



Published in final edited form as:

*Paediatr Perinat Epidemiol.* 2015 July ; 29(4): 360–371. doi:10.1111/ppe.12201.

## Design and conduct of an internet-based preconception cohort study in North America: Pregnancy Study Online (PRESTO)

Lauren A Wise<sup>1,2</sup>, Kenneth J. Rothman<sup>1,3</sup>, Ellen M. Mikkelsen<sup>4</sup>, Joseph B. Stanford<sup>5</sup>, Amelia K. Wesselink<sup>1</sup>, Craig McKinnon<sup>1</sup>, Siobhan M. Gruschow<sup>1,2</sup>, Casie E. Horgan<sup>1</sup>, Aleta S. Wiley<sup>1</sup>, Kristen A. Hahn<sup>1</sup>, Henrik Toft Sørensen<sup>4</sup>, and Elizabeth E. Hatch<sup>1</sup>

<sup>1</sup>Department of Epidemiology, Boston University School of Public Health, Massachusetts

<sup>2</sup>Slone Epidemiology Center, Boston University, Massachusetts

<sup>3</sup>RTI Health Solutions, Research Triangle Park, North Carolina

<sup>4</sup>Department of Clinical Epidemiology, Aarhus University Hospital, Aarhus, Denmark

<sup>5</sup>Department of Family and Preventive Medicine, University of Utah, Salt Lake City, Utah

### Abstract

**Background**—We launched the Boston University Pregnancy Study Online (PRESTO) to assess the feasibility of carrying out an internet-based preconception cohort study in the U.S. and Canada.

**Methods**—We recruited female participants age 21–45 and their male partners through internet advertisements, word of mouth, and flyers. Female participants were randomized with 50% probability to receive a subscription to FertilityFriend.com (FF), a web-based program that collects real-time data on menstrual characteristics. We compared recruitment methods within PRESTO, assessed the cost-efficiency of PRESTO relative to its Danish counterpart (Smart-Gravid), and validated retrospectively-reported date of last menstrual period (LMP) against FF data.

**Results**—After 99 weeks of recruitment (2013–2015), 2,421 women enrolled; 1,384 (57%) invited their male partners to participate, of whom 693 (50%) enrolled. Baseline characteristics were balanced across randomization groups. Cohort retention was similar among those randomized vs. not randomized to FF (84% vs. 81%). At study enrollment, 56%, 22%, and 22% couples had been trying to conceive for <3, 3–5, and 6 months, respectively. The cost per subject enrolled was \$146 (2013 \$US), which was similar to our companion Danish study and half that of a traditional cohort study. Among FF users who conceived, >97% reported their LMP on the PRESTO questionnaire within 1 day of the LMP recorded via FF.

**Conclusions**—Use of the internet as a method of recruitment and follow-up in a North American preconception cohort study was feasible and cost-effective.

---

**Disclaimer:** All VIP Fertility Friend memberships were donated free of charge to the PRESTO research efforts. No NIH funding was used to purchase any Fertility Friend memberships.

## Keywords

fertility; internet; prospective studies; methods; mobile apps

---

## Introduction

About 15% of North American couples experience infertility, clinically defined as the inability to conceive within 12 months of unprotected intercourse.<sup>1</sup> Pregnancy rates have declined over the last several decades,<sup>2,3</sup> but few modifiable risk factors for infertility have been identified. In addition to female and male age,<sup>4</sup> several lifestyle and behavioral factors have been associated with delayed time-to-pregnancy (TTP), including smoking, obesity, and extremes of physical activity.<sup>5</sup> Results have been inconsistent across studies, most likely due to differences in exposure measurement and study design.

Prospective TTP studies are expensive and time-consuming, given the considerable effort placed on recruiting and retaining a cohort of participants that provide high-quality data. Furthermore, couples planning a pregnancy are not easily identifiable by traditional methods. Our earlier NIH-funded study, Snart-Gravid (“Soon Pregnant”)<sup>6,7</sup> and its successor study, Snart-Forældre (“Soon Parents”), evaluated the feasibility of internet-based recruitment and follow-up of a preconception cohort in Denmark.<sup>7</sup> Snart-Gravid enrolled over 5,000 women and achieved a cost per enrolled participant that was half that of a traditional prospective cohort study.<sup>7,8</sup>

The extent to which an internet-based study could be successful in North America, where cultural and privacy norms are different, is unclear. We launched the Boston University Pregnancy Study Online (PRESTO) to assess the feasibility of carrying out an internet-based preconception cohort study in the U.S. and Canada; to randomize female participants with 50% probability to receive a premium subscription to FertilityFriend.com (FF); and to assess determinants of subfertility. Similar to its Danish counterpart, PRESTO enrolls women before pregnancy and collects data prospectively. Despite the growing number of prospective TTP studies,<sup>6,9–11</sup> most TTP studies have been small, confined to a single geographic region, and retrospective in design.<sup>12</sup>

In the present report, we compare recruitment methods, assess the cost efficiency of PRESTO relative to Snart-Gravid, and validate retrospective reporting of last menstrual period (LMP) dates with prospectively-collected menstrual cycle data from FF. As secondary aims, we examine predictors of male participation, assess whether FF is associated with cohort retention, and identify predictors of software use among women randomized to FF.

## Methods

### Study Design

PRESTO is a prospective cohort study of couples residing in the U.S. and Canada that began recruitment in 2013. Eligible women are aged 21–45 years, not using contraception or fertility treatments, in a stable relationship with a male partner, planning a pregnancy and

not currently pregnant, and willing to participate in a 12-month study. There are no restrictions for male partners, other than that their female partner is already enrolled.

### Enrollment

Enrollment and primary data collection are accomplished via the study website (<http://presto.bu.edu>) and email (bupresto@bu.edu). Potential female participants read an online consent form, complete a screening questionnaire, and provide a valid e-mail address that is subsequently confirmed. The confirmatory e-mail includes a link that directs women to a comprehensive online baseline questionnaire (median completion time: 32 minutes). Women are then encouraged to invite their male partner to complete an optional one-time baseline questionnaire (median completion time: 14 minutes). Ten days after enrollment, female participants are invited to complete the Dietary Health Questionnaire (DHQ) II,<sup>13</sup> a web-based food frequency questionnaire developed and validated by the National Cancer Institute (median completion time: 40 minutes).

