternal serum screening as compared to pregnant women (58%). In the event of a positive fetal diagnosis of Down syndrome, 56% of pregnant women stated that they would continue with the pregnancy as compared with 65.7% of non-pregnant women. This finding appears counter-intuitive as the subjective attachment to the fetus associated with an ongoing pregnancy would have been anticipated to negate against women’s consideration of selective abortion (Johansson-Stenman and Svedsater, 2007). Nevertheless, a greater proportion of pregnant women (32%) expressed uncertainty with respect to decision-making in the event of a positive diagnosis of Down syndrome as compared to non-pregnant women (23.3%). Pregnant and non-pregnant women’s attitudes towards Down syndrome and selective abortion were not statistically significantly different (p=0.737).

Main effects were estimated for all attributes studied with both pregnant and non-pregnant women reporting that the accuracy of screening test results had a statistically significant impact on their screening choices. Pregnant women amplified an increasing demand for more accurate test results when making decisions about a screening test only or a screening test in conjunction with diagnostic testing. By contrast, non-pregnant women expressed indifference towards tests offering between 75% and 85% accuracy. These women tended to ‘shy away’ from the offer of 100% accurate results when considering a screening test only, preferring instead a test offering 95% accuracy. Pregnant and non-pregnant women were found to have different preferences with respect to timing of screening tests during pregnancy. Pregnant women expressed significant positive preferences for earlier screening when deciding about a screening test only or a screening test in conjunction with diagnostic testing. These women expressed clear preferences for first trimester screening and indicated increasing demand for screening as pregnancy progressed. Non-pregnant women’s preferences for timing of tests were statistically insignificant when considering a screening test only. They had no strong preferences for first or second trimester screening when diagnostic testing became a decision-making factor.

The associated risk of miscarrying a baby unaffected by Down syndrome as a result of diagnostic testing was found to significantly affect decision-making for both pregnant and non-pregnant women (p<0.01). Pregnant women were indifferent to risk of miscarriage between 1% and 3% but indicated greatest reduction in demand for screening tests when such risk exceeded 3%. By contrast, non-pregnant women displayed a steadily decreasing demand for screening as risk of miscarriage increased. Pregnant women expressed slightly greater preferences for obstetricians rather than midwives as sources of information when considering diagnostic testing in conjunction with screening tests. This was not the case for non-pregnant women.

**Discussion**

**Main findings**

The findings of this study show that, for the attributes studied, both pregnant and non-pregnant women preferred combined biochemical and ultrasound scan screening tests conducted during the first trimester of pregnancy offering results with accuracy levels of 90% and above, and where the risk of subsequent diagnostic procedure-associated...
miscarriage was 2% and below. They also preferred to have screening information provided by healthcare professionals during face-to-face consultations, screening tests that provided results within two days, and to pay a nominal cost of £10 for any such test. Women who participated in this study valued information about the risk of potential subsequent diagnostic procedure-associated miscarriage provided that risk remained low. However, they preferred screening tests offering more accurate results over no risk of miscarriage. Although women in the diary and focus group studies expressed a very strong preference for 0% ‘risk of miscarriage’ (Reid, 2010), the findings of this study point to value of choice experiments in ‘teasing out’ women’s preferences. Given that the majority of women who participated reported that they would continue with the pregnancy in the event of a positive antenatal diagnosis of fetal Down syndrome, it may be suggested that preferences for earlier and more accurate screening test results were indicative of a straightforward desire for information or a ‘need to know’. While this study demonstrated that pregnant women and non-pregnant women of childbearing age shared similar relative values regarding accuracy of screening test results and the risk of diagnostic procedure-associated miscarriage, they differed in the value they attached to the timing of screening tests, with pregnant women placing more value than non-pregnant women on tests carried out during the first trimester.

Strengths and limitations
A major strength of this study is the use of screening attributes derived from focused ethnographic study data, thereby reflecting as closely as possible the key drivers influencing women’s choices (Reid et al., 2008). While this study helps to ‘tease out’ women’s screening choices, it is not possible to generalise the findings as women in this study were self-selected and selection bias may have impacted on the results. Future research would benefit from a larger sample size incorporating a wider demographic profile.

In addition, the study raises methodological issues with respect to the development and use of choice experiments. Firstly, we have used a within-sample design to ascertain the impact of the outcome attribute, ‘risk of miscarriage’, upon women’s screening decisions by constructing two experiments completed by the same sample of women. A between-sample experiment may have provided an alternative design. However, there is still some debate in the literature as to whether hypothetical bias is much minimised with a between-sample over a within-sample study design (Johansson-Stenman and Svendsater, 2007; List et al., 2006; Lusk and Schroeder, 2003). The recent ISPOR task force report (Johnson et al., 2013), which evaluated alternative approaches to experimental design for choice experiments, made no reference to a preference for one such design over the other. Secondly, it is probable that the statistical power to estimate main effects may have been enhanced by utilising a minimum sample size for each of the two experiments rather than a combined minimum sample size. However, sample size calculations are recognised as being particularly difficult for choice experiments (Rose and Bliemer, 2013; Bridges et al., 2011). Traditional orthogonal designs and existing sampling theories do not adequately address the issue, and no single strategy is available that can be used to state unequivocally what sample size will be required for a choice experiment (Rose and Bliemer, 2013). Hence researchers have had to resort to rules of thumb as in this study.

Interpretation
This study is the first to describe the preferences of pregnant and non-pregnant women of childbearing age for antenatal screening for Down syndrome in Northern Ireland. The authors are aware of no other studies that have used discrete choice methods to quantify the preferences of both these groups of women. In addition, this study was undertaken in a unique setting where the offer of screening is not at present universal and selective abortion on the grounds of positive fetal diagnosis of Down syndrome is legally inaccessible.

