Published in final edited form as:

Epidemiology. 2009 January ; 20(1): 56–59. doi:10.1097/EDE.0b013e31818ef47e.

Validity of Self-Reported Time to Pregnancy

Maureen A. Cooney^a, Germaine M. Buck Louis^a, Rajeshwari Sundaram^a, Bridget M. McGuinness^b, and Courtney D. Lynch^a

^aDivision of Epidemiology, Statistics and Prevention Research, Eunice Kennedy Shriver National Institute of Child Health and Human Development, Rockville, MD

^bDepartment Social and Preventive Medicine, University at Buffalo, Buffalo, NY

Abstract

Background—The reliability of retrospective time to pregnancy (TTP) has been established, but its validity has been assessed in only 1 study, which had a short follow-up.

Methods—Ninety-nine women enrolled a decade earlier in a prospective TTP study were queried by means of mailed questionnaires about the duration of time they had required to become pregnant. Their responses were compared with their earlier data from daily diaries (gold standard).

Results—One-third of women could not recall their earlier TTP either in menstrual cycles or calendar months. Only 17%-19% of women recalled their TTP exactly. Agreement increased to 41%-51%, 65%-72%, and 72%-77% when defined as ± 1 , ± 2 , and ± 3 months, respectively. Women with longer observed TTPs or previous pregnancies were more likely to under-report their TTP.

Conclusions—The findings raise questions about the commonly assumed validity of self-reported TTP. Recalled TTP may introduce error when estimating fecundability or classifying couples' fecundity status.

Time to pregnancy (TTP) is a functional measure of human fecundity representing successful completion of several conditional processes underlying human conception and implantation. TTP estimates fecundability, historically defined as the probability of conception among women exposed to unprotected sexual intercourse in the absence of lactational anovulation, pregnancy, or sterility.¹ Time may be defined as the number of calendar months or menstrual cycles, and data can be collected retrospectively or prospectively.

TTP can be used to assess behavioral or environmental factors. Some factors associated with a longer TTP include a body mass index 30.0 kg/m or more, menstrual cycles longer than 35 days, menstrual bleeding for 4 or fewer days, smoking more than 20 cigarettes per day, caffeine consumption of more than 300 mg/d among smokers, and blood mercury concentrations of greater than 1.2 $\mu\text{g/L}$. The range of effect size varies (18%-78%)

Copyright © 2008 by Lippincott Williams & Wilkins

Correspondence: Germaine M. Buck Louis, Epidemiology Branch, Division of Epidemiology, Statistics and Prevention Research, Eunice Kennedy Shriver National Institute of Child Health and Human Development, 6100 Executive Boulevard, Rm. 7B03, Rockville, MD 20852. louisg@mail.nih.gov..

underscoring the importance of accurate TTP data.²⁻⁷ TTP has also been reported to be associated with adverse perinatal outcomes^{8,9} and, more recently, speculated to be in the pathway of urologic or gynecologic disorders.^{10,11}

Although several reliability studies for TTP have been published,¹²⁻¹⁵ to our knowledge, only 1 validation study has been reported. Zielhuis et al¹⁶ had prospective data on TTP for 100 women, and then collected self-reported TTP 3-20 months later by using 1 of 3 questionnaire approaches: in-person, telephone, or e-mail. Mean differences in TTP were 0.1, 2.5, and 0.6 months for the 3 approaches, respectively, which the authors interpreted as supporting validity. Joffe et al¹⁵ assessed TTP at 2 points with about a 14 plus-year interval among women giving birth in what is frequently cited as a validation study. However, there was no gold standard (prospective measurement); instead, the authors constructed an initial TTP derived from women's answers to questions on changes in contraceptive use, pregnancy, births, or miscarriages in the preceding 12 months.