After completing the baseline questionnaire, women are randomized with 50% probability to receive a complimentary premium subscription to FertilityFriend.com (FF), a menstrual cycle charting and fertility information software program accessible via computer or smartphone. Existing FF users are not eligible for randomization. At every stage of the study (baseline, diet, and follow-up questionnaires), we can quantify the number of eligible participants who did not start their questionnaires or only filled out part of it. We can also compare demographic characteristics of "partial responders" with "responders" because demographic data were collected at the start of the baseline questionnaire. All data are protected using secure socket layering (SSL) encryption technology and are implemented by means of unique login names, passwords, and unique IP addresses of users' computers. On a weekly basis, updated FF data are available to PRESTO investigators via download from a secure password-protected server.

### Incentives

In addition to FF, participants are entered into several lotteries, depending on their level of participation. One of every 500 women who complete the dietary survey are entered into a lottery to win a \$100 grocery store gift card and 1/500 women who complete all of the required follow-ups are entered into a drawing to win a \$200 gift card. Women who complete the dietary survey receive a summary of their intake (macro-and micro-nutrients) and USDA recommendations for selected nutrients. To enhance male participation, 1/250 couples are randomly selected to win an iPad mini if the male partner completes his survey.

With the exception of FF, we avoid using incentives in the recruitment phase of PRESTO out of concern that it may encourage initial participation for financial rather than altruistic reasons, thereby attracting individuals who are less intent on completing the study.

### Advertising

We implemented various advertising strategies and compared their effectiveness in participant recruitment. These methods included targeted banner advertisements on Facebook.com, Twitter.com, and other social-networking sites; health-related websites such

as WebMD.com; pregnancy-related websites; and parenting blogs. We purchased “e-mail blasts” from TheBump.com, which emailed advertisements to registrants living in selected U.S. metropolitan areas. Finally, we assessed the effectiveness of traditional methods such as posting flyers in local shops and community-health centers, and advertising in print magazines and newspapers. We also mailed recruitment postcards to newly-married women residing in Massachusetts, identified from commercial mailing lists purchased from TheKnot.com.

## Data Collection

### Tracking and Follow-up

After completing the baseline questionnaire, women are contacted by email every 8 weeks for up to 12 months or until self-reported pregnancy (i.e., primary endpoint of interest), whichever occurs first. A reminder email is sent every 3 days for up to 9 days (1 initial invitation plus 3 reminders) until the questionnaire is completed. A 21-day window is given to complete the follow-up questionnaires. After >1 missed cycle of follow-up, personalized reminder emails are sent to non-respondents. If non-response appears to be due to technical problems or a changed email address, we contact the subject by telephone. The study website provides information on how a participant can discontinue participation; a “withdrawal” form is available for this purpose.

### Variables

Data self-reported on the female baseline questionnaire include: contact information (name, telephone number, birth date, primary and back-up email addresses, and home address); biographical and SES data (marital status, length of partnership, education, occupation, income); reproductive and medical history (menstrual history, cycle length, contraceptive history, gravidity, parity, birth characteristics, infertility history, medication use, medical illnesses); the Major Depression Inventory (MDI);<sup>14</sup> the Perceived Stress Scale (PSS-10);<sup>15</sup> lifestyle factors (height, weight, waist and hip circumferences, physical activity, alcohol intake, smoking, caffeine, supplement use, intercourse frequency, doing anything to improve chances of conception--e.g., charting menstrual cycles, monitoring cervical fluid quality, using ovulation predictor kit); and male characteristics (age, education, height, weight, and smoking history).

The male baseline questionnaire collects self-reported data on demographics, medical history, pubertal timing, height, weight, waist circumference, physical activity, alcohol, tobacco, caffeine, medication and supplement use, intercourse frequency, infertility, and exposure to various heat sources (e.g., saunas).

The bimonthly female follow-up questionnaires assess changes in exposures and pregnancy status. Key variables related to fecundability, such as frequency of intercourse, time apart from male partner, and doing anything to improve chances of conception, are also included. The early pregnancy questionnaire, to which a woman immediately transitions when she reports a pregnancy on the follow-up questionnaire, assesses pregnancy confirmation method, nausea and vomiting, and changes in exposures such as coffee intake and

medication use. The late pregnancy questionnaire, designed chiefly to identify late miscarriage, is completed at 32 weeks' gestation.

### **Linkage with Birth Registry**

PRESTO links questionnaire data with birth registry data from the Massachusetts Department of Public Health, which provides additional data on birth weight, gestational age, and medical interventions in pregnancy and delivery. We are seeking linkage with other birth registries in states with high numbers of participants (e.g., PA and CA).

### **Biospecimen Collection**

In a pilot project, PRESTO invited women with attempt times at study entry of 0, 1, or 2 cycles and who live or work in the Greater Boston area to provide blood and urine samples at the Boston University School of Public Health. Women were paid \$100 for their visit to the laboratory until March of 2015, after which women were paid \$50 because we added a 2<sup>nd</sup> trimester visit (\$50 compensation). We successfully collected biospecimens from 69 participants during the preconception period (58% of those invited agreed to participate). Since March of 2015, we have collected biospecimens from 7 pregnant women during the 2<sup>nd</sup> trimester (50% of those invited agreed to participate).

### **Data Analysis**

We assessed baseline characteristics of women by FF randomization status. We used log-binomial regression models to assess predictors of FF use and predictors of male participation. Probability ratios (PR) and 95% confidence intervals (CI) were adjusted for potential confounders.

We assessed the level of agreement between LMP dates reported on PRESTO questionnaires and those recorded using FF software. On each PRESTO questionnaire, women reported the first day of their LMP (i.e., first day of bleeding, not spotting). LMP dates in FF were identified using daily-recorded data on the presence and heaviness of menstrual flow. The first day of menstrual flow in a series of consecutive days with “light”, “medium”, or “heavy” flow (not spotting), was taken to be the LMP date.

In a previous publication, our research team compared the relative cost efficiency (in 2008 US\$) of Smart-Gravid (\$160 per participant enrolled) with a hypothetical study of identical scope conducted without the internet (\$322 per participant enrolled).<sup>7</sup> Adjusted for inflation to 2013 US\$, these costs are \$173 and \$348 per participant enrolled, respectively. We estimated costs for PRESTO in an identical manner, based on our actual study budget excluding institutional indirect costs.

