



Combined oral contraceptives, venous thromboembolism, and the problem of interpreting large but incomplete datasets

Jürgen Dinger,¹ Samuel Shapiro²

¹Director, ZEG - Berlin Center for Epidemiology and Health Research, Berlin, Germany
²Visting Professor of Epidemiology, Department of Epidemiology, University of Cape Town, Cape Town, South Africa

Correspondence to

Dr Jürgen Dinger, ZEG - Berlin Center for Epidemiology and Health Research, Invalidenstrasse 115, 10115 Berlin, Germany; dinger@zeg-berlin.de

Received 11 November 2011
 Accepted 14 November 2011
 Published Online First
 29 November 2011

Background

In 2009, Lidegaard *et al.*¹ published findings in the *British Medical Journal*, derived from a Danish retrospective cohort study of the risk of venous thromboembolism (VTE) associated with the use of combined oral contraceptives (COCs). Their analysis was based on data derived from national health registries, and they concluded that “oral contraceptives with desogestrel, gestodene, or drospirenone were associated with a significantly higher risk of VTE than oral contraceptives with levonorgestrel”. That report has previously been reviewed in this Journal² and at an international workshop.³

Subsequently, because of methodological limitations in the Danish study, the European Medicines Agency (EMA) requested a re-analysis. The conduct of the re-analysis was overseen by an independent three-member steering committee [KJ Rothman (Chairman), FE Skjeldestad (nominated by Lidegaard, and a co-author of the published re-analysis)⁴ and S Shapiro (nominated by Bayer Schering, and a co-author of this commentary)].

The completed re-analysis, together with a commentary by the steering committee, an additional commentary by S Shapiro, and an audit requested by the steering committee, was submitted to the EMA, and an abbreviated version of the EMA submission has recently been published.⁴ Here we review the publication of the re-analysis submitted to the EMA. The publication manuscript was not submitted to the steering committee for review before publication.

In the published re-analysis⁴ the authors claim to have confirmed the conclusions reached in the original analysis:¹ if anything, the risk estimates were even higher

than the original estimates. A difficulty, however, is that several numbers reported in the publication differ from those mentioned in the re-analysis submitted to EMA (one example is given below).

Since the mid-1990s there has been heated debate regarding the risk of VTE associated with the use of different progestogens, and those who have followed the discussion can only note with concern its confrontational and increasingly sharp tone, which, unfortunately, is also reflected in the published responses to the re-analysis,^{5–7} and more particularly in the authors' replies.^{8,9}

The heat of the debate may have something to do with the massive number of pending claims for compensation filed against the manufacturers in the USA by users of newer COCs. The potential sums involved could cover the annual budget of a small country. The scientists involved are being subjected to pressure from the media, from the plaintiffs and their attorneys, as well as from the affected manufacturers. In this atmosphere, nearly everyone – ourselves included – is confronted with insinuations of a conflict of interest. And the effects of a debate conducted in such a highly publicised, polemical, and *ad hominem* manner can be devastating. In the course of the ‘pill scare’ of the mid-1990s, for example, the use of all oral contraceptives declined, and there was a substantial increase in the incidence of abortions.¹⁰ Here we wish, if possible, to avoid unnecessary escalation, and to confine ourselves to a consideration of relevant facts and methodological concerns.

The results of the Danish re-analysis will be viewed by its authors, by the public, and presumably by some specialists as well,

as evidence of progestogen-related differences among COCs with regard to the risk of VTE. We question this view for a number of reasons. Below, in addressing the question of whether the re-analysis shows sufficient evidence for differential effects of progestogens on the risk of VTE, we limit ourselves to three areas of concern.

Relevance and presentation of different analyses

One important objective of the Danish re-analysis was to account for bias due to different times and durations of market availability of different progestogens, particularly drospirenone (DRSP) and levonorgestrel (LNG). The use of DRSP could only have commenced in 2001, the year in which it was introduced in Denmark. By contrast, LNG was introduced decades before the Danish registry of Medicinal Products was established in 1994. For symmetry, the analysis should therefore have excluded women whose use of LNG commenced before 2001: in this subgroup VTE-susceptible subjects could have been depleted. It is important to note that this susceptibility reflects not only genetic risk, but also long-standing predisposing conditions such as obesity or family history, risks associated with lifestyle, working conditions and personal circumstances.

