

The Desire to Avoid Pregnancy Scale: clinical considerations and comparison with other questions about pregnancy preferences.

Authors:

Jennifer A Hall,¹

Geraldine Barrett,¹

Judith Stephenson,¹

Natalie Edelman,^{2, 3}

Corinne H Rocca⁴

¹Reproductive Health Research Department, UCL Elizabeth Garrett Anderson Institute for Women's Health, London, UK

²School of Sport & Health Sciences, University of Brighton, Brighton & Hove, UK

³Primary Care & Public Health, Brighton & Sussex Medical School, Brighton & Hove, UK

⁴Advancing New Standards in Reproductive Health (ANSIRH), Department of Obstetrics, Gynecology and Reproductive Sciences, University of California, San Francisco (UCSF) School of Medicine, Oakland, CA 94612, USA

Corresponding author: Jennifer Anne Hall, UCL Elizabeth Garrett Anderson Institute for Women's Health, Reproductive Health Research Department Room 237B, 74 Huntley Street, London, WC1E 6AU. Jennifer.hall@ucl.ac.uk

KEYWORDS

Pregnancy planning; pregnancy prevention; sensitivity; specificity; prediction; desire to avoid pregnancy

ABSTRACT

Background

Clinicians and people of reproductive age would benefit from a reliable way to identify who is, or is not, likely to become pregnant in the next year, to direct health advice. The 14-item Desire to Avoid Pregnancy (DAP) Scale is predictive of pregnancy; this paper compares it with other ways of assessing pregnancy preferences to shortlist options for clinical implementation.

Methods

A cohort of 994 UK women of reproductive age completed the DAP and other questions about pregnancy preferences, including the Attitude towards Potential Pregnancy Scale (APPS), at baseline, and reported on pregnancies quarterly for a year. For each question, DAP item, and combinations of DAP items, we examined the predictive ability, sensitivity, specificity, area under the receiver operating curve (AUROC), and positive and negative predictive values.

Results

The AUROCs and predictive ability of the APPS and DAP single items were weaker than the full DAP, though all except one had acceptable AUROCs (>0.7). The most predictive individual DAP item was 'It would be a good thing for me if I became pregnant in the next three months', where women who strongly agreed had a 66.7% chance of pregnancy within 12 months and the AUROC was acceptable (0.77).

Conclusion

We recommend exploring the acceptability to women and healthcare professionals of asking a single DAP item ('I wouldn't mind if I became pregnant in the next three months'), possibly in combination with additional DAP items. This will help to guide the provision of information and services to support reproductive preferences.

Key messages

- Clinicians do not currently have a valid and reliable way of asking people about their pregnancy preferences.
- A single item from the Desire to Avoid Pregnancy Scale is effective at identifying who is likely to become pregnant in the next year; other questions, including the APPS, are less discriminative but may be more acceptable.
- The acceptability to women and health professionals of different ways of asking people about their pregnancy preferences in health care settings should be explored.

INTRODUCTION

While there are a multitude of measures and screening tools available to predict pregnancy-related conditions, such as pre-eclampsia¹ or gestational diabetes,² there are no valid and reliable clinical tools to identify women^a who are likely to become pregnant.¹ Such a tool would have great utility for clinicians working with women of reproductive age as it could be used to guide discussions on either preparation for pregnancy or contraceptive options, depending on if and when a pregnancy might be desired.