Because PRESTO follow-up is still ongoing and many women have not had the opportunity to complete 12 months of observation, we defined “response” in this report as the proportion completing 1 follow-up questionnaire(s) or who notified us that they had experienced a pregnancy or miscarriage, had initiated fertility treatment, or were no longer trying to conceive. Response proportions were restricted to the subset of women who had the

opportunity to complete a follow-up questionnaire (8 weeks+21 days=11 weeks since completion of baseline questionnaire).

## Results

### Recruitment and Retention

After 99 weeks of recruitment (June 17, 2013 through May 10, 2015), 3,805 female screener questionnaires were completed, identifying 3,365 (88%) eligible women. Of those eligible, 2,421 (72%) enrolled by completing the female baseline questionnaire (Figure 1). Among eligible women, 21% (N=696) never started their baseline questionnaire and 7.4% (N=249) started but never finished their baseline questionnaire. Enrolled women had been trying to conceive for fewer months at entry (mean: 4.4 vs. 8.2 months) and had higher educational attainment (mean: 15.8 vs. 15.1 years) than non-enrolled women; there were no material differences by age (data not shown).

Table 1 shows baseline characteristics of female participants across enrollment groups, including women randomized to FF (N=935), women randomized not to receive FF (N=949), and women excluded from randomization because they were already FF users (N=537). Completion of 1 follow-up questionnaire in each of these groups was 751 (84%), 730 (81%), and 416 (82%), respectively. Among the 778 women who completed the early pregnancy questionnaire and were in their third trimester of pregnancy, 613 (79%) completed their late pregnancy questionnaire.

Table 2 shows response rates for the 1,350 women who had the opportunity to complete 12 months of follow-up (i.e., those enrolled through May 10, 2014). Cycle-specific response was >86% in all cycles. After 12 months of follow-up, 58% of women (N=789) became pregnant, 11% (N=148) started fertility treatment, 3% (N=45) were no longer trying to conceive, and 8% (N=109) were not yet pregnant. The remaining 19% (N=259) were lost to follow-up: 15% (N=202) were lost due to non-response during follow-up and 4% (N=57) women actively withdrew from the study.

Figure 2 shows the number of women enrolled per week and cumulative enrollment. Among the various paid methods of recruitment using similar types of banner advertising (Figure 3), Facebook was the most reliable and cost-efficient method (N=1,426 at a cost of \$27.77/participant). Advertising on Facebook also generated periodic increases in enrollment resulting from publicity on other websites where we did not advertise. For example, during week 33, a participant who learned about PRESTO via Facebook posted about the study on Reddit.com, a social networking site, which triggered PRESTO's largest peak in recruitment (N=192; Figure 2). Two similar, but smaller peaks in recruitment (weeks 35 and 72) resulted from a Reddit.com posting. The next most cost-effective vendor was TheBump.com, for which most participants were recruited via e-mail advertisements to registered members. Less cost-effective websites included Google.com (\$47.69/participant), Twitter.com (\$84.36/participant), WebMD.com (\$96.15/participant), and "mommy blogs" (\$84.09/participant; Figure 3). Our experience with blogs was inconsistent: Boston Mamas (<http://bostonmamas.com/>) and Boston Moms (<http://boston.com/life/moms>) yielded zero

participants, while Birthing Beautiful Ideas (<http://birthingbeautifulideas.com/>) yielded a cost-effective \$6.82 per participant enrolled.

### Baseline Characteristics

The median age for female participants was 30 years (range: 21–44) and for their male partners, 31 years (range: 20–60) (Table 1). More than 95% of female participants had >12 years of education, 34% had a college degree, and 39% had a graduate degree. Twenty-four percent were overweight (BMI 25–29) and 27% were obese (BMI ≥ 30). Most self-identified as “White, non-Hispanic” (83%), with 3% identifying as “Black, non-Hispanic,” 2% as “Asian, non-Hispanic,” 6% as “Hispanic/Latina,” and 4% as mixed/other race. Nearly 95% were married, and household income was \$50,000 or greater for 75% of participants. About 27% were residents of Massachusetts.

At study entry, 56%, 22%, and 22% had been trying to conceive for <3, 3–5, and ≥ 6 months, respectively. Fewer than half (48%) had ever been pregnant and 30% were parous; 14% reported a history of infertility. The most common methods of last contraception were barrier methods (39%) and oral contraceptives (36%).

### Male Participation

Of 2,421 women enrolled, 1,384 (57%) invited their male partners to participate. Of those invited, 693 (50%) enrolled (Figure 1). Women who were married and more educated, and whose partners were more educated, were more likely to invite their partners (Supplemental Table 1). Black women were less likely to invite their partners. Female and male age, income, and parity were not appreciably associated with partner invitation.

Men who were younger, more educated, and whose partners were more educated were more likely to enroll, as were men with shorter attempt times at study entry. Male partners of Black and Hispanic/Latina women were less likely to enroll. Marital status, parity, and income were not appreciably associated with male enrollment.

### Fertility Friend

Randomization to FF resulted in evenly-balanced baseline characteristics across randomization groups (Table 1). Of the 935 women randomized to receive FF, 49% (n=456) reported any FF use (Supplemental Table 2). Younger women, more educated women and women with shorter attempt times at study entry were more likely to use FF, while women recruited through Facebook and users of alternate fertility-tracking software were less likely to use FF. The most commonly-recorded features in FF were date of menses, intercourse, and cervical mucus quality (Supplemental Table 3).

### Relative Validity of Self-Reported LMP

Among the 392 women randomized to FF who provided FF data, 93% reported their FF and baseline questionnaire LMP dates within 1 day of each other (median time between reports=15 days) (Table 3). Among the 85 women who were randomized to FF, who provided FF data, and who conceived during the study period, 98% reported their FF and

early pregnancy questionnaire LMP dates within 1 day of each other (median time between reports=60 days).

### Cost

The largest costs were research personnel (53%) and subject recruitment (19%; Table 4). Subject recruitment costs included advertising and study incentives. Website maintenance and follow-up represented 5%, FF memberships 12%, and general set-up 11% of total costs. The overall cost per subject enrolled in PRESTO was \$146 (2013 \$US), which was similar to Snart-Gravid (\$173 in 2013 \$US).<sup>7</sup> The main differences in costs between the two studies were advertising (more expensive in PRESTO) and personnel costs (less expensive in PRESTO). For instance, advertising costs in PRESTO were \$28.05 per participant enrolled relative to \$17.53 per participant enrolled in Snart-Gravid (2013 US\$). Given that PRESTO has not yet reached its target enrollment (2,500 participants), we anticipate that its overall cost per participant enrolled will eventually exceed that of Snart-Gravid as recruitment continues.