For these reasons it was stipulated, in advance, that a comparison of women who had never previously used a COC (first-time users) and who also started COC use during the same time interval (i.e. not before 2001) would be the most valid comparison, provided sufficiently large cohorts could be accrued and compared. This goal was partially acknowledged in the report submitted to the EMA.

The focus on first-time use (or at least, the best approximation to it: for example, women who had not used a COC before 2001 according to the information available in the registry being designated as starters) becomes even more important in view of the limited validity of data, and of various sources of bias and confounding considered below. It is likely that correctly specified starter cohorts of COC users, in each of which new and established progestogens are in use, would be more similar with regard to the prevalence of prognostic factors compared to non-starters. Therefore, a lack of information on relevant prognostic factors would probably have less impact on the risk estimates. In addition, bias caused by the limited validity of information on outcomes and exposure (see below) would probably be less marked in this setting.

In line with these requirements, in one analysis submitted to the steering committee, women who first used a COC from 2001 onwards, and who were not recorded in the Danish registry as having used any COC before between 1995 (the registry was initiated in 1994) and 2001, were compared. Among DRSP and LNG users there were 60 and 11 confirmed cases of VTE (Table 16 of Analysis 3 in the EMA report), and the adjusted relative risk was 1.0 [upper 95% confidence limit approximately 2 (our calculation)].

The number of events was not large, but a ratio of 2 between the relative risk and its upper 95% confidence limit was still statistically meaningful, and it did not suggest a difference. Surprisingly, the results of this analysis, specified and agreed to *a priori*, were not reported in the published re-analysis. That finding should have been communicated and discussed, so that readers could have come to their own conclusions.⁷

Validity of data

(1) The independent audit requested by the steering committee revealed, among other things, a lack of the following: a formal *a priori* statistical analysis plan; standard operating procedures; quality control of statistical programming; and documentation of programming and analyses. Transparency and traceability of the analyses, an important standard for clinical and epidemiological studies of regulatory relevance, was not met in the re-analysis.

(2) The analysis and re-analysis of the compared cohorts in the Danish Cohort Study were based on incidence rates of VTE associated with the use of COCs containing different progestogens. For this purpose the number of VTEs for each COC cohort was needed, and the validity of the study was crucially dependent on the validity of that information.

Data on VTE were obtained from the Danish registry of patients. In the original publication¹ the investigators estimated that 10% of the diagnoses “were uncertain”. It was difficult to reconcile this estimate with the results of another Danish Cohort Study of cases of VTE over the age of 50 years, conducted by Severinsen and her colleagues.^{11–13} In her study, a review of medical records revealed that the registry diagnoses of VTE were incorrect in 25% of cases diagnosed in hospital wards, and in 69% of cases diagnosed in emergency departments; the latter cases constituted 41% of the total.

The principal investigator informed the steering committee that (a) VTE diagnoses from emergency departments were not used and (b) that the remaining difference between the studies could be explained by the age difference of the study populations. In addition, a complete chart review to establish false-positive as well as false-negative diagnoses in the registry was not possible due to Danish legislation. The steering committee therefore requested that the proportion of false-positive cases should be estimated, based on a sample of 200 registry-recorded diagnoses of VTE. When this was done, in line with the results of Severinsen’s study, the validation study showed that 26% of the registry-recorded diagnoses, not 10% as previously claimed, were false-positives. Consequently, in order to reduce the impact of false-positive diagnoses it was agreed that a ‘confirmed’ diagnosis would be defined as a woman with registry-recorded VTE who received anti-coagulants for at least 4 weeks; for that definition the proportion of false-positives was small. However, the information on VTE in the study was still inaccurate.

First, the exclusion of emergency room diagnoses and the exclusion of hospital ward diagnoses without registry-documented anticoagulatory treatment of 4 weeks led to under-ascertainment of VTE. Second, there was disproportionate exclusion of potential VTE cases among the compared cohorts, as shown by the ratio of confirmed to unconfirmed diagnoses of VTE. Based on data in Table 3 of the re-analysis publication⁴ the ratio was 1.2 for non-users of COCs, 2.8 for users of COCs with LNG plus 30–40 μg ethinylestradiol (EE), and 5.1 for COCs with desogestrel plus 30–40 μg EE (the highest value). Thus it is likely