The Desire to Avoid Pregnancy (DAP) Scale was developed in the United States (USA) in 2019,³ and validated in the United Kingdom (UK) in 2022, where it was also shown to be highly predictive of pregnancy, with women with the lowest DAP score having an 80% chance of becoming pregnant within 12 months.⁴ Subsequently the sensitivity and specificity of the DAP have been described (0.78 and 0.81 respectively, at a cut-point ≤ 2)⁵ and further analysis has shown how DAP score is associated with socio-demographic factors,⁵ findings which are in keeping with the literature on pregnancy preferences.⁶⁻⁸

While there may be ways of incorporating the DAP scale into clinical practice, e.g., pre-consultation digital use or self-completion, a tool with 14-items is likely to be impractically lengthy in many clinical consultations. Use of selected items from the DAP or other questions, may provide a brief and viable alternative if they distinguish sufficiently between who will and who will not become pregnant. While other, shorter ways of assessing pregnancy preferences exist, such as the Attitude toward Potential Pregnancy Scale (APPS)⁹ and the 'One Key Question[®]' (OKQ),¹⁰ there are no data on their predictive ability and, until the DAP scale, there had not been a 'gold standard' to compare them to.

The aims of this paper are to: 1) evaluate the predictive ability, sensitivity, specificity, area under the receiver operating curve (AUROC) and positive and negative predictive values (PPV and NPV) of different methods of assessing pregnancy preferences; and 2) examine the same criteria for individual DAP items and selected combinations of DAP items. Using the findings of these two aims, options are shortlisted for consideration for clinical implementation.

METHODS

Sample

We analysed data from a cohort of 994 non-pregnant women, aged 15 and over, who were recruited in the UK in October 2018 and followed up every three months for one year. The full details of recruitment and participation are described elsewhere⁴ but, in brief, people who self-reported as female, were pre-menopausal and not sterilised were recruited through advertising at a school, a university, a sexual health clinic and a pregnancy termination clinic, as well as through online recruitment via both paid advertisements (Instagram and Facebook) and sharing

^aWhere we refer to 'women' this should be taken to include people who do not identify as women but who have the capability to become pregnant.

through networks.⁴ The survey was programmed in RedCap^{11,12} and included the DAP scale, other questions about pregnancy preferences and socio-demographics. At each quarterly follow-up, participants were asked whether they were currently pregnant or had been pregnant since the last survey. The women who did not complete 12month follow-up were not significantly different in sociodemographic characteristics from those who did (suggesting that there is no selection bias in the loss to follow up).⁵

Measures

Outcome

We created a binary variable of any incident pregnancy between baseline and 12months. Given low attrition and to ease interpretation, we included participants in pregnancy denominators until they were lost to follow-up and report percentages rather than rates.

Attitude toward Potential Pregnancy Scale

The Attitude toward Potential Pregnancy Scale (APPS) is a five-item measure of a woman's emotional outlook regarding a potential pregnancy (Table 1).⁹ Each item is scored one-to-five on a visual anchored scale and summed to give an overall score from five to 25; higher scores represent a more positive attitude to pregnancy. The APPS had a Cronbach's alpha of 0.86 among 130 participants in the USA but has not been examined in the UK.

Other pregnancy preference questions

We compared the DAP score with three other questions asking about pregnancy preferences (Table 1). The 'feelings question' was used in the ADAPT study;¹³ the 'trying' and 'thinking' questions are ones that clinicians in the UK have told us are the kind of questions they currently use. We did not use the OKQ® 'Would you like to become pregnant in the next year?' as this is a proprietary tool that is unlikely to be available for widespread implementation in the UK, but the 'thinking' question is similar.

DAP Scale

The DAP scale is a psychometrically validated measure covering three conceptual domains: 1) cognitive desires and preferences; 2) affective feelings and attitudes; and 3) anticipated practical consequences (Table 1).³ Each item uses a Likert scale, scored zero-to-four, to ask women how much they agree or disagree with a statement about either becoming pregnant in the next three months or having a baby in the next year. The DAP was developed following an extensive item development process and item response theory was utilised to create the final tool. Responses are averaged, producing a total score between zero and four; higher scores represent a higher desire to avoid pregnancy.

Table 1 Wording of questions about preferences regarding future pregnancy.

