Reproductive Health

COMMENTARY   **Open Access**

CrossMark

Published analysis of contraceptive effectiveness of Daysy and DaysyView app is fatally flawed

Chelsea B. Polis[1,2]

Abstract

Background: In March 2018, Dr. Martin C. Koch and colleagues published an analysis purporting to measure the effectiveness of the Daysy device and DaysyView app for the prevention of unintended pregnancy. Unfortunately, the analysis was flawed in multiple ways which render the estimates unreliable. Unreliable estimates of contraceptive effectiveness can endanger public health.

Main body: This commentary details multiple concerns pertaining to the collection and analysis of data in Koch et al. 2018. A key concern pertains to the inappropriate exclusion of all women with fewer than 13 cycles of use from the Pearl Index calculations, which has no basis in standard effectiveness calculations. Multiple additional methodological concerns, as well as prior attempts to directly convey concerns to the manufacturer regarding marketing materials based on prior inaccurate analyses, are also discussed.

Conclusion: The Koch et al. 2018 publication produced unreliable estimates of contraceptive effectiveness for the Daysy device and DaysyView app, which are likely substantially higher than the actual contraceptive effectiveness of the device and app. Those estimates are being used in marketing materials which may inappropriately inflate consumer confidence and leave consumers more vulnerable than expected to the risk of unintended pregnancy. Prior attempts to directly convey concerns to the manufacturer of this device were unsuccessful in preventing publication of subsequent inaccurate analyses. To protect public health, concerns with this analysis should be documented in the published literature, the Koch et al. 2018 analysis should be retracted, and marketing materials on contraceptive effectiveness should be subjected to appropriate oversight.

Keywords: Daysy, DaysyView, Fertility awareness-based methods, Contraceptive effectiveness

Background

In March 2018, Martin C. Koch et al. published a paper entitled "Improving usability and pregnancy rates of a fertility monitor by an additional mobile application: results of a retrospective efficacy study of Daysy and DaysyView app" [1]. Unfortunately, this paper contains fatal flaws in the estimation of effectiveness of the Daysy device and DaysyView app for prevention of unintended pregnancy, which render the published effectiveness estimates unreliable. The published estimates are likely to be substantially higher than the actual contraceptive effectiveness of the device and app.

This commentary argues that this paper merits retraction [2], given the demonstrated potential for these estimates to be used in marketing materials, leading to public confusion about the actual contraceptive effectiveness of the Daysy device and DaysyView app. The manufacturer of Daysy (Valley Electronics AG, Zurich, Switzerland) has cited the Koch et al. 2018 publication in contraceptive effectiveness claims in marketing materials (Fig. 1). This could lead to inappropriately inflated consumer confidence in the contraceptive effectiveness of Daysy and DaysyView, and could leave consumers more vulnerable to the risk of unintended pregnancy.

Main text

The Koch et al. 2018 effectiveness estimates are fatally flawed and unreliable for multiple reasons. First, their

Correspondence: cpolis@guttmacher.org
[1]Guttmacher Institute, 125 Maiden Lane, 7th Floor, New York, NY 10038, USA
[2]Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD 21205, USA