### Comment

Appreciable numbers of women and their male partners were willing to enroll in an internet-based preconception cohort study in the U.S. and Canada. The proportion eligible to participate based on initial web-based screening was high (89%), suggesting effective ad targeting and self pre-screening by those accessing the website. In other words, the stated screening criteria worked well in reducing the number of individuals one would have to screen at the population level to find the targeted at risk population. This approach channels the sampling framework to those nearing or at risk for pregnancy, which enhances study efficiency given that only 1% of women are estimated to be planning pregnancy at any point in time.<sup>10</sup> We maintained relatively high rates of follow-up regardless of randomization to FF.

There was high agreement between questionnaire and FF data on LMP dates, indicating good reliability of internet-based data. Our comparison of questionnaire data and FF data may have overestimated the extent to which women validly report their LMP dates given that women could have referred to FF while completing the PRESTO questionnaires. The value of tracking the beginning dates of each menstrual cycle has long been recognized in clinical practice and epidemiologic research.<sup>16,17</sup> Many women attempting pregnancy are not familiar with the nature and timing of the fertile window, or how to identify the days of the menstrual cycle when intercourse is most likely to result in conception.<sup>18</sup> A resource that educates them on how to identify the fertile window may therefore provide useful information to accelerate TTP.<sup>19</sup> The internet provides a powerful platform to disseminate resources for menstrual cycle and fertility tracking.

Although there are several mobile computing charting applications (“apps”) to track menstrual cycles and fertility,<sup>20</sup> there has been little scientific evaluation of their efficacy. FF is a well-established, computerized menstrual cycle charting system in use online since 1998 and as a mobile phone app since 2008, and claiming over 650,000 pregnancies among its users to date.<sup>21</sup> FF allows women to enter multiple signs and symptoms of the menstrual



cycle, including cervical mucus and urinary hormones, the biomarkers that have the most evidence supporting their use to identify days when intercourse is most likely to result in pregnancy.<sup>22</sup>

The cost per participant enrolled in PRESTO was similar to Smart-Gravid, though Smart-Gravid did not use incentives and advertising costs were higher in PRESTO. The response proportion was lower in PRESTO (82%) than Smart-Gravid (84%).<sup>7</sup> The slightly lower response proportions in PRESTO may reflect greater concerns about data privacy and potential for breach of confidentiality among North Americans. Although such a breach might be perceived to have greater consequences in countries without universal health care (e.g., U.S.) than in those with universal health care (e.g., Denmark and Canada), response proportions were higher among U.S. than Canadian participants (84% vs. 73%). Thus, there are likely other yet-to-be-identified cross-national differences in attitudes and behaviors toward research participation.

About 83% of PRESTO participants self-identified as non-Hispanic White, allowing for meaningful subgroup analyses among minority participants. Although we expected advertising on Twitter to increase minority enrollment because Black women are more likely to use the social-networking site,<sup>23</sup> Twitter was not an effective method of participant recruitment. Another possible barrier to the recruitment of minorities was the English language of the questionnaires. Future translation of the questionnaires into other languages such as Spanish may increase the diversity of our cohort.

PRESTO's restriction to women younger than 21 years of age may limit the generalizability of results to younger, more fecund women. For example, younger women may differ from older women in terms of educational attainment and other sociodemographic characteristics that could potentially modify the extent to which primary associations of interest affect fertility and pregnancy outcomes.

Men are less likely to participate in epidemiologic studies.<sup>24,25</sup> Despite being an optional component of PRESTO, 57% of women invited their male partners to participate, and 50% of those invited enrolled. Educated, married, White Non-Hispanic women were more likely to invite their male partners. Once invited, younger age, greater education, White non-Hispanic race/ethnicity, and shorter attempt time at study entry predicted male participation. These patterns of male participation are generally consistent with those found in a U.S.-based prospective cohort study of male reproductive health.<sup>25</sup>

Patterns of internet use are similar in the U.S. and Canada,<sup>26,27</sup> with more than 85% of adults reporting use. The prevalence of internet use among reproductive age individuals is even higher, given that older individuals are less likely to use the internet. Women and men use the internet in similar proportions.<sup>26,27</sup> White and Asian non-Hispanics have a higher prevalence of internet usage than Hispanics and Black non-Hispanics, but the prevalence was at least 60% in all racial/ethnic groups.<sup>26</sup> Likewise, individuals across most education and income levels have relatively high internet usage; the only group with usage below 50% were households with incomes <\$25,000 per year.<sup>26,27</sup>

Internet-based recruitment and follow-up can be carried out at a lower cost relative to other studies because the internet removes the dependence of in-person interviews or mailed questionnaires, both of which are time consuming and expensive.<sup>7,28</sup> It is easier to maintain regular contact with subjects through a study website and periodic email reminders. An internet-based study may also be more effective at targeting pregnancy planners, a population that rarely makes its pregnancy intentions known publicly.

Pregnancy planners are unlikely to be searching for health information about conception (unless they are experiencing subfertility) and this may explain why Google's "active" advertising model, which is designed to show ads when specific "key words" are searched, was less effective. Facebook's "passive" advertising strategy, showing ads on News Feeds or computer screen side-panels, appeared to work more effectively in reaching our population. The ability to target Facebook advertisements by relationship status (e.g., newlywed) or parental status (e.g., parents of toddlers), and demographics such as gender, age, and geographic region, allowed us to direct ads to those likely to be eligible for PRESTO. A Facebook page (<http://facebook.com/bostonuniversitypresto>) provides information about PRESTO and lends legitimacy to the study.

The internet is frequently used as a method of recruitment and follow-up in clinical trials, and many prospective cohort studies have incorporated internet-based data collection into ongoing studies. To our knowledge, few studies have relied solely on the internet for recruitment and follow-up. The Nutrinet Sante study is enrolling and following 100,000 participants from the general population for ten years to investigate links between diet and health.<sup>29</sup> Other internet-based studies have been conducted within occupational groups,<sup>30</sup> among patients with a specific disease,<sup>31</sup> and among reproductive-aged women.<sup>32,33</sup> These studies generally indicate that the internet is cost-effective and acceptable to participants.