Question wording	Response options	Domain
DAP (with label for reference in text)		
I wouldn't mind if I became pregnant in the next 3 months <i>Pregnant: wouldn't mind</i>	0 Strongly agree – 4 strongly disagree	1
It would be a good thing for me if I became pregnant in the next 3 months <i>Pregnant: good thing for me</i>	0 Strongly agree – 4 strongly disagree	1
Thinking about becoming pregnant in the next 3 months makes me feel unhappy. <i>Pregnant: unhappy</i>	4 Strongly agree – 0 strongly disagree	2
Thinking about becoming pregnant in the next 3 months makes me feel excited. <i>Pregnant: excited</i>	0 Strongly agree – 4 strongly disagree	2
Becoming pregnant in the next 3 months would bring me closer to my main partner <i>Pregnant: closer to partner</i>	0 Strongly agree – 4 strongly disagree	3
I want to have a baby within the next year. <i>Baby: want</i>	0 Strongly agree – 4 strongly disagree	1
If I had a baby in the next year, it would be bad for my life. <i>Baby: bad for life</i>	4 Strongly agree – 0 strongly disagree	1
It would be a positive addition to my life to have a baby in the next year. <i>Baby: positive addition to life</i>	0 Strongly agree – 4 strongly disagree	1
It would be the end of the world for me to have a baby in the next year <i>Baby: end of the world for me</i>	4 Strongly agree – 0 strongly disagree	1
Thinking about having a baby within the next year makes me smile <i>Baby: makes me smile</i>	0 Strongly agree – 4 strongly disagree	2
Thinking about having a baby within the next year make me feel stressed out. <i>Baby: stressed out</i>	4 Strongly agree – 0 strongly disagree	2
I would feel a loss of freedom if I had a baby in the next year. <i>Baby: loss of freedom</i>	4 Strongly agree – 0 strongly disagree	3
If I had a baby in the next year, it would be hard for me to manage raising the child. <i>Baby: hard for me to manage</i>	4 Strongly agree – 0 strongly disagree	3
I would worry that having a baby in the next year would make it harder for me to achieve other things in my life. <i>Baby: hard to achieve other things</i>	4 Strongly agree – 0 strongly disagree	3
APPS		
How much do you want to be pregnant now?	1 Not at all – 5 Very much	
How important is it to you to avoid become pregnant now?	1 Not at all – 5 Very much	
How worried would you be if you were pregnant now?	1 Not at all – 5 Very much	
How upset would you be if you were pregnant now?	1 Not at all – 5 Very much	
How happy would you be if you were pregnant now?	1 Not at all – 5 Very much	
Feeling question		
Overall, when I think about become pregnant in the next 3 months, I feel:	1 Mostly positive 2 More positive than negative 3 In the middle between positive and negative 4 I don't feel strongly one way or the other 5 More negative than positive 6 Most negative	Not applicable
Trying question		
Are you currently trying to get pregnant?	1 Yes 0 No	
Thinking question		
Are you thinking about trying to get pregnant in the next year?	2 Yes 1 Maybe 0 No	

Domains: 1 = Cognitive desires and preferences, 2 = Affective feelings and attitudes, 3 = Anticipated practical consequences.

Further information on the DAP wording and instructions for use are available here:

https://www.ansirh.org/sites/default/files/publications/files/desire_to_avoid_pregnancy_scale_english_6_22_20.pdf

Plan of Analysis

Aim 1: Evaluating the performance of other methods of assessing pregnancy preferences

We assessed the APPS's reliability (internal consistency) with Cronbach's alpha (>0.7 considered acceptable)¹⁴ and checked that all item-rest correlations were positive and >0.2 .¹⁵ We examined construct (structural) validity via principal components analysis and considered the scale structurally valid if all items loaded on to one component with an Eigenvalue >1 .¹⁶

To compare the assessment of pregnancy preferences by the DAP versus the APPS and the three individual questions ('feeling', 'trying' and 'thinking'), we examined the relationships among them. As the DAP and the APPS yield continuous scores, we used the Pearson correlation coefficient to examine the strength of association, expecting a negative correlation (i.e., as DAP score increases, APPS score decreases), considering a strong correlation to be lower than -0.7 . This also served as a test of the APPS's concurrent validity. The range of DAP scores within, and distribution of DAP scores across, the response options of each question was examined. The Kruskal-Wallis test was used to assess whether differences in median DAP score across response options within each question were significant.