BMC
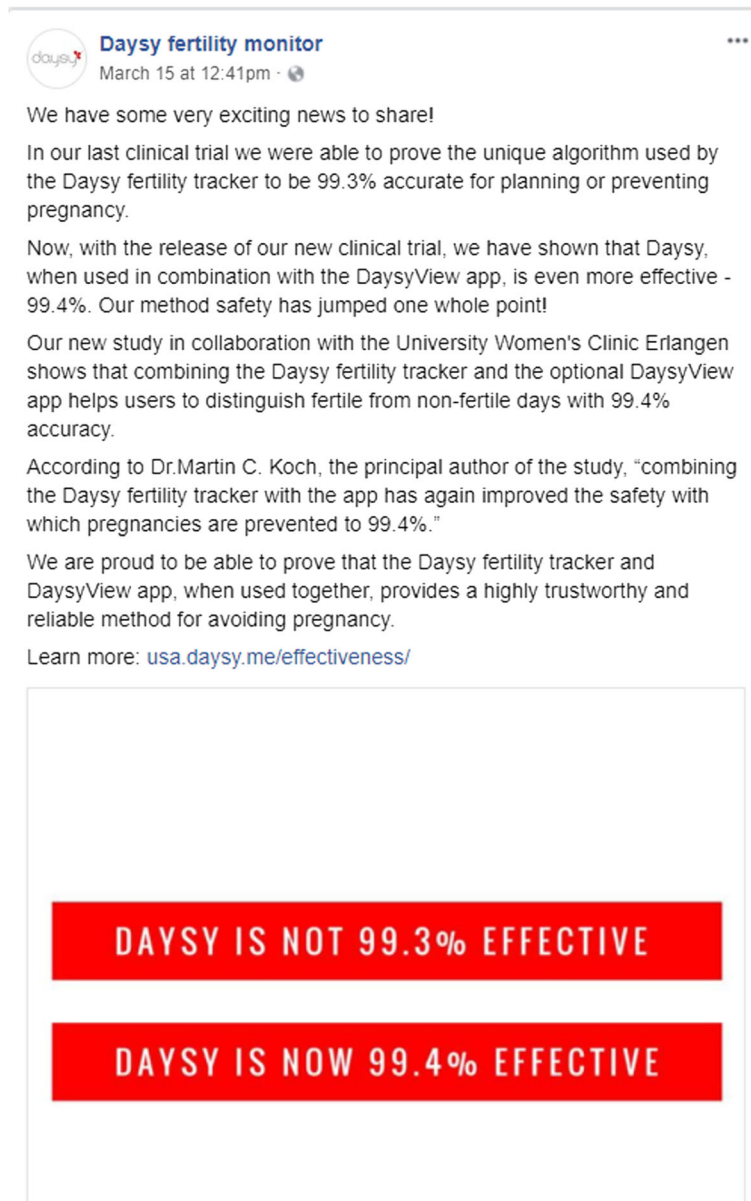
Fig. 1 DaysyUSA Facebook post claiming a 99.4% effectiveness of Daysy and the DaysyView app (based on results of the inaccurate Koch et al. 2018 analysis)

Pearl Index calculations (which are less preferable than life table estimates [3], yet are the only estimates acknowledged in the abstract of the Koch et al. 2018 publication and the Valley Electronics marketing language shown in Fig. 1) are void of meaning since the authors inappropriately excluded from the Pearl Index calculations all women with fewer than 13 cycles of use, ignoring cycle and pregnancy information from the majority of study participants. This approach has no basis in standard effectiveness estimation approaches, and would severely underestimate unintended pregnancy rates by inappropriately excluding women who may be at greatest risk of unintended pregnancy (i.e., those who are least experienced with use of the contraceptive method and most fertile) [4]. Table 2 of Koch et al. 2018 suggests that at least 10 pregnancies occurred to women with fewer than 13 cycles (and potentially more, given broader concerns about pregnancy ascertainment, as discussed below), but these pregnancies were not included in the Pearl Index calculations.

In addition, the investigators did not prospectively collect information regarding perfect or imperfect use

of the method for each cycle. Instead, they asked about overall perfect/imperfect use (intended to reflect the entire duration of use) via a retrospective survey. This precludes their ability to correctly calculate perfect use unintended pregnancy rates, in either the Pearl Index or life table calculations [5]. Thus, the so-called "perfect use" unintended pregnancy rates in Koch et al. 2018 are not comparable to correctly estimated perfect use pregnancy rates for other contraceptive methods in other studies.

Importantly, even if the inappropriate exclusion of pregnancies were corrected, the underlying data remain unreliable, given the weak approach to pregnancy ascertainment. First, the survey participation rate was low; only 13% (798/ 6278) of the invited population participated. Furthermore, eligibility for participation was limited to registered users with a DaysyView account, but no information is provided on what proportion of overall Daysy users this would represent. This makes the results vulnerable to selection bias and incomplete ascertainment of pregnancies among all Daysy users. Second, the authors note that if a woman indicated on the retrospective survey that she had an "unwanted" pregnancy, *then* the pregnancy was "verified from the user's dataset" using a definition of pregnancy of "elevated temperature of longer than 18 days, or if the user stopped using the device during the luteal phase." However, the authors do not state that temperature or device use information was evaluated among *all* women (to attempt to prospectively identify all pregnancies). If the criteria for examining the temperature/device use data were premised upon the woman self-reporting an "unwanted" pregnancy on the survey, an unknown number of pregnancies may have been excluded, given that unwanted pregnancies, including those ending in abortion, may be underreported [6, 7]. Furthermore, the survey question asked women to report "unwanted" pregnancies occurring during Daysy use. The term "unwanted" is not synonymous with "unintended" [8], and it is unclear how women may have interpreted this question – but standard contraceptive effectiveness estimates include all unintended pregnancies (including mistimed and unwanted pregnancies). Under-ascertainment of unintended pregnancies would lead to inflated estimates of contraceptive effectiveness.