Internet-based recruitment has been criticized because the characteristics of those with and without internet access differ and, among those with internet access, those who choose to participate in studies may differ from those who do not.<sup>34</sup> However, studies have shown that even when participation at cohort enrollment is related to characteristics such as age, parity, or smoking, measures of association are not necessarily biased due to self-selection.<sup>35-37</sup> Concerns about selection bias stemming from length-biased recruitment of longer TTPs can be assessed by stratifying by attempt time at study entry and, if needed, controlled by restricting analysis to couples with shorter attempt times at study entry.

The present study demonstrates the feasibility of using the internet as a method of recruitment, follow-up, and data collection in a North American preconception cohort study. The cost of carrying out an internet-based preconception cohort study was estimated to be half that of a traditional epidemiologic study. These findings may inform recruitment strategies for other internet-based cohort studies in North America.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

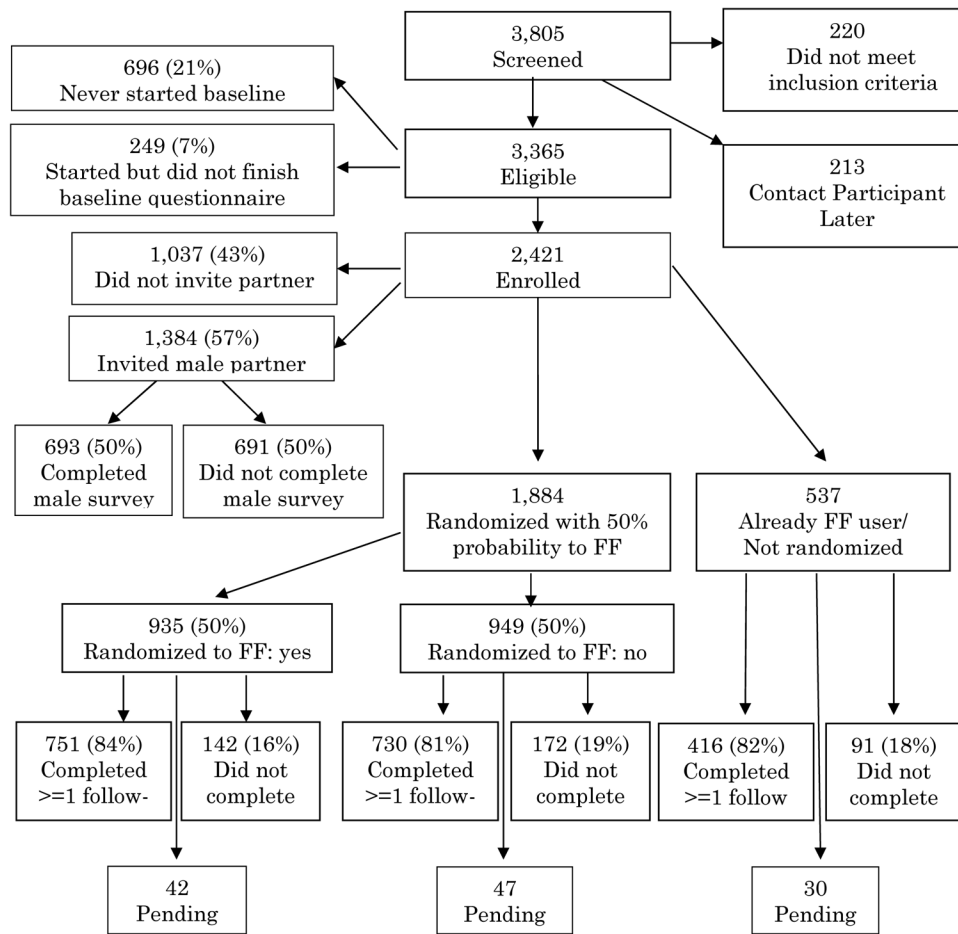
We acknowledge the contributions of PRESTO participants and staff. We thank Mr. Michael Bairos for his technical support with developing the web-based infrastructure of PRESTO and Dr. Donna Baird for her feedback on questionnaire design. We are grateful to Drs. Frederic Montoya and Heather Bromberg for their generous donation of FertilityFriend.com memberships. Lastly, we thank Ms. Alexandra E. Kriss, Ms. Alina Chaiyasarikul, Ms. Alanna Raskin, Ms. Tara Bylsma, Mr. James Michiel, Dr. Rose G. Radin, Dr. Sam Ransbotham, Dr. Stephanie Jernigan, and Dr. Matthew O. Hunt for their technical assistance and instrumental support of the study.