To investigate how well each approach predicted actual pregnancy, we modelled the probability of pregnancy using logistic regression models and examined the sensitivity, specificity, AUROC, PPV and NPV, using the Youden index suggested cut-point.¹⁷ The AUROC represents the tool's ability to identify who will become pregnant. An AUROC of $0.7-0.8$ was considered acceptable, $0.8-0.9$ excellent and >0.9 outstanding.¹⁸

Aim 2: Examination of the performance of individual and combinations of DAP items

We explored the predictive ability of each individual DAP item, as well as selected combinations of items, using logistic regression, with incident pregnancy as the outcome. The item combinations were designed to ensure coverage of the three conceptual domains, positively and negatively framed items, different time frames and use of 'pregnancy' and 'baby', as well as the predictive ability of the items. The combinations were cross-validated by developing them on one half of the data (created with a random split) and tested on the other half to avoid over-fitting and give a more accurate estimate of how the questions perform outside the data they were developed from.^{19 20} The results from the testing data are presented. The predictive ability, sensitivity, specificity, AUROC, PPV and NPV of the selected items/combinations were calculated (using the Youden index suggested cut-point)¹⁷ and compared with the total DAP scale.

To inform which question(s) provided the best balance between brevity and predictive ability, the performance of the single questions, the APPS, and selected DAP items/combinations were compared using the predictive ability, AUROC and number of items/questions, to make

recommendations on which item/question(s) should be taken forward for further consideration.

Patient and Public Involvement

We undertook public involvement in the development of our overall programme of research on pregnancy planning; the P3 Study. Findings were frequently discussed with the P3 Study's PPI group to inform the next steps and the group is currently planning our wider dissemination.

RESULTS

Sample

As previously described,⁴ the baseline cohort of 994 women were aged 15-50 years (median=31, IQR=23-36, mean=29.7). Most (82%) were in a relationship, 82% described themselves as heterosexual and 84% identified as white. The sample was quite highly educated with 39% having an undergraduate degree and 31% a postgraduate or professional qualification. 57% had one or more children in their household. The dataset is available in the UCL Research Data Repository.²¹

Aim 1: Evaluating the performance of other methods of assessing pregnancy preferences

The full range of APPS scores (5-25) was reported in our cohort. The Cronbach's alpha for the APPS was 0.93, all item-rest correlations were >0.2 and positive, and all items loaded on to one component with an Eigenvalue of 3.87. The Pearson correlation coefficient between the APPS and DAP was -0.893 (p<0.001) showing a strong negative correlation (i.e., as a woman's attitude towards pregnancy becomes more positive her desire to avoid pregnancy reduces.) The median DAP score was statistically significantly different across the response options for the APPS and each of the 'feeling', 'trying' and 'thinking' questions (p-values in Table 2). For example, women who felt 'mostly positive' about pregnancy had a median DAP score of 0.64 (low desire to avoid pregnancy) whereas women who felt 'mostly negative' about pregnancy had a median DAP score of 3.50 (high desire to avoid pregnancy).

Table 2 Relationship with DAP and predicted probability, sensitivity, specificity, AuROC, PPV and NPV of other questions.

	n	Median DAP score	Inter-quartile range	p value	Predicted probability of pregnancy	95%CI	Sensitivity	Specificity	AuROC	PPV	NPV
APPS score[^]											
5	265	3.64	3.29 - 3.86	<0.001	0.03	0.019 - 0.044	78%	74%	0.76	36%	95%
10	59	2.50	2.29 - 2.71		0.09	0.063 - 0.107					
15	35	1.71	1.50 - 2.07		0.21	0.173 - 0.243					
20	13	1.07	0.86 - 1.50		0.47	0.362 - 0.490					
25	42	0.57	0.36 - 0.64		0.68	0.594 - 0.761					