Several other aspects of the questionnaire also raise concern. For example, a survey question presumably intended to ascertain if (at baseline) Daysy was being used to avoid or attempt pregnancy offered the following response options: (A) To avoid a pregnancy, (B) For family planning, (C) Both. However, response options A and B (and therefore, also C) are poorly distinguished and likely to be perceived synonymously by some respondents, meaning that even this already limited measure of baseline pregnancy intentions is unreliable. Also, this retrospective survey was unable to capture and address changing pregnancy intentions over time to attempt to prospectively characterize pregnancies as intended (and therefore, excluded from estimates of contraceptive effectiveness) or unintended (and therefore, included in estimates of contraceptive effectiveness).

While some issues in data collection and analysis are not unique to this study, they nonetheless raise additional concern about the accuracy of the estimates. For example, no inclusion/exclusion criteria were described to ensure that the analytic population was at meaningful risk of pregnancy (i.e., excluding women likely to be subfertile). Also, the majority of survey respondents (64%) reported concurrent use of contraceptive methods in addition to Daysy; the potential confounding effect of use of other methods on the effectiveness estimates is not addressed in the analysis.

For these reasons, this analysis produced estimates which cannot be understood as reflective of the true (and still unknown) contraceptive effectiveness of the Daysy device and DaysyView app. Our forthcoming systematic review of studies assessing the effectiveness of various fertility awareness-based methods (FABMs) [9] carefully considers the quality of prospective studies on effectiveness of various FABMs. Studies of extremely poor quality, such as those in Koch et al. 2018, must not be understood as providing reliable evidence on contraceptive effectiveness.

The manufacturer of the Daysy device, Valley Electronics, has made inaccurate marketing statements in the past (Figs. 2 and 3), based on previously published analyses [10, 11] which purported to assess contraceptive effectiveness of their products, but which also contained fatal flaws.[1] In September 2017, I contacted the Director of Medical Affairs for Valley Electronics to directly express concerns regarding the methodological issues in these prior analyses [10, 11], and flag concern regarding the reliability of this information. I encouraged the company to "alert Daysy users and potential customers that at present, Daysy lacks sufficient evidence to be promoted as a tool for pregnancy prevention," and to "focus instead on collecting high quality evidence to adequately assess the true effectiveness of this device" [12]. I was informed that my concerns would be discussed with the advisory board. However, the exchange appears to have been unsuccessful in encouraging the company to ensure accuracy in subsequently published estimates and marketing language. Therefore, additional steps are needed to protect scientific integrity and public health.

