



HAL
open science

Estimating Pregnancy Loss Trends: Important Work With Imperfect Data

Heini Väisänen, Hanna Remes

► **To cite this version:**

Heini Väisänen, Hanna Remes. Estimating Pregnancy Loss Trends: Important Work With Imperfect Data. Paediatric and Perinatal Epidemiology, 2025, Online First, <10.1111/ppe.70092>. <hal-05338753>

HAL Id: hal-05338753

<https://hal.science/hal-05338753v1>

Submitted on 30 Oct 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons CC BY 4.0 - Attribution - International License

Commentary

Estimating pregnancy loss trends: important work with imperfect data

Heini Väisänen¹ & Hanna Remes²

Affiliations

1. Institut National d'Etudes Démographiques (INED), Aubervilliers, France.
2. University of Helsinki, Helsinki, Finland.

Corresponding author

Hanna Remes (Email: hanna.remes@helsinki.fi)

Funding

The authors were funded by the European Union (ERC, SOC-MISC, Grant ID: 101077594). Views and opinions expressed are, however, those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

Disclosures

The authors declare no conflicts of interest.

ORCID

Heini Väisänen <https://orcid.org/0000-0002-5494-0415>

Hanna Remes <https://orcid.org/0000-0002-9291-703X>

Submitted versionⁱ: *this commentary has been published by Paediatric and Perinatal Epidemiology, see the version of record here: <https://doi.org/10.1111/ppe.70092>*

ⁱ Which corresponds exactly to the accepted version, as this type of text is not peer-reviewed.

Despite a longstanding interest in improving data quality and measurement on abortion, there is a lack of comprehensive methodological work on spontaneous pregnancy loss. Yet, reporting pregnancy losses in surveys could be affected by stigma and emotional distress¹, recall bias², varying likelihood of early pregnancy recognition among different population sub-groups,³ and reporting other pregnancy outcomes, such as induced abortions, as miscarriages.⁴ Studies are also often restricted by insufficient sample sizes, particularly among population sub-groups. While much of scientific research tends to focus on topics for which there is good-quality data, efforts to overcome data limitations in studying topics that are more difficult to measure are commendable. Currently, a better understanding of pregnancy loss, its risk factors and consequences, as well as monitoring of prevalence and population sub-group disparities, is ever more important as the provision of sexual and reproductive health services has become increasingly politicised, resulting in deteriorating care.⁵ Thus, the article by Forrest and colleagues⁶ in this issue of *Paediatric and Perinatal Epidemiology* is a welcome contribution by applying the enhanced Modified Kalman Filter (eMKF) to study pregnancy loss trends by age in metropolitan and non-metropolitan populations in the US. Research examining these trends by metropolitan vs. non-metropolitan context has not been possible previously due to small sample sizes in non-metropolitan populations. Some key challenges in any methodological work on pregnancy loss, in general, and in the study by Forrest et al.⁶, in particular, are discussed below.

1 Population at 'risk': who should we study?

When monitoring pregnancy loss trends and sub-group disparities, who should be considered at risk of pregnancy loss remains murky. On the one hand, one cannot experience a pregnancy loss unless pregnant, which could mean only including pregnant individuals in the denominator. On the other hand, pregnancy loss, miscarriage in particular, can be seen within a continuum of subfertility, in which case, including those aiming to conceive or at risk of pregnancy, remains relevant. However, we often lack detailed information on the respondents' pregnancies and their outcomes, sexual activity, pregnancy intentions, contraceptive use, and diagnosed or suspected fecundity problems, making it difficult to define the study population in this way. Yet,

differences in pregnancy loss risk by population sub-groups may depend on who is included in the denominator: a cohort study in the UK showed that when the reproductive lifespan of all women was studied, miscarriage risk differed by socioeconomic characteristics, but when only those who were pregnant were included, no such differences were found.⁷ Therefore, it is essential to clearly define who is included in a pregnancy loss study and provide a rationale for this inclusion.

Forrest et al.⁶ mention including ever pregnant women and studying women's risk of pregnancy loss, but their analyses are conducted at the pregnancy level. This can make it difficult to understand the trend estimates, since they reflect the share of pregnancy losses relative to live births (with induced abortions excluded and all spontaneous losses included in the outcome of interest), rather than any individual's or population sub-group's probability of experiencing a pregnancy loss. This approach does not account for the clustering of pregnancies at the individual level, which may result in, for example, artificially narrow confidence intervals or a disproportionate impact on the results from individuals with multiple pregnancy losses.

2 Which pregnancy experiences count?

Pregnancy losses may occur at different stages of gestation. However, there is no internationally agreed upon definition of miscarriage—and, consequently, stillbirth—as a function of gestational age.⁸ In addition, in rare cases people may not know whether they experienced an early miscarriage or an ectopic pregnancy, if the pregnancy ended before it was visible in an ultrasound. Forrest et al.⁶ overcome the issue by including miscarriages, ectopic pregnancies, and stillbirths under one term 'pregnancy loss'. While this approach avoids the problem that these outcomes can sometimes be difficult to distinguish, it may obscure differences in subpopulation risks of pregnancy loss if the risk factors of interest affect these outcomes in different ways. While stillbirth risk varies by socioeconomic and racial/ethnic characteristics,⁹ the few existing studies on miscarriage show mixed results,⁷ thus indicating that the risk may indeed vary depending on the specific type of pregnancy loss.