Overall, when I think about becoming pregnant in the next 3 months, I feel:										
Mostly positive	117	0.64	0.43 - 0.93	<0.001	0.56	0.47 - 0.65				
More positive than negative	84	1.29	1.07 - 1.64		0.37	0.26 - 0.48				
In the middle	107	2.00	1.57 - 2.36		0.14	0.07 - 0.21				
Don't feel strongly	21	2.14	2.00 - 2.43		0.07	-0.06 - 0.19	75%	76%	0.75	37% 94%
More negative than positive	264	2.64	2.36 - 3.00		0.09	0.06 - 0.13				
Mostly negative	395	3.50	3.14 - 3.79		0.03	0.02 - 0.05				
Are you currently trying to get pregnant?										
Yes	907	0.64	0.36 - 0.86	<0.001	0.69	0.58 - 0.50	37%	97%	0.67	69% 89%
No	80	2.79	2.07 - 3.43		0.11	0.09 - 0.13				
Are you thinking about trying to get pregnant in the next year?										
Yes	707	1.21	0.79 - 1.71	<0.001	0.53	0.42 - 0.63				
Maybe	97	1.86	1.36 - 2.21		0.15	0.07 - 0.22	71%	84%	0.77	34% 96%
No	96	3.07	2.57 - 3.57		0.04	0.20 - 0.06				

^ APPS score is continuous from 5-25, 5 points shown for illustration therefore n does not add up to the full sample

The AUROCs and predictive ability of the APPS and the single questions were weaker than the full DAP (AUROC 0.87, predictive ability 79.4%), though all except the 'trying' question had acceptable AUROCs of >0.7.

Aim 2: Evaluating the performance of individual and combinations of DAP items

All DAP items and domains were associated with pregnancy. The best performing individual DAP item in terms of pregnancy prediction was 'Pregnancy: good thing for me', where there was a 66.7% chance of pregnancy within 12months among women who strongly agreed (Table 3). At a cut point of 2.5 this item also had an acceptable AUROC (0.77).

Adding a second item of 'Baby: end of the world for me' improved the specificity and PPV, without affecting the AUROC. The highest AUROC (0.70) was achieved with the combination of three items 'Pregnancy: good thing for me', 'Baby: want' and 'Baby: end of the world for me', which also had the highest PPV (57.1%). The item combinations had AUROCs between 0.77-0.79, in line with that of the individual item 'Pregnancy: good thing for me' suggesting little additional gain.

Table 3 Exploration of the predicted probability, sensitivity, specificity, AuROC, PPV and NPV of selected individual and combined DAP items.

Description	Predicted probability of pregnancy at DAP score	Empirical cut-point	Sensitivity	Specificity	AuROC	PPV	NPV	Domains covered	Positive & Negatively worded items
-------------	---	---------------------	-------------	-------------	-------	-----	-----	-----------------	------------------------------------