**Source of funding:** NICHD grant R21HD072326.

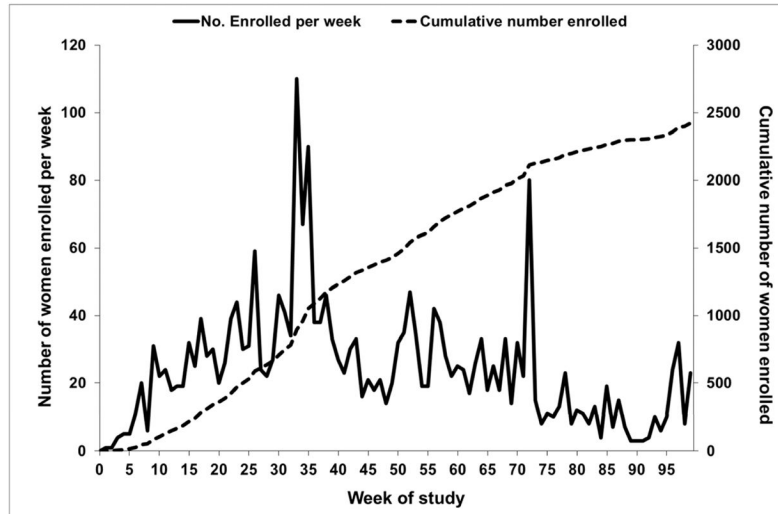
## References

1. Thoma ME, McLain AC, Louis JF, King RB, Trumble AC, Sundaram R, et al. Prevalence of infertility in the United States as estimated by the current duration approach and a traditional constructed approach. *Fertil Steril*. 2013; 99:1324–31. e1. [PubMed: 23290741]
2. Curtin SC, Abma JC, Ventura SJ, Henshaw SK. Pregnancy rates for U.S. women continue to drop. *NCHS Data Brief*. 2013:1–8. [PubMed: 24314113]
3. Martin JA, Hamilton BE, Osterman MJ, Curtin SC, Matthews TJ. Births: final data for 2013. *National Vital Statistics Reports*. 2015; 64:1–65.
4. Female age-related fertility decline. Committee Opinion No 589. *Fertil Steril*. 2014; 101:633–4. [PubMed: 24559617]
5. Buck Louis GM, Lynch CD, Cooney MA. Environmental influences on female fecundity and fertility. *Semin Reprod Med*. 2006; 24:147–55. [PubMed: 16804813]
6. Mikkelsen EM, Hatch EE, Wise LA, Rothman KJ, Riis A, Sorensen HT. Cohort Profile: The Danish Web-based Pregnancy Planning Study--‘Snart-Gravid’. *Int J Epidemiol*. 2009; 38:938–43. [PubMed: 18782897]
7. Huybrechts KF, Mikkelsen EM, Christensen T, Riis AH, Hatch EE, Wise LA, et al. A successful implementation of e-epidemiology: the Danish pregnancy planning study ‘Snart-Gravid’. *Eur J Epidemiol*. 2010
8. Rothman KJ, Mikkelsen EM, Riis A, Sørensen HT, Wise LA, Hatch EE. Randomized Trial of Questionnaire Length. *Epidemiology*. 2009; 20:154. [PubMed: 19234404]
9. Steiner AZ, Long DL, Tanner C, Herring AH. Effect of vaginal lubricants on natural fertility. *Obstet Gynecol*. 2012; 120:44–51. [PubMed: 22914390]
10. Buck GM, Lynch CD, Stanford JB, Sweeney AM, Schieve LA, Rockett JC, et al. Prospective pregnancy study designs for assessing reproductive and developmental toxicants. *Environ Health Perspect*. 2004; 112:79–86. [PubMed: 14698935]
11. Schisterman EF, Silver RM, Perkins NJ, Mumford SL, Whitcomb BW, Stanford JB, et al. A randomised trial to evaluate the effects of low-dose aspirin in gestation and reproduction: design and baseline characteristics. *Paediatr Perinat Epidemiol*. 2013; 27:598–609. [PubMed: 24118062]
12. Joffe M, Key J, Best N, Keiding N, Scheike T, Jensen TK. Studying time to pregnancy by use of a retrospective design. *Am J Epidemiol*. 2005; 162:115–24. [PubMed: 15972942]
13. National Cancer Institute. [Last accessed: March 10, 2012] Diet History Questionnaire II: Web-based DHQ. <http://riskfactor.cancer.gov/dhq2/webquest/>
14. Olsen LR, Jensen DV, Noerholm V, Martiny K, Bech P. The internal and external validity of the Major Depression Inventory in measuring severity of depressive states. *Psychol Med*. 2003; 33:351–6. [PubMed: 12622314]
15. Cohen S, Kamarck T, Mermelstein R. A global measure of perceived stress. *J Health Soc Behav*. 1983; 24:385–96. [PubMed: 6668417]
16. Diaz A, Laufer MR, Breech LL. Menstruation in girls and adolescents: using the menstrual cycle as a vital sign. *Pediatrics*. 2006; 118:2245–50. [PubMed: 17079600]
17. Dasharathy SS, Mumford SL, Pollack AZ, Perkins NJ, Mattison DR, Wactawski-Wende J, et al. Menstrual bleeding patterns among regularly menstruating women. *Am J Epidemiol*. 2012; 175:536–45. [PubMed: 22350580]

