BJPsych The British Journal of Psychiatry (2025)
226, 6, 334–336. doi: 10.1192/bjp.2025.119

Guest Editorial

Methodological obstacles in studies linking hormonal contraception and depression

Søren Vinther Larsen, Vibe G. Frokjaer and Brice Ozenne

Summary

This editorial describes common methodological obstacles, including the healthy user bias, that are encountered in research studying the link between hormonal contraceptive use and depression and provides considerations for future directions.

Keywords

Hormonal contraceptive; depressive disorder; mood disorder; observational study; statistical methodology.

Copyright and usage

© The Author(s), 2025. Published by Cambridge University Press on behalf of Royal College of Psychiatrists.

> 'If the randomized trial data were also incorrectly analyzed by deleting data from the early follow-up, an apparently beneficial estimate of hormone therapy would similarly have been found.'[1]

With more than 250 million hormonal contraceptive users globally, the relationship between hormonal contraceptive use and mental well-being is a subject of growing interest and debate. Mood symptoms significantly contribute to early discontinuation; as many as 30% of oral contraceptive users discontinue within the first 6 months, with a considerable portion attributing their discontinuation to mood-related side symptoms. This creates concern about the implications for mental health and is also a significant barrier in many studies aiming to unravel potential links between hormonal contraceptive use and mood disorders. We claim that this is likely to be a critical source of inconsistency in the current literature. In this editorial, we discuss the methodological obstacles faced in attempts to study the relationship between hormonal contraceptive use and depression, while highlighting reflections for future directions.

Methodological obstacles

No strong evidence exists to support a causal link between hormonal contraceptive use and depression on the basis of the few randomised studies conducted, although these studies have provided evidence coherent with mood-deteriorating effects. Randomised trials face significant challenges, including maintaining blinding and obtaining sufficient sample sizes and follow-up times to assess long-term safety with respect to rare outcomes. Another major limitation is the failure to include hormonal contraceptive-naive individuals. As many as 80% of women (in terms of individuals assigned female at birth) in developed countries have used hormonal contraception at age 20 years. Consequently, women who tolerate hormonal contraception well are more inclined to start on hormonal contraception again and are thus more likely to participate in randomised trials. This limits the generalisability of the results, especially regarding the frequency of the outcome, to the general population. All these challenges mean that observational studies provide a pragmatic alternative.

Over the past decade, access to national healthcare data has enabled investigation of the link between hormonal contraception and depression at a country population level with several years of follow-up. However, studies have shown mixed and even contradictory findings. Many cohort studies aim to replicate a hypothetical randomised trial to address the causal question of interest. Although this approach is promising, cohort studies, especially if retrospective, face critical methodological obstacles (beyond lack of randomisation) that need to be carefully considered and discussed when drawing causal conclusions. On the basis of the emulated target trial framework,[1] we summarise the key methodological deviations from the target trial in terms of five obstacles that are prone to causing bias and are potential sources of the conflicting findings in the literature (Fig. 1). In particular, the 'healthy user bias' is likely to occur in hormonal contraceptive studies owing to (a) high rates of mood-related discontinuation and (b) depletion of individuals most susceptible to developing depression, leading to a potentially selected sample under study.

Obstacle 1: definition of time zero

In randomised studies, the follow-up time often starts at a time point (i.e. time zero) aligned with intervention allocation and specification of participant eligibility. Retrospective cohort studies often lack this alignment owing to the lack of an obvious time zero for the control group and the limited ability of healthcare data to capture the time of drug initiation. The former limitation can be handled by using a comparator design, if applicable; for example, we previously compared the risk of depression among first-time users of three different intrauterine devices.[2] The latter limitation means that only a subset of a cohort has complete follow-up times; thus, to preserve power, researchers often choose to also include prevalent users, that is, participants who initiated hormonal contraception a long time before time zero. This 'prevalent user' design mimics a target trial that omits the first part of the follow-up and includes only those still eligible after a certain time after time zero.[1] For example, the Royal College of General Practitioners' Oral Contraception Study in the UK included 62% prevalent users of oral contraception, with one-third having used it for more than 2 years.[3] However, this design is prone to the 'healthy user bias'; this is a type of selection bias[4] in which susceptible individuals are excluded or overrepresented in the control group, as they are likely to have developed depression or discontinued contraceptive owing to mood-related symptoms before the time of inclusion. The prevalent user design can result in contradictory findings compared with those of randomised trials and emulated target trials. For instance, two Swedish register-based studies yielded conflicting findings: one, using the prevalent user design, found similar incidence rates in individuals using combined oral contraceptives (risk ratio 1.01, 95% CI: 0.98–1.02);[5] whereas the other, emulating a target trial, showed an increased instantaneous risk compared with never-users (hazard ratio 1.36, 95% CI: 1.34–1.38).[6]

https://doi.org/10.1192/bjp.2025.119 Published online by Cambridge University Press