METHODS

Study Design and Data Collection

Prospective TTP—The referent study population comprised women members of the New York State Angler Cohort Study, established in 1991 to characterize species-specific fish consumption among reproductive-age couples.¹⁷ Women who stated that they were considering future pregnancies were recontacted in 1996 and asked to take part in a prospective TTP study; 113 of 244 eligible women agreed to participate. Fourteen women became pregnant before enrollment, leaving 99 women for study. Women were interviewed upon discontinuing contraception and trained in the accurate use of home pregnancy test kits and the completion of daily diaries (ie, sexual intercourse, menstruation, pregnancy test results, and related behaviors). Women were followed until they had a positive home pregnancy test result on the expected day of menses or up to 12 menstrual cycles with sexual intercourse in the estimated fertile window. We relied on the Ogino-Knaus method for estimating the likely date of ovulation, counting back 14 days from the end of the cycle.^{18,19}

Retrospective TTP—In 2006, letters on the original study letterhead were sent to the 99 women informing them of their laboratory values and asking them to complete a short questionnaire focusing on how well women remember what they were doing while trying to become pregnant. To stimulate recall for the correct pregnancy, we included the infant's birth date and baby photo. Retrospective TTP was defined on the basis of women's responses to the following question, which was first asked in calendar months and then in menstrual cycles: "Do you remember how many [calendar months/menstrual cycles] it took you to become pregnant?" If the woman said yes, she was asked to give the number. Human subjects' approval was obtained in accordance with the Declaration of Helsinki of the World Medical Association.

Statistical Analysis—Maternal characteristics at baseline were compared by participation status in 2006 and the women's ability to recall TTP, using the Fisher exact test or the *t* test. Prospective TTP was defined separately as the number of (1) (rounded) calendar months,

and (2) menstrual cycles. Error in TTP recall was calculated by subtracting the prospective from the retrospective TTP, yielding the amount of error (in months or cycles) for each woman with both measures. In this case, the error is a measure of statistical bias of recall.

The Jonckheere-Terpstra test was used to assess trends between the observed (prospective) TTP and the absolute error in recalled TTP.²⁰ We tested the null hypothesis of no trend against the alternative hypothesis of an increase in frequency of absolute error with an increase in observed TTP for each unit of TTP time; right-tailed *P* values were calculated to determine the direction of error in reporting.

Using multiple linear regression, we modeled the association between error of recall for TTP (in months and cycles) in relation to presumed fecundity determinants for women with both TTP measures. Covariates included age (<30 or ≥30 years); gravidity (nulligravida or gravid); education; observed (prospective) TTP; use of alcohol, multivitamins, and cigarettes while attempting pregnancy; and sport fish consumption. Variables were retained if they were a priori identified as confounders and were associated with error in recall at the 5% level. Both normality and multiple linear regression assumptions were upheld.

RESULTS

The sample comprised white mostly college educated (96%) gravid (81%) women. Eighty-nine women (90%) returned questionnaires: 46 women did not report both TTP measures, that is, there were 27 missing prospective and 19 missing retrospective TTPs. Reasons for missing prospective TTPs included: infertility (*n* = 9); withdrawals (*n* = 16), and pregnant at enrollment (*n* = 2). Women missing retrospective TTPs simply could not remember.

Sixty-seven percent of women reported that they could recall TTP in months and 59% in cycles. No differences were observed for maternal characteristics or pregnancy outcomes in relation to participation or ability to recall TTP (data not shown).

Figure 1 reflects the distribution of prospective against retrospective TTP, whereas Figure 2 reflects the distribution of error across the continuum of prospective TTP. Women over- and under-reported TTP both in months and cycles. Nineteen percent of women reported TTP accurately; this increased to 51%, 72%, and 77% when agreement was expanded by ± 1 , ± 2 , and ± 3 months, respectively. Given the small cohort size, TTP was categorized (1, 2, 3-5, 6-8, >8 months/cycles) to avoid digit preferences of 6 and 9 months in retrospectively reported TTP (Table 1). The size of the error in retrospective recall of TTP tended to increase with increases in the true (prospective) TTP.

Gravidity and a longer prospective TTP were associated with error in recall of TTP regardless of the unit of measurement for TTP (Table 2). Specifically, the error in recall of TTP for gravid women is an estimated -5.4 months relative to nulligravid women and the error in recall of TTP changes by -0.9 for every month increase in prospective TTP. Findings were similar when TTP was expressed in cycles.