	0	4								
Total DAP: <i>All 14 items</i>	79.4%	0.9%	1.96	78.0%	81.0%	0.85	43.0%	95.0%	All	Yes
Single DAP item: <i>Pregnant: good thing for me.</i>	66.7%	2.1%	2.5	68.1%	86.3%	0.77	49.0%	93.3%	1	Positive only
Single DAP item: <i>Baby: end of the world for me.</i>	27.4%	–	0.5	65.3%	77.7%	0.72	36.2%	92.0%	1	Negative only
Two DAP items: <i>Pregnant: good thing for me & Baby: end of the world for me.</i>	71%	0.7%	1.25	65.3%	90.1%	0.78	56.0%	93.1%	1	Yes
Three DAP items: <i>Pregnant: good thing for me, Pregnant: excited, Baby: hard to achieve other things.</i>	83%	2.8%	2.15	62.5%	87.1%	0.75	48.4%	92.3%	All	Yes
Three DAP items: <i>Pregnant: good thing for me, Pregnant: excited, Baby: end of the world for me.</i>	73%	1.2%	1.49	66.7%	87.4%	0.77	50.5%	93.1%	1,2	Yes
Three DAP items: <i>Pregnant: good thing for me, Baby: end of the world for me, Baby: makes me smile.</i>	68%	1.0%	1.49	69.4%	84.4%	0.77	46.3%	93.5%	1,2	Yes
Three DAP items: <i>Pregnant: good thing for me, Baby: want, Baby: end of the world for me.</i>	72%	13.9%	1.49	66.7%	90.3%	0.79	57.1%	93.3%	1	Yes
Three DAP items: <i>Pregnant: good thing for me, Baby: end of the world for me, Baby: hard for me to manage.</i>	76%	8.8%	1.82	75.0%	77.2%	0.76	38.8%	94.1%	1,3	Yes
Four DAP items: <i>Pregnant: good thing for me, Pregnant: excited, Baby: want, Baby: end of the world for me.</i>	73%	1.4%	1.63	66.7%	89.2%	0.78	54.5%	93.2%	1,2	Yes
Four DAP items: <i>Pregnant: good thing for me, Pregnant: excited, Baby: end of the world for me, Baby: hard to achieve other things.</i>	81%	1.2%	1.88	66.7%	83.9%	0.75	44.4%	92.9%	All	Yes
Six DAP items: <i>Pregnant: wouldn't mind, Pregnant: unhappy, Baby: want, Baby: positive addition to life, Baby: makes me smile, Baby: hard for me to manage.</i>	73%	1.6%	1.93	75.0%	80.9%	0.78	43.2%	94.4%	All	Yes

DISCUSSION

This is the first examination of the predictive ability of two measures and three single questions about pregnancy preferences on a sample that is broadly representative of women of reproductive age in the UK. Good correlation was previously seen between DAP score and ‘One

Key Question® (OKQ) in the USA.²² As that study noted, DAP scores ranged widely within each OKQ response option; women who responded to the OKQ with ‘want to get pregnant in future but not in next year’ had DAP scores ranging from 1–4. Our data demonstrated the same patterns; women who reported feeling ‘mostly negative’ towards pregnancy had DAP scores ranging from 0.286–4, this shows the nuance that the DAP can capture and demonstrates the heterogeneity missed by a single question.

In terms of the AUROC and the predicted probability of pregnancy in the next year, the performance of the APPS and the other questions was weaker than the complete DAP, though to varying degrees. The APPS and the other questions are all shorter than the DAP and therefore potentially less burdensome for clinical use. Arguably, their poorer performance is offset by their brevity. Our data show that the single question ‘Are you currently trying to get pregnant?’ had low sensitivity (37%) and the lowest AUROC (0.67). While women answering yes to this question were highly likely to become pregnant (PPV 69%) there were more pregnancies among women who answered no. Questions like this, which force women to answer yes-or-no, fail to recognise the complexity of the concept of pregnancy preferences.²³

When considering which question(s) might be best for clinical use, the trade-off between sensitivity and specificity, and the PPV and NPV is important. In addition, the number of questions that need to be asked, and the complexity of combining them to achieve a score, could be a barrier in consultations, therefore having fewer questions will likely increase uptake. As our analysis shows, single questions generally have lower AUROCs than a set of questions, however selected items from the DAP, which have been developed based on rigorous theoretical groundwork, do have higher AUROCs than the less carefully constructed questions. Neither the DAP nor the APPS were designed to be spoken, whereas the other questions lend themselves more easily to being asked verbally.

Strengths and limitations

We used a large, broadly representative dataset⁴ and a validated measure (the DAP) to assess the performance of a range of questions for assessing pregnancy preferences and have provided preliminary evidence that the APPS is valid in the UK. Given the limitations on the OKQ we were not able to include it in our research. While we used a split dataset to conduct cross-validation within our sample, which is a strength, confirmation of our findings in other populations would provide further confidence in selecting the most appropriate question(s). Finally, the selected DAP items/combinations were based on their potential as screening tools for pregnancy, as well as theoretical considerations; item-response theory-based analysis of the psychometric properties of these combinations could be conducted.