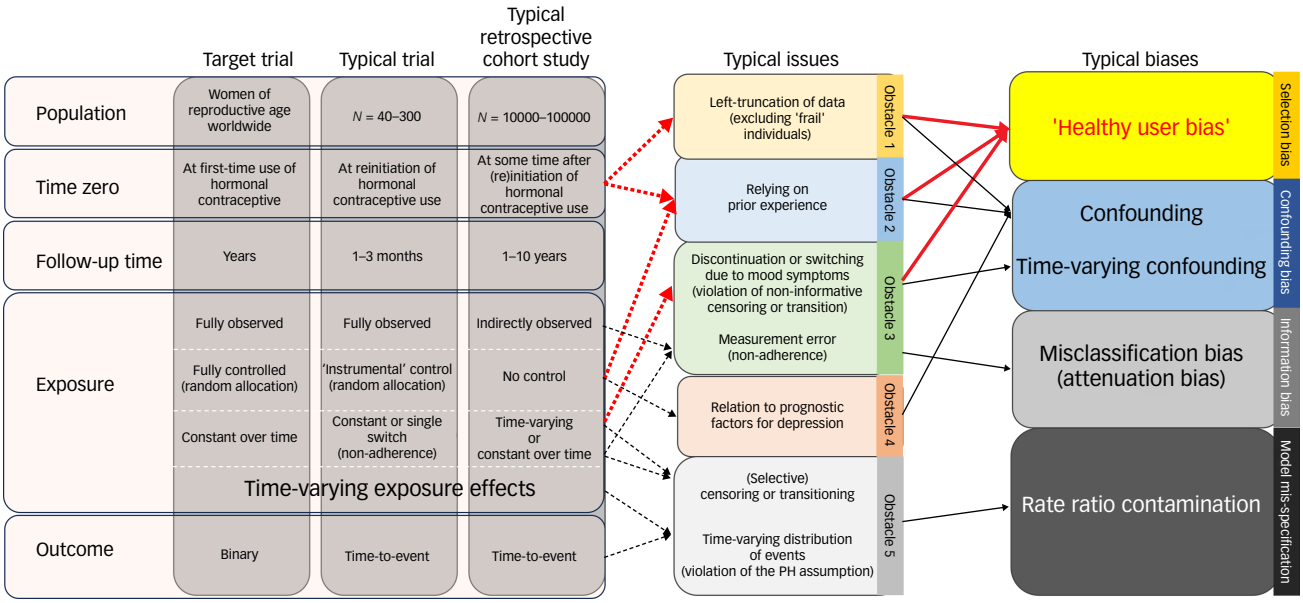


Fig. 1 The comparison between the target trial, the typical trial and the typical retrospective cohort study investigating the relationship between hormonal contraceptive use and depression. It highlights the typical issues observed in the typical retrospective cohort study due to methodological design deviations from the target trial (dashed line arrows) and the consequent typical biases (solid line arrows) with a special emphasis on 'the healthy user bias' (bold arrows). PH, proportional hazard.

Obstacle 2: handling of previous treatment

Earlier exposure to hormonal contraceptive use can affect the decision to use hormonal contraception. This can also give rise to confounding and healthy user bias, as individuals who discontinued before time zero owing to mood-related symptoms are less likely to start again. Thus, even with careful efforts to avoid the prevalent user design, estimates can also be biased if prior hormonal contraceptive use is not accounted for. This can be mitigated by excluding prior users or statistically, e.g. via stratification. When prior users were included in the two Swedish studies, combined oral contraceptive use became associated with a protective effect (risk ratio 0.89, 95% CI: 0.87–0.91) in the study with the prevalent user design, and in the emulated target trial it actually changed from being harmful (hazard ratio 1.22, 95% CI: 1.19–1.24) to being protective (hazard ratio 0.92, 95% CI: 0.91–0.94).[6] Thus, such design discrepancies can effectively lead to contradictory conclusions.