An earlier study used a screening options questionnaire to survey 1127 women attending antenatal clinics in England for Down syndrome screening (Spencer and Aitken, 2004). Adopting a false-positive rate of 5% and tests offering 90% accurate results, it was found that 75% of women preferred first trimester screening, which used serum screening and ultrasound scan, as compared to integrated first and second trimester screening (24%) and second trimester screening (1%). Women were also more likely to prefer a one hour result option. These findings are consistent with the findings reported in this study in terms of method and timing of screening. However, the study did not explore the impact that false-positive rates might have on women’s choices. The authors therefore cannot discount that if accuracy of screening test results were fixed at the same level for all tests and false-positive rates varied, women may show a preference for tests with the lowest follow-up rates.

Two further studies investigated how women’s and healthcare professionals’ preferences for screening differed, not only in terms of timing and accuracy of test results, but also in terms of the risk of subsequent diagnostic procedure-associated miscarriage. Using conjoint analysis, a study carried out in England found that while both groups valued screening tests carried out during early pregnancy with associated highly accurate test results, professionals valued earlier tests more highly than women (Bishop et al., 2004). Similar to the findings of the authors’ study, women had a weaker preference for an earlier test relative to accuracy of test results. These findings were confirmed by a replicative Australian study (Lewis et al., 2006). However, both studies conflated the risk of diagnostic procedure-associated miscarriage with screening tests and reported that women’s screening preferences were largely determined by this attribute. By contrast, the authors’ study found that women’s preferences for screening tests only would be determined by offering more timely results whereas women’s preferences for screening considered in conjunction with diagnostic testing would be determined by offering more accurate test results.
One reason forwarded by earlier studies for women's preference for earlier screening relates to the possibility of a first trimester selective abortion in the event of a positive diagnosis of fetal Down syndrome (Lewis et al, 2006; Bishop et al, 2004; Spencer and Aiken, 2004). However, the authors’ study found that only 11.4% of pregnant and non-pregnant women of childbearing age would not continue with such a pregnancy. Even when those women whose uncertainty is taken into consideration, such a proportion is much lower than the previously reported 48% of the general population of Northern Ireland who conditionally approved of selective abortion (Halman et al, 2001). A retrospective study conducted in Uruguay, where abortion is also legally inaccessible, found that 89% of women would terminate a Down syndrome affected pregnancy (Quadrelli et al, 2007). Although, such diverse findings may be partially explainable in terms of sampling issues and study timing, further sensitive exploration of women’s views regarding selective abortion is warranted.

This study’s finding that women derived greater utility from paying a nominal cost for screening as compared to the approximate or higher costs is contradictory to other reported findings (Seror and Ville, 2009). Using hierarchical cluster analysis, it was found that women in France were willing to pay on average €96 for a screening test, more than twice the sum of €39 which was charged and subsequently reimbursed by the French health insurance system (Seror and Ville, 2009). In addition, women deciding about screening in conjunction with future risk of diagnostic procedure-associated miscarriage were willing to pay €150 or more for screening, whereas women in the authors’ study were indifferent to cost when making such decisions. Women in some countries may be accustomed to paying ‘out-of-pocket’ expenses in an insurance-based healthcare system as compared to women in a ‘national’ healthcare system where services are free at the point of use.

The authors’ study should also be considered within the context of current screening policy in the UK, which recommends that all pregnant women be offered the ‘combined test’ and, if accepted, this test should be performed before the end of the first trimester (UK National Screening Committee, 2011). This test should provide women with results offering a detection rate of more than 90%, for a screen positive rate of less than 2% (of affected pregnancies). The waiting time for results is to be determined locally. The risk of subsequent diagnostic procedure-associated miscarriage is recognised as being between 1% and 2% (UK National Screening Committee, 2006). Such policy recommendations concur with the preferences of women in this study. However, these recommendations are not currently implemented in Northern Ireland where an ad hoc offer of second trimester screening is made to high-risk pregnant women. In addition, women’s preferences for screening tests offering results with accuracy levels of 90% and above is of particular importance given that women in Northern Ireland are screened using the triple test with a detection rate of 67% to 70% for a false-positive rate of 3% (UK National Screening Committee, 2011; List et al, 2006). However, the combined test as recommended involves a two-stage testing approach during the first trimester and this study did not establish women’s acceptability of such an approach. Finally, these findings pose a dilemma for women, healthcare professionals and policy decision-makers where the lack of access to selective abortion in Northern Ireland appears to be incongruent with national and international screening pathways (National Screening Unit, 2009; UK National Screening Committee, 2009; NICE, 2008; American College of Obstetricians and Gynecologists Committee on Practice Bulletins, 2007), and the principles of equity and choice which underpin current maternity care policy (DHSSPS, 2012; DH, 2007).

### Conclusion

Women’s preferences for antenatal screening tests for Down syndrome involve complex trade-offs and interactions between attributes pertaining both to the process and outcome of screening. Policies for the implementation of antenatal screening for Down syndrome need to be informed by this study and future research which consider these preferences to ensure that the needs of current and potential future service users are met. Particular attention needs to be given to the development of a clear regional policy which is congruent with the current national UK screening strategy. Effective pre-test counselling and informed consent processes are essential to ensure that women understand the attributes of any screening test and its possible implications.

### References


© 2014 The Royal College of Midwives. Evidence Based Midwifery 12(4): 112-120

References continued