Conclusions

The scientific and reproductive health communities have a responsibility to protect public health by ensuring that
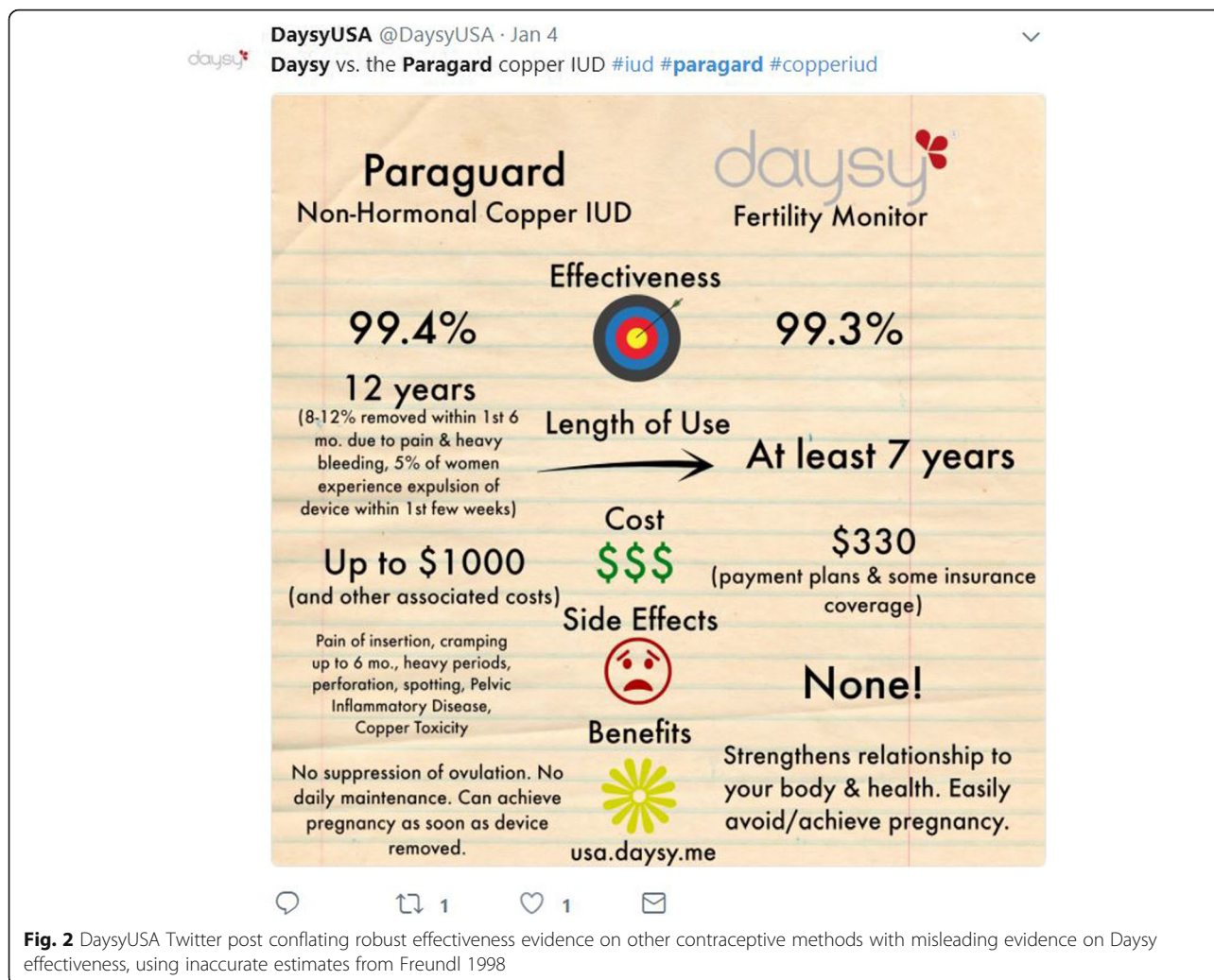
Fig. 2 DaysyUSA Twitter post conflating robust effectiveness evidence on other contraceptive methods with misleading evidence on Daysy effectiveness, using inaccurate estimates from Freundl 1998

inaccurately conducted analyses of contraceptive effectiveness do not put unsuspecting consumers at a greater than expected risk of experiencing an unintended pregnancy. Women and couples interested in using any form of contraception, including an FABM such as Daysy, deserve robust effectiveness data on which to base their contraceptive decisions. Unfortunately, the published Koch et al. 2018 estimates are fatally flawed and inaccurate, and are likely to be substantially higher than the actual contraceptive effectiveness of the device and app. By using marketing language based on inaccurate analyses, Valley Electronics falsely
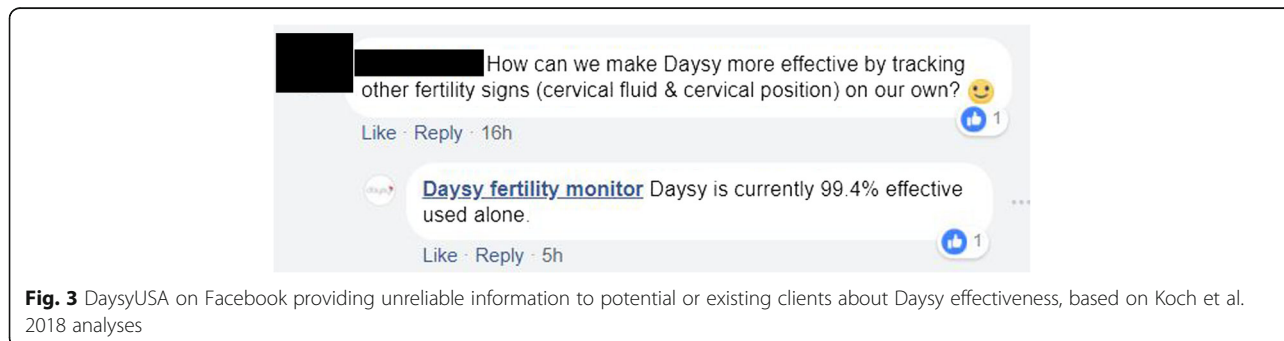

Fig. 3 DaysyUSA on Facebook providing unreliable information to potential or existing clients about Daysy effectiveness, based on Koch et al. 2018 analyses

increases consumer confidence in the effectiveness of the Daysy device and the DaysyView app, which could endanger the well-being of their customers. Prior efforts to directly communicate methodological concerns to the company do not appear to have led to enhanced caution in ensuring the accuracy of their published effectiveness estimates or marketing language. Therefore, to protect public health, scientific integrity, and potential consumers, these concerns should be documented in the published literature, [13] the Koch et al. 2018 analysis should be retracted, [2] and marketing materials on contraceptive effectiveness should be subjected to appropriate oversight.