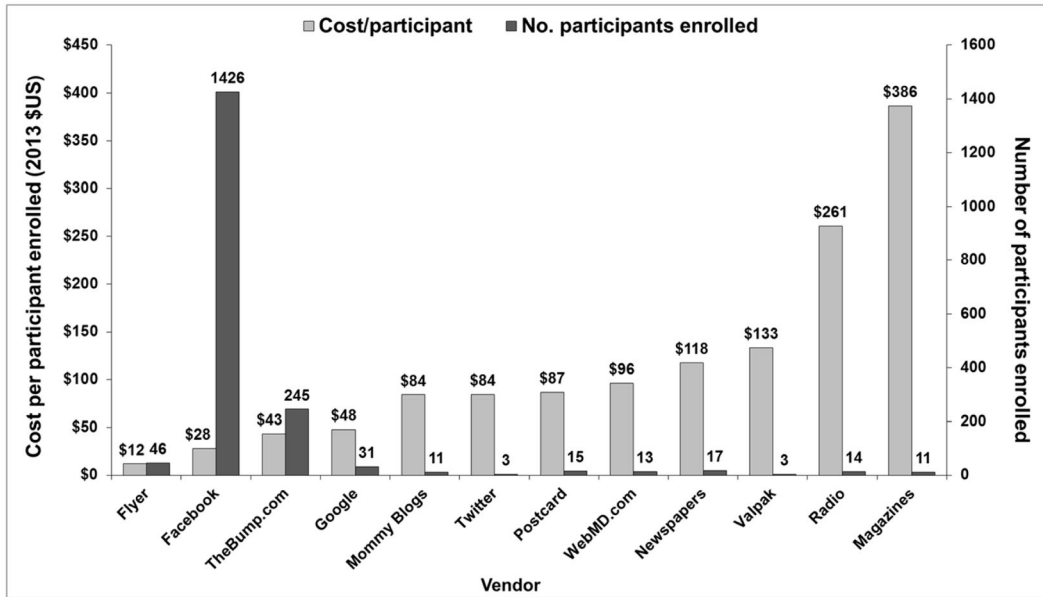
18. Lundsberg LS, Pal L, Garipey AM, Xu X, Chu MC, Illuzzi JL. Knowledge, attitudes, and practices regarding conception and fertility: a population-based survey among reproductive-age United States women. *Fertil Steril*. 2014; 101:767–74. [PubMed: 24484995]
19. Stanford JB, Dunson DB. Effects of sexual intercourse patterns in time to pregnancy studies. *Am J Epidemiol*. 2007; 165:1088–95. [PubMed: 17289774]
20. [accessed February 16, 2014] <http://techcrunch.com/2014/07/15/vcs-find-fertile-ground-in-womens-health/>
21. <http://www.fertilityfriend.com/about.html>, accessed February 16, 2015. (Accessed at
22. Stanford JB, White GL, Hatasaka H. Timing intercourse to achieve pregnancy: current evidence. *Obstet Gynecol*. 2002; 100:1333–41. [PubMed: 12468181]
23. [Last accessed: March 10, 2012] Twitter Usage in America. 2010. [http://www.edisonresearch.com/home/archives/2010/04/twitter\\_usage\\_in\\_america\\_2010\\_1.php](http://www.edisonresearch.com/home/archives/2010/04/twitter_usage_in_america_2010_1.php)
24. Galea S, Tracy M. Participation rates in epidemiologic studies. *Ann Epidemiol*. 2007; 17:643–53. [PubMed: 17553702]
25. Olshan AF, Perreault SD, Bradley L, Buus RM, Strader LF, Jeffay SC, et al. The healthy men study: design and recruitment considerations for environmental epidemiologic studies in male reproductive health. *Fertility and Sterility*. 87:554–64. [PubMed: 17140573]
26. File, T.; Ryan, C. American Community Survey Reports, ACS-28. Vol. 2014. U.S. Census Bureau; Washington, DC: 2014. Computer and internet use in the United States: 2013.
27. Internet-use by individuals, by selected characteristics. Statistics Canada; <<http://www.statcan.gc.ca/tables-tableaux/sum-som/101/cst01/comm35a-eng.htm>> [Accessed 2010 May 11]
28. Harris ML, Herbert D, Loxton D, Dobson A, Wigginton B, Lucke JC. Recruiting young women for health surveys: traditional random sampling methods are not cost-effective. *Aust N Z J Public Health*. 2014; 38:495. [PubMed: 25269982]
29. Andreeva VA, Galan P, Julia C, Castetbon K, Kesse-Guyot E, Hercberg S. Assessment of response consistency and respective participant profiles in the Internet-based NutriNet-Sante Cohort. *Am J Epidemiol*. 2014; 179:910–6. [PubMed: 24521560]
30. Turner C, Bain C, Schluter PJ, Yorkston E, Bogossian F, McClure R, et al. Cohort Profile: The Nurses and Midwives e-Cohort Study--a novel electronic longitudinal study. *Int J Epidemiol*. 2009; 38:53–60. [PubMed: 18202083]
31. Long MD, Kappelman MD, Martin CF, Lewis JD, Mayer L, Kinneer PM, et al. Development of an internet-based cohort of patients with inflammatory bowel diseases (CCFA Partners): methodology and initial results. *Inflamm Bowel Dis*. 2012; 18:2099–106. [PubMed: 22287300]
32. Mishra GD, Hockey R, Powers J, Loxton D, Tooth L, Rowlands I, et al. Recruitment via the Internet and social networking sites: the 1989–1995 cohort of the Australian Longitudinal Study on Women's Health. *J Med Internet Res*. 2014; 16:e279. [PubMed: 25514159]
33. Richiardi L, Baussano I, Vizzini L, Douwes J, Pearce N, Merletti F. Feasibility of recruiting a birth cohort through the Internet: the experience of the NINFEA cohort. *Eur J Epidemiol*. 2007; 22:831–7. [PubMed: 17955333]
34. Keiding N, Slama R. Commentary: time-to-pregnancy in the Real World. *Epidemiology*. 2015; 26:119–21. [PubMed: 25437318]
35. Nohr EA, Frydenberg M, Henriksen TB, Olsen J. Does low participation in cohort studies induce bias? *Epidemiology*. 2006; 17:413–8. [PubMed: 16755269]
36. Nilsen RM, Vollset SE, Gjessing HK, Skjaerven R, Melve KK, Schreuder P, et al. Self-selection and bias in a large prospective pregnancy cohort in Norway. *Paediatr Perinat Epidemiol*. 2009; 23:597–608. [PubMed: 19840297]
37. Hatch EE, Hahn KA, Wise LA, Mikkelsen EM, Kumar R, Fox MP, et al. Evaluation of Selection Bias in an Internet-based Study of Pregnancy Planners. *Epidemiology*. 2015 in press.



**Figure 1.** Flow Chart of Enrollment and Follow-up in PRESTO (2013–2015)



**Figure 2.** Total number of female participants enrolled in PRESTO (2013–2015)  
 \*Peaks at 33 (n=110), 35 (n=90), and 72 (n=80) weeks are result of PRESTO participant posting on Reddit.com



**Figure 3.**

PRESTO: Cost per participant enrolled and total number enrolled, by vendor <sup>a</sup>

<sup>a</sup>Includes paid advertising only. Women were permitted to report more than 1 way of hearing about the study (2% of women reported >1 source). “Mommy blogs” includes Birthing Beautiful Ideas (11 participants at \$6.82 per participant enrolled), Boston Mamas (0 participants for total of \$450 for advertising), and BostonMoms (0 participants for total of \$400 advertising). “Magazines” includes Boston Parents’ Paper (6 participants at \$125 per participant enrolled), The Knot (1 participant at \$500 per participant enrolled), and Improper Bostonian (4 participants at \$750 per participant enrolled).

**Table 1**

Baseline characteristics of women enrolled in PRESTO (2013–2015).

Characteristic	Enrolled Women	Randomized to FF		Already a FF user
		Yes	No	
Number of women (%)	2,421 (100%)	935 (38.6%)	949 (39.2%)	537 (22.2%)
Age at baseline, years (median, range)	30 (21–44)	30 (21–44)	30 (21–44)	29 (21–42)
Age at baseline, years (%)				
<25	9.1	10.0	8.4	8.9
25–29	40.0	38.2	40.2	42.8
30–34	36.9	37.8	36.5	36.3
35–39	12.2	12.2	13.1	10.6
40	1.8	1.9	1.9	1.3
Male partner's age, years (median, range)	31 (20–60)	32 (20–60)	32 (20–55)	31 (21–49)
Attempt time at study entry, months (%)				
<3	55.7	56.7	57.7	50.5
3–5	21.9	22.0	20.6	24.0
6–11	13.8	12.4	12.8	17.9
12	8.6	8.9	9.0	7.6
Education, years (%) <sup>a</sup>				
<12	0.7	0.9	0.8	0.2
12	3.5	3.5	4.7	1.3
13–15	22.7	23.5	22.3	21.8
16	34.3	35.0	32.1	37.1
17	38.8	37.1	39.9	39.7
Body mass index, kg/m <sup>2</sup> (%)				
<18.5	2.1	2.0	1.7	2.8
18.5–24.9	47.5	47.0	48.0	47.6
25.0–29.9	23.6	24.1	24.2	21.5
30.0	26.9	26.9	26.1	28.2
Race (%)				
White, non-Hispanic	83.4	81.2	82.4	89.2
Black, non-Hispanic	3.2	3.3	3.9	1.9
Asian, non-Hispanic	2.2	2.7	2.0	1.5
Hispanic/Latina	6.3	6.7	6.9	4.7
Mixed/other race	3.8	4.7	3.6	2.4
Married (%)	94.4	94.6	93.9	95.2
Household income, 2013 US\$ (%)				
<\$50,000	20.4	22.8	20.1	16.8
\$50,000–\$99,999	35.9	32.8	35.8	41.3
\$100,000–\$149,999	24.0	25.0	22.8	24.6
\$150,000	15.2	14.4	16.0	15.1
Refused/Don't know/Missing	4.5	4.9	5.3	2.2