Obstacle 3: handling of non-adherence

Treatment transitions (starting, stopping or switching hormonal contraceptive type) may also occur during follow-up. Many cohort studies mitigate this by allowing individuals to transition between treatment arms; however, this may introduce time-varying confounding and selection bias (including the healthy user bias) when it is related to mood symptoms, e.g. if mood symptoms lead to hormonal contraceptive discontinuation. The Royal College of General Practitioners' Oral Contraception Study showed that 27% discontinue within the first year despite continued contraceptive need, 8% owing to side effects and 40% owing to intercurrent morbidity, including 3% because of depression.[3] Accordingly, former users had higher risks of depression (risk ratio 1.26, $P < 0.01$) and hospital admission with depression (risk ratio 2.76, $P < 0.01$) compared with never-users. Correspondingly, a UK Biobank study showed that recent discontinuation of hormonal contraceptive use was associated with an increased risk of depression (hazard ratio 1.17, 95% CI: 1.08–1.27).[7] This indicates that treatment transition is probably not independent of the probability of developing depression, i.e. it violates the assumption of non-informative transition. Similarly, censoring individuals at treatment transition could also introduce bias and violate the assumption of non-informative censoring. On the other hand, in an intention-to-treat analysis in which treatment transitions are ignored, non-adherence can lead to misclassification bias, which is non-differential if unrelated (resulting in attenuation bias, a major problem in equivalence or non-inferiority trials) and differential if related to the subsequent risk of the outcome (e.g. if women discontinue owing to mood symptoms). Adherence-adjusted effect estimates using inverse-probability-of-treatment or censoring weighting could mitigate these issues; however, this requires time-dependent information on depression risk factors that predict non-adherence, which is often not available.

Obstacle 4: handling of confounding

A limitation of observational studies is the lack of randomisation and hence the risk of confounding; as comparison groups are self-selected by the decision to initiate hormonal contraception, they may differ in terms of various predisposing factors for depression. As residual confounding will always be a matter for discussion in observational studies, use of strategies to assess whether residual confounding is likely to explain an association is warranted, for instance, use of negative control outcomes or sibling analyses to approximate potential unmeasured familial confounding.[7]

An attempt should be made to account for confounders at the time of treatment initiation, including confounders related to prior use (e.g. confounding by indication or prior mental disorder). However, in studies with a prevalent user design, in which time zero is not aligned with the time of treatment initiation, potential prognostic factors (e.g. anxiety symptoms at time zero) can be intermediate variables, i.e. developed owing to the treatment, making adjustment inappropriate. On the other hand, failing to account for potential healthy user bias can also induce confounding, e.g. if treatment-related anxiety resulted in treatment

https://doi.org/10.1192/bjp.2025.119 Published online by Cambridge University Press

discontinuation and avoidance of later reinitiation. This dilemma can be avoided when time zero is synchronised with the time of allocation, i.e. the new-user design.[3] Nevertheless, confounding may still occur during follow-up owing to time-varying exposure (e.g. if treatment-induced anxiety resulted in treatment discontinuation). When this is the case, traditional methods such as Cox regression or time-dependent Cox regression are not suitable for handling time-varying confounding; instead, dedicated methods should be used, e.g. the inverse-probability-of-treatment weighting. The intention is to reweight the users and non-users in a time-varying fashion such that their covariate distributions are comparable.[8]

Obstacle 5: handling of time-varying effects

The Cox model is commonly used to summarise a treatment effect over a period of time into a single number, the hazard ratio. In the absence of censoring, the hazard ratio can be thought of as a weighted average of the risk ratio over time, with weights proportional to the number of people at risk. However, if the effect varies with duration of use, the estimate is very much dependent on the duration of the follow-up period.[9] For example, if the risk of depression is high during early follow-up and low or even diminished at late follow-up, the estimated risk will be averaged over the whole follow-up period, potentially masking noteworthy effects. In the presence of censoring, the weights also depend on the censoring distribution, making the hazard ratio difficult to understand and not comparable across studies unless the contraceptive effect is constant over time. Even in this case, the measured effect may still be time-varying if susceptible individuals tend to discontinue over time owing to mood-related symptoms. Such time-varying effects may especially lead to biases if time zero has not been aligned with the treatment allocation, as the initial treatment phase is ignored. Modelling time-varying treatment effects is possible but challenging; for example, it requires time zero to be aligned with treatment allocation, as well as careful interpretation. Indeed, unmodelled patient heterogeneity (frailty) can be mistaken for apparent attenuation of the treatment effect over time. This is why reporting on a risk scale (e.g. the difference in 1-year risks) is generally recommended over use of hazard ratios. Alternatively, the use of time-restricted analyses with increasing longer follow-up times (not to be interchanged with period-specific effects, as this would be susceptible to built-in selection bias) has been attempted.[9] The hazard ratio increased from the first month of treatment and peaked 6 months after initiation, with a hazard ratio of 1.58 (95% CI: 1.54–1.62), whereafter it started to decrease, although the hazard ratio remained elevated overall with follow-up time restricted to 2 years.[6] Time-restricted analyses are especially informative with regard to early effects; however, long-term restricted analysis would still average effects over a long period of time, so baseline-adjusted survival curves will be more informative with regard to late effects,.[9]