CONCLUSION

Discussions about pregnancy preferences are important, regardless of whether the patient wants to become pregnant in the future or not. Ensuring that those who do not wish to become pregnant have the right support to avoid pregnancy is just as important as identifying those

who might benefit from pre-pregnancy health advice or those who may benefit from both. Equally, for those who have never formally considered their preferences, it provides an opportunity to empower them and increase their agency by highlighting that people do have choices about pregnancy and parenthood (recognising the effect of external factors) and encouraging them to explore their aspirations.

In the context of a face-to-face clinical encounter, where the full DAP is less likely to be suitable, a tool to assess people's preferences regarding a future pregnancy needs to be both practical (short) and discriminative i.e. identify who is and is not likely to become pregnant in the short term so that the appropriate advice can be given. Based on our findings we recommend exploring the acceptability to women and health care professionals of a single item from the DAP ('I wouldn't mind if I became pregnant in the next three months') or a combination of this with two additional DAP items (either 'Baby: want' and 'Baby: end of the world for me' or 'Pregnant: excited' and 'Baby: hard to achieve other things') adapted from the written format to a spoken one.

Acknowledgements

We would like to thank all the women who took part in the P3 study, the P3 Study steering committee and Catherine Stewart for her support in preparing the manuscript for submission.

Contributors

JH: guarantor, conceptualisation, funding acquisition, methodology, investigation, data curation, formal analysis, writing-original draft, visualisation, writing-review and editing, project administration. GB: conceptualisation, methodology, writing-review and editing. JS: conceptualisation, writing-review and editing CR: conceptualisation, methodology, writing-review and editing. NE: methodology, writing-review and editing.

Funding

The study was funded by an NIHR Advanced Fellowship held by JH (PDF-2017-10-021). The funder had no role in the study design; in the collection, analysis and interpretation of the data; in the writing of the report; or in the decision to submit the paper for publication.

Competing interests

The authors declare that they have no conflicts of interest.

Data availability statement

The dataset is available in the UCL Research Data Repository, doi 10.5522/04/21263007.