DISCUSSION

This validation study found TTP to be poorly recalled by women about 10 years later. TTP had originally been measured in an intensive prospective study with daily diaries, while attempting to become pregnant. Among the 67% of women who said they could recall TTP, only 17%-19% had exact agreement in either calendar months or menstrual cycles. Accuracy increased to 77% when agreement was allowed within ± 3 months.

Errors in retrospectively reported TTP occurred in both directions. Furthermore, the magnitude of error was comparable to the effect size reported for some exposures or behaviors suggesting that errors may mask or spuriously identify fecundity determinants. Reporting errors may misclassify women or couples with regard to fecundity status and may contribute to the estimated 20% of “subfecund” couples reported to conceive spontaneously in the year after an infertility evaluation (where subfecundity is defined by recalled attempt time).²¹ The extent to which such misclassification would affect research findings focusing on life course issues remains to be established.

Although provocative, our findings need to be interpreted in the context of a cohort of limited size. Also, we cannot rule out the possibility that women provided information for a pregnancy other than the one captured in the original study. We believe this is unlikely, given the prompts used to stimulate recall. Another possible explanation is that women were hesitant to answer, knowing we had the “answer.” This may have contributed to the high percentage of women who reported not being able to recall. Some women may already have been trying to become pregnant before enrollment, but this premise is inconsistent with the observed bidirectional error.

Our findings coupled with those from the earlier validity study suggest that retrospective TTP may have short-term but not long-term validity. Given the increasing use of TTP as a measure of fecundity and as a predictor of adverse perinatal outcomes and adult-onset disease, other researchers with prospective pregnancy data are encouraged to conduct formal validation studies of retrospective TTP to delineate further the utility of this measure.

Acknowledgments

We thank Rafael Mikolajczyk and Enrique Schisterman for providing comments on earlier versions of the paper.

Supported by the Great Lakes Protection Fund (RM791-3021), the Agency for Toxic Substances and Disease Registry (H751 ATH 298338), and the Intramural Research Program of the Eunice Kennedy Shriver National Institute of Child Health and Human Development.

REFERENCES

1. Gini MC. Premieres recherches sur la fecondabilite de la femme. *Proc Int Math Congr.* 1924; 2:889–892.
2. Gesink L, Maclehorse RF, Longnecker MP. Obesity and time to pregnancy. *Human Reprod.* 2007; 22:414–420.
3. Jensen TK, Scheike T, Keiding N, et al. Fecundability in relation to body mass and menstrual cycle patterns. *Epidemiology.* 1999; 10:422–428. [PubMed: 10401878]
4. Small CM, Manatunga AK, Klein M, et al. Menstrual cycle characteristics: associations with fertility and spontaneous abortion. *Epidemiology.* 2006; 17:52–60. [PubMed: 16357595]