Forrest et al.⁶ test the robustness of their results by excluding pregnancies reported as unwanted from the analysis and including induced abortions in the denominator. The former criterion aims to solve the issue of non-planned pregnancies being typically recognised later, which could mean that some early miscarriages are ignored, whereas the latter approach better takes into account that some pregnancies ending in abortions would have been miscarried. While neither analysis dramatically changed the overall results,⁶ it demonstrated how impactful such analytical decisions can be among population sub-groups. The disproportionately high proportion of pregnancy loss among the youngest age group in the main analyses seemed more reasonable after abortions were included, indicating that a relatively low proportion of pregnancies among this age group ended in live births due to a relatively high prevalence of induced abortion. Following from the argument that those intending a pregnancy are more aware of a pregnancy loss than those experiencing an unwanted, or ‘mistimed’ pregnancy, an interesting additional analysis could have focused on intended pregnancies only.

Overall, any analysis of pregnancy loss using self-reported data will include some spillovers from one pregnancy category to another, whether it is fluctuations between abortions and miscarriages⁴, or different types of spontaneous pregnancy loss. Thus, separating different pregnancy outcomes and including all pregnancies (also abortions) in the denominator whenever possible, given the overlap and misclassification between different categories, is the recommended approach.

3 Potential pitfalls of model-based estimation: a need for multiple approaches

Any model-based approach to estimate pregnancy loss is a simplification of reality based on imperfect data. When it comes to such limitations of their method, Forrest et al.⁶ mentioned for instance, not being able to adjust their results for other sociodemographic characteristics, since the eMKF method only allows two stratifying factors (in this case, metropolitan status and age), and discuss some of the technical details of the modelling approach. However, given that one of the aims of the paper was to test the feasibility of using the eMKF method for analysing pregnancy loss, a more comprehensive discussion of the method could have been warranted,

especially focusing on how the limitations of the model may affect the results and why. Moreover, while smoothing out random variation, the eMKF model can explore trends and differences between sub-population groups when only limited sample sizes are available; the model does not overcome the issues with potential erroneous reporting of miscarriage in surveys or clustering of pregnancy experiences within individuals. Exploring multiple analysis approaches in future studies and discussing their implications for the results is key to a more comprehensive understanding of pregnancy loss and its risk factors.

The analytic sample used in Forrest et al.⁶ only included pregnancies up to the year 2018 and thus precedes the Dobbs decision overturning the right to have an abortion before foetal viability in the US. Future studies should examine how the landscape of pregnancy loss and reporting in surveys is changing and its impacts on population health monitoring. Poorer access to obstetrics and gynaecology care, especially in rural areas and states with strict abortion restrictions,⁵ is likely to increase the probability of spontaneous pregnancy loss and perinatal death, as well as maternal health complications, making it important to monitor future trends. On the reporting side, it could lead, for instance, to more induced abortions being reported as spontaneous pregnancy losses, which makes said monitoring more complicated. In addition, future studies should investigate the specific factors behind the increasing trend of pregnancy losses in the US, as demonstrated by Forrest et al⁶, because policy implications differ based on the underlying reasons. For example, if miscarriages increase due to the effects of climate change or pollution, the implications are different than if they seem to increase due to changes in reporting behaviours. Therefore, innovative approaches, including model-based estimates, are needed to better understand this pressing public health issue.

References:

1. Bardos J, Hercz D, Friedenthal J, Missmer SA, Williams Z. A National Survey on Public Perceptions of Miscarriage. *Obstetrics & Gynecology*. 2015;125(6):1313-1320. doi:10.1097/AOG.0000000000000859
2. Lang K, Nuevo-Chiquero A. Trends in Self-reported Spontaneous Abortions: 1970–2000. *Demography*. 2012;49(3):989-1009. doi:10.1007/s13524-012-0113-0
3. Watson K, Angelotta C. The frequency of pregnancy recognition across the gestational spectrum and its consequences in the United States. *Perspect Sexual Reproductive*. 2022;54(2):32-37. doi:10.1363/psrh.12192
4. VandeVusse AJ, Mueller J, Kirstein M, Strong J, Lindberg LD. “Technically an abortion”: Understanding perceptions and definitions of abortion in the United States. *Soc Sci Med*. 2023;335:116216. doi:10.1016/j.socscimed.2023.116216
5. Nandi P, Roncari DM, Werner EF, Gilbert AL, Ramos SZ. Navigating Miscarriage Management Post-Dobbs: Health Risks and Ethical Dilemmas. *Womens Health Issues*. 2024;34(5):449-454. doi:10.1016/j.whi.2024.05.004
6. Forrest SE, Rossen LM, Ahrens KA. Trends in risk of pregnancy loss among U.S. women by metropolitan status, 2000-2018. *Paediatric and Perinatal Epidemiology*. Published online 2025. doi:10.1111/ppe.70066
7. Väisänen H, Keenan K. Measuring social inequalities in self-reported miscarriage experiences in the United Kingdom. *J Crit Public Health*. Published online March 26, 2025. doi:10.55016/ojs/jcph.vi.79555
8. Prior M, Bagness C, Brewin J, et al. Priorities for research in miscarriage: a priority setting partnership between people affected by miscarriage and professionals following the James Lind Alliance methodology. *BMJ Open*. 2017;7(8):e016571. doi:10.1136/bmjopen-2017-016571
9. Adams A, Dongarwar D, Shay L, et al. Social Determinants of Health and Risk of Stillbirth in the United States. *Am J Perinatol*. 2024;41(S 1):e477-e485. doi:10.1055/s-0042-1756141