In summary, both randomised and observational studies face challenges in evaluating the causal link between hormonal contraceptive use and depression, especially selection bias and non-adherence or switch of treatment, which gives rise to the healthy user bias. Other challenge that were not discussed in this editorial but complicate this research area further include the potential presence of effect heterogeneity, in which certain subgroups are more susceptible to mood-related side effects, whereas others may experience beneficial effects on mood. Considering the potential impact on women's mental health and the consequences of non-adherence due to mood symptoms, such as unintended pregnancies, it is essential to mitigate these methodological challenges in future research by minimising bias through careful study design and analytical approaches.

Søren Vinther Larsen, Neurobiology Research Unit, Rigshospitalet, Copenhagen, Denmark; **Vibe G. Frokjaer**, Neurobiology Research Unit, Rigshospitalet, Copenhagen, Denmark; Department of Psychiatry, Psychiatric Centre Copenhagen, Copenhagen, Denmark; and Department of Clinical Medicine, University of Copenhagen, Copenhagen, Denmark; **Brice Ozenne**, Neurobiology Research Unit, Rigshospitalet, Copenhagen, Denmark; and Department of Public Health, Section of Biostatistics, University of Copenhagen, Copenhagen, Denmark

Correspondence: Søren Vinther Larsen. Email: soren.vinther@nru.dk

First received 13 Jan 2025, final revision 31 Mar 2025, accepted 2 Apr 2025, first published online 11 Jun 2025

Data availability

Data availability is not applicable to this article as no new data were created or analysed in this study.

Acknowledgement

We thank Professor Per Kragh Andersen of the Section of Biostatistics, Department of Public Health, University of Copenhagen, for his insights and feedback on an earlier version of the manuscript. He received no compensation for his role in the study.

Author contributions

S.V.L., V.G.F. and B.O. contributed to concept development and to drafting and critical revision of the manuscript for important intellectual content. S.V.L., V.G.F. and B.O. approved the final draft.

Funding

S.V.L. is funded by a Lundbeck Foundation Early-Career Clinician Scientist grant on the project titled 'Reproductive mental health: Optimizing treatment of depression in women using oral contraceptive' (grant identifier: R450-2023-1488, received in 2024). The grant includes salary payment and operational costs. The funder had no role in the design or conduct of the study; preparation, review or approval of the manuscript; or the decision to submit the manuscript for publication.

Declaration of interest

V.G.F. has received honoraria for lectures for Lundbeck Pharma A/S, Janssen-Cilag A/S, Gedeon-Richter A/S and Ferring Pharmaceuticals A/S.

References

1 Hernán MA, Wang W, Leaf DE. Target trial emulation: a framework for causal inference from observational data. *JAMA* 2022; **328**: 2446–7.

2 Larsen SV, Mikkelsen AP, Ozenne B, Munk-Olsen T, Lidegaard Ø, Frokjaer VG. Association between intrauterine system hormone dosage and depression risk. *Am J Psychiatry* 2024; **181**: 834–41.

3 Royal College of General Practitioners. *Oral Contraceptives and Health: An Interim Report from the Oral Contraception Study of the Royal College of General Practitioners*. Pitman, 1974.

4 Ray WA. Evaluating medication effects outside of clinical trials: new-user designs. *Am J Epidemiol* 2003; **158**: 915–20.

5 Lundin C, Wikman A, Lampa E, Bixo M, Gemzell-Danielsson K, Wikman P, et al. There is no association between combined oral hormonal contraceptives and depression: a Swedish register-based cohort study. *BJOG* 2022; **129**: 917–25.

6 Johansson T, Karlsson T, Lidegaard Ø, Johansson Å. Contemporary hormonal contraception and risk of depression – a Swedish nationwide register-based emulated target trial. To be published in *Lancet* [Preprint] 2024. Available from: https://ssrn.com/abstract=4709955.

7 Johansson T, Vinther Larsen S, Bui M, Ek WE, Karlsson T, Johansson Å. Population-based cohort study of oral contraceptive use and risk of depression. *Epidemiol Psychiatr Sci* 2023; **32**: e39.

8 Mansournia MA, Etminan M, Danaei G, Kaufman JS, Collins G. Handling time varying confounding in observational research. *BMJ* 2017; **359**: j4587.

9 Hernán MA. The hazards of hazard ratios. *Epidemiology* 2010; **21**: 13–5.

https://doi.org/10.1192/bjp.2025.119 Published online by Cambridge University Press