REFERENCES

1. Tan MY, Wright D, Syngelaki A, et al. Comparison of diagnostic accuracy of early screening for pre-eclampsia by NICE guidelines and a method combining maternal factors and biomarkers: results of SPREE. *Ultrasound Obstet Gynecol* 2018;51(6):743-50. doi: 10.1002/uog.19039 [published Online First: 2018/03/15]
2. Artzi NS, Shilo S, Hadar E, et al. Prediction of gestational diabetes based on nationwide electronic health records. *Nature Medicine* 2020;26(1):71-76. doi: 10.1038/s41591-019-0724-8
3. Rocca CH, Ralph LJ, Wilson M, et al. Psychometric Evaluation of an Instrument to Measure Prospective Pregnancy Preferences: The Desire to Avoid Pregnancy Scale. *Med Care* 2019;57(2):152-58. doi: 10.1097/MLR.0000000000001048 [published Online First: 2018/12/15]
4. Hall J, Barrett G, Rocca C. Evaluation of the Desire to Avoid Pregnancy Scale in the UK: a psychometric analysis including predictive validity. *BMJ Open* 2022;12(7):e060287. doi: 10.1136/bmjopen-2021-060287 [published Online First: 2022/07/26]
5. Hall J, Barrett G, Stephenson J, et al. Predictive Ability of the Desire to Avoid Pregnancy Scale. *medRxiv* 2022:2022.10.17.22281028. doi: 10.1101/2022.10.17.22281028
6. Hill B, Ling M, Mishra G, et al. Lifestyle and Psychological Factors Associated with Pregnancy Intentions: Findings from a Longitudinal Cohort Study of Australian Women. *Int J Environ Res Public Health* 2019;16(24):5094. doi: 10.3390/ijerph16245094
7. Lang AY, Harrison CL, Barrett G, et al. Opportunities for enhancing pregnancy planning and preconception health behaviours of Australian women. *Women Birth* 2021;34(2):e153-e61. doi: 10.1016/j.wombi.2020.02.022 [published Online First: 2020/04/22]
8. Geist C, Everett BG, Simmons RG, et al. Changing lives, dynamic plans: Prospective assessment of 12-month changes in pregnancy timing intentions and personal circumstances using data from HER Salt Lake. *PLoS One* 2021;16(9):e0257411. doi: 10.1371/journal.pone.0257411 [published Online First: 2021/09/21]
9. Paterno MT, Han H-R. Development and psychometric testing of the attitude toward potential pregnancy scale. *J Obstet Gynecol Neonatal Nurs* 2014;43(6):710-18. doi: 10.1111/1552-6909.12511 [published Online First: 2014/10/14]
10. Allen D, Hunter MS, Wood S, et al. One Key Question((R)): First Things First in Reproductive Health. *Matern Child Health J* 2017;21(3):387-92. doi: 10.1007/s10995-017-2283-2 [published Online First: 2017/02/22]
11. Harris PA, Taylor R, Minor BL, et al. The REDCap consortium: Building an international community of software platform partners. *J Biomed Inform* 2019;95:103208. doi: 10.1016/j.jbi.2019.103208 [published Online First: 2019/05/13]
12. Harris PA, Taylor R, Thielke R, et al. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform* 2009;42(2):377-81. doi: 10.1016/j.jbi.2008.08.010 [published Online First: 2008/10/22]
13. Rocca CH, Parra M, Munoz I, et al. Comparison of Pregnancy Preferences Preceding vs Year 1 of the COVID-19 Pandemic. *JAMA Netw Open* 2022;5(7):e2220093. doi: 10.1001/jamanetworkopen.2022.20093 [published Online First: 20220701]

14. Cronbach LJ. Essentials of psychologic testing. 5th ed. New York: Harper & Row 1990.
15. Health Measurement Scales: a practical guide to their development and use [program]. Oxford: Oxford University Press, 2008.
16. Kline P. The new psychometrics: science, psychology and measurement. London: Routledge 1998.
17. Youden WJ. Index for rating diagnostic tests. *Cancer* 1950;3(1):32-35. doi: [https://doi.org/10.1002/1097-0142\(1950\)3:1<32::AID-CNCR2820030106>3.0.CO;2-3](https://doi.org/10.1002/1097-0142(1950)3:1<32::AID-CNCR2820030106>3.0.CO;2-3)
18. Hosmer DW, Lemeshow S. Applied Logistic Regression. 2nd ed. New York, NY: John Wiley and Sons 2000:160-64.
19. Koul A, Becchio C, Cavallo A. Cross-Validation Approaches for Replicability in Psychology. *Front Psychol* 2018;9:1117. doi: 10.3389/fpsyg.2018.01117 [published Online First: 2018/07/24]
20. Yarkoni T, Westfall J. Choosing Prediction Over Explanation in Psychology: Lessons From Machine Learning. *Perspect Psychol Sci* 2017;12(6):1100-22. doi: 10.1177/1745691617693393 [published Online First: 2017/08/26]
21. Hall J. Data for DAP predictive ability and comparison. UCL Research Data Repository, 2022.
22. Stulberg DB, Datta A, White VanGompel E, et al. One Key Question(R) and the Desire to Avoid Pregnancy Scale: A comparison of two approaches to asking about pregnancy preferences. *Contraception* 2020;101(4):231-36. doi: 10.1016/j.contraception.2019.12.010 [published Online First: 2020/01/15]
23. Barrett G, Wellings K. What is a 'planned' pregnancy? Empirical data from a British study. *Social Science and Medicine* 2002;55:545-57.