Endnotes

[1]For example, in addition to multiple concerns pertaining to the methods of pregnancy ascertainment in Freundl 1998 (collected via a mailed, retrospective questionnaire), the calculation which generated a 0.7 perfect use Pearl Index [10] used an incorrect denominator which included all cycles, rather than only cycles of perfect use, which would inflate perfect use effectiveness rates [3].

Abbreviation
FABM: Fertility awareness-based method

Availability of data and materials
Data sharing is not applicable to this article as no datasets were generated or analyzed during the writing of this commentary.

Authors' contributions
CBP conceived of and wrote the commentary. The author read and approved the final manuscript.

Authors' information
CBP is a Senior Research Scientist at the Guttmacher Institute and holds an Associate affiliation in the Department of Epidemiology at the Johns Hopkins Bloomberg School of Public Health. Author affiliation included for informational purposes only, as this work was not conducted under the auspices of the Guttmacher Institute. The views expressed herein are those of the authors and do not necessarily reflect the views of the Guttmacher Institute.

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests. As described in the commentary, CBP reached out directly with the Director of Medical Affairs for Valley Electronics in multiple emails during September 2017 to directly raise concern regarding prior inaccurate analyses and their use in marketing materials for Valley Electronics products. Also as noted, CBP is participating in the conduct of a forthcoming systematic review on the effectiveness of fertility awareness-based methods for pregnancy prevention.

Publisher's Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References
1. Koch MC, Lermann J, van de Roemer N, Renner SK, Burghaus S, Hackl J, et al. Improving usability and pregnancy rates of a fertility monitor by an additional mobile application: results of a retrospective efficacy study of Daysy and DaysyView app. Reprod Health. 2018;15:37.
2. Committee on Publication Ethics Retraction Guidelines [Internet]. COPE; 2009. Available from: https://publicationethics.org/files/retraction%20guidelines_0.pdf. Accessed 30 Mar 2018.
3. Trussell J. Methodological pitfalls in the analysis of contraceptive failure. Stat Med. 1991;10:201–20.
4. Trussell J, Portman D. The creeping pearl: why has the rate of contraceptive failure increased in clinical trials of combined hormonal contraceptive pills? Contraception. 2013;88:604–10.
5. Dominik R, Trussell J, Walsh T. Failure rates among perfect users and during perfect use: a distinction that matters. Contraception. 1999;60:315–20.
6. Jones RK, Kost K. Underreporting of induced and spontaneous abortion in the United States: an analysis of the 2002 National Survey of family growth. Stud Fam Plan. 2007;38:187–97.
7. Jagannathan R. Relying on surveys to understand abortion behavior: some cautionary evidence. Am J Public Health. 2001;91:1825–31.
8. Klerman LV. The intendedness of pregnancy: a concept in transition. Matern Child Health J. 2000;4:155–62.
9. Peragallo Urrutia R, Polis CB, Jensen ET, Greene ME, Kennedy E, Stanford JB. Effectiveness of fertility awareness-based methods for pregnancy prevention: a systematic review. Obstet Gynecol. In press
10. Freundl G, Frank-Herrmann P, Godehardt E, Klemm R, Bachhofer M. Retrospective clinical trial of contraceptive effectiveness of the electronic fertility indicator Ladycomp/Babycomp. Adv Contracept. 1998;14:97–108.
11. Demianczyk A, Michaluk K. Evaluation of the effectiveness of selected natural fertility symptoms used for contraception: estimation of the pearl index of lady-comp, pearly and Daysy cycle computers based on 10 years of observation in the polish market. Ginekol Pol. 2016;87:793–7.
12. Polis CB. Email from Chelsea B. Polis to Niels van de Roemer, director. Medical Affairs. 2017;
13. Marcus A, Oransky I. Science publishing: The paper is not sacred. Nature. 2011;480:449–50.