Characteristic	Enrolled Women	Randomized to FF		Already a FF user
		Yes	No	
Gravid (%)	48.4	48.5	47.1	50.5
Parous (%)	30.2	30.3	31.0	28.9
History of infertility (%) <sup>b</sup>	14.0	13.9	14.6	13.2
History of infertility among those who have tried to conceive previously (%) <sup>c</sup>	21.5	22.1	19.3	21.2
Smoking Status (%)				
Non-smoker	88.1	86.3	88.5	90.5
Smoker (Occasional)	4.3	6.0	3.6	2.4
Smoker (Regular)	7.6	7.7	7.9	7.1
Last method of contraception (%) <sup>c</sup>				
Oral contraceptives	36.4	35.7	35.6	38.8
Other hormonal methods	3.6	4.1	4.2	1.5
Barrier methods	39.0	38.2	40.3	38.0
Withdrawal/rhythm/other	21.1	21.9	19.9	21.7
Use of multivitamins, prenatal vitamins, or folate supplements (%)	79.9	75.1	78.0	91.6
Recruited via Facebook (%)	58.9	68.0	66.6	29.4
Resident of Massachusetts (%)	26.5	28.2	28.0	20.9
Resident of Canada (%)	10.9	10.3	12.7	8.7

FF=FertilityFriend.com; TTP=time-to-pregnancy. Columns 2 through 4 sum to total number of participants in study (column 1).

<sup>a</sup>59 participants missing data on education of male partner.

<sup>b</sup>Defined as trying to conceive without success for  $\geq 12$ , including all participants.

<sup>c</sup>Defined as trying to conceive without success for  $\geq 12$  months, excluding 814 women who reported never having tried to conceive previously.

<sup>d</sup>“Other hormonal methods” include contraceptive patch, injectables, or implantable rods. “Other” methods include calendar method, monitoring of cervical mucus, menstrual charting, abstinence, douching, or natural family planning.

**Table 2** Response rates at each follow-up questionnaire, based on the 1,350 participants with opportunity to complete 12 months of follow-up<sup>a</sup>

Follow-up Cycle	Population at-risk	Completed follow-up				Lost to follow-up			Risk of dropout	Conditional probability of remaining under follow-up	Cumulative probability of remaining under follow-up (90% CI)	
		Total	Pregnant <sup>b</sup>	Started fertility treatment	No longer trying to conceive	Total	Non-response	Withdrawn				
				Q	WF	Q	WF					
1	1350	287	264	11	1	8	3	137	135	2	0.11	0.89 (0.87, 0.90)
2	926	224	185	32	0	6	1	24	22	2	0.03	0.86 (0.84, 0.88)
3	678	159	119	31	3	4	2	21	8	13	0.04	0.83 (0.81, 0.85)
4	498	135	93	29	2	7	4	25	10	15	0.06	0.78 (0.76, 0.80)
5	338	74	49	18	0	7	0	41	21	20	0.14	0.68 (0.64, 0.71)
6	223	103	79	20	1	0	3	11	6	5	0.06	0.63 (0.59, 0.67)
Total		982	789	141	7	32	13	259	202	57		

Abbreviations: Q=reported on questionnaire; WF=reported on withdrawal form.

<sup>a</sup>This table is structured as a life table. For example, the population at-risk in cycle 2 (N=926) can be derived from the cycle 1 information: 1,350 at risk - 264 women who were censored because they became pregnant - 12 women who were censored because they started fertility treatment - 11 women who were censored because they reported no longer trying to conceive - 137 women who were lost to follow-up = 926 women under observation at the start of cycle 2.

<sup>b</sup>Includes pregnancy losses.

**Table 3**

Difference in reported dates of last menstrual period (LMP) comparing self-reported questionnaire data and data from Fertility Friend software application<sup>a</sup>

Difference in reported LMP (days)	Women at baseline <sup>b</sup> (N=392)		Women who conceived during study period <sup>c</sup> (N=85)	
	Baseline questionnaire vs. FF		Early pregnancy questionnaire vs. FF	
	N	%	N	%
-3	12	3.0	0	0.0
-2	5	1.3	0	0.0
-1	14	3.6	2	2.3
0	331	84.4	77	90.6
1	21	5.4	4	4.7
2	1	0.3	2	2.4
3	8	2.0	0	0.0

<sup>a</sup>Restricted to women who were randomized to Fertility Friend and who provided complete data on LMP dates from the Fertility Friend software application (b) at time of baseline questionnaire or (c) at time of conception during study period.

**Table 4**

Costs of recruiting and following 2,421 female PRESTO participants

<b>Study cost components</b>	
General set-up	
Website construction	\$20,183
Development of e-mail reminder system	\$3,441
Other <sup>a</sup>	\$20,321
Subject recruitment (advertisement, media strategy, incentives <sup>b</sup> )	\$67,898
Website maintenance and follow-up	\$16,831
FertilityFriend.com VIP memberships (in-kind donation: \$45 × 935)	\$42,075
Research personnel <sup>c</sup>	\$187,432
<b>Total costs</b>	<b>\$353,181</b>
<b>Per subject cost (2013 US\$)</b>	<b>\$146</b>

<sup>a</sup> Includes costs associated with quality assurance, system documentation, and coordination between research and system development teams.

<sup>b</sup> Includes lotteries but not FertilityFriend.com memberships. FertilityFriend.com memberships were donated in-kind and no NIH funds were used to cover this expense.

<sup>c</sup> Includes for unpaid internships completed by undergraduate and graduate students at Boston University (assuming 10 hours/week for 12 weeks at entry level wage (\$12/hour) + student fringe rate of 9.5%).