5. Curtis KM, Savitz DA, Arbuckle TE. Effects of cigarette smoking, caffeine consumption, and alcohol intake on fecundability. *Am J Epidemiol.* 1997; 146:32–41. [PubMed: 9215221]
6. Jensen TK, Henriksen TB, Hjollund NH, et al. Caffeine intake and fecundability: a follow-up study among 430 Danish couples planning their first pregnancy. *Reprod Toxicol.* 1998; 12:289–295. [PubMed: 9628552]
7. Cole DC, Wainman B, Sanin LH, et al. Environmental contaminant levels and fecundability among non-smoking couples. *Reprod Toxicol.* 2006; 22:13–19. [PubMed: 16439098]
8. Axmon A, Hagmar L. Time to pregnancy and pregnancy outcome. *Fertil Steril.* 2005; 84:966–974. [PubMed: 16213851]
9. Williams MA, Goldman MB, Mittendorf R, et al. Subfertility and the risk of low birth weight. *Fertil Steril.* 1991; 56:668–671. [PubMed: 1915940]
10. Skakkebaek NE, Rajpert-De Meyts E, Main KM. Testicular dysgenesis syndrome: an increasingly common developmental disorder with environmental aspects. *Human Reprod.* 2001; 16:972–978.
11. Buck Louis, GM.; Cooney, MA. Effects of environmental contaminants on ovarian function and fertility. In: Gonzalez-Bulnes, A., editor. *Novel Concepts in Ovarian Endocrinology*. Research Signpost; Kerala, India: 2007. p. 249–268.
12. Baird DD, Weinberg CR, Rowland AS. Reporting errors in time-to-pregnancy data collected with a short questionnaire. Impact on power and estimation of fecundability ratios. *Am J Epidemiol.* 1991; 133:1282–1290. [PubMed: 2063836]
13. Joffe M. Feasibility of studying subfertility using retrospective self reports. *J Epidemiol Community Health.* 1989; 43:268–274. [PubMed: 2607307]
14. Joffe M, Villard L, Li Z, et al. Long-term recall of time-to-pregnancy. *Fertil Steril.* 1993; 60:99–104. [PubMed: 8513966]
15. Joffe M, Villard L, Li Z, et al. A time to pregnancy questionnaire designed for long term recall: validity in Oxford, England. *J Epidemiol Community Health.* 1995; 49:314–319. [PubMed: 7629471]
16. Zielhuis GA, Hulscher ME, Florack EI. Validity and reliability of a questionnaire on fecundability. *Int J Epidemiol.* 1992; 21:1151–1156. [PubMed: 1483821]
17. Vena JE, Buck GM, Kostyniak P, et al. The New York Angler Cohort Study: exposure characterization and reproductive and developmental health. *Toxicol Ind Health.* 1996; 12:327–334. [PubMed: 8843550]
18. Ogino K. Ovulationstermin und Konzeptionstermin. *Zentralbl F Gynak.* 1930; 54:464–479.
19. Knaus H. Eine neue Methods zur Bestimmung des Ovulationstermines. *Zentralbl F Gynak.* 1929; 53:2193.
20. Hollander, M.; Wolfe, DA. *Nonparametric Statistical Methods*. Wiley; New York: 1999.
21. van der Steeg JW, Steures P, Eijkemans MJ, et al. Pregnancy is predictable: a large-scale prospective external validation of the prediction of spontaneous pregnancy in subfertile couples. *Human Reprod.* 2007; 22:536–542.

FIGURE 1.

Scatterplot of prospective compared with retrospective time to pregnancy. Circle size reflects the frequency of reporting ranging from 1 (smallest circle) to 5 (largest circle). An extreme retrospective value of 32 months (with a prospective value of 3 months) and a prospective value of 27 months (with a retrospective value of 4 months) were removed to better capture the pattern of responses.

FIGURE 2.

Distribution of the error in the retrospective report of time to pregnancy compared with the prospectively recorded measurement (gold standard) by unit of time. An extreme retrospective value of 32 months (with a prospective value of 3 months) and a prospective value of 27 months (with a retrospective value of 4 months) were removed to better capture the pattern of responses.

TABLE 1

Percent Distribution of Margin of Error in Retrospective Report of TTP by Prospective Measure and Unit of Time

Error (in Calendar Months, n = 43)				
Prospective TTP	0	±1	±2	>2
1	36	29	21	14
2	11	67	11	11
3-5	22	11	44	22
6-8	0	17	17	67
>8	0	40	0	60
Trend $P < 0.01^a$				
Error (in Menstrual Cycles, n = 38)				
	0	±1	±2	>2
1	35	20	25	20
2	17	33	17	33
3-5	25	0	25	50
6-8	20	20	0	60
>8	0	0	33	67
Trend $P = 0.02^a$				

^aBased on the Jonckheere-Terpstra test.

TABLE 2

Determinants of Error in Retrospective Report of Time to Pregnancy by Unit of Time

	Months (95% CI)
Calendar mo	
TTP (per mo)	-0.9 (-1.2 to -0.6)
Any previous pregnancy	-5.4 (-9.4 to -1.4)
Menstrual cycles	
TTP (per cycle)	-0.9 (-1.4 to -0.4)
Any previous pregnancy	-5.6 (-9.6 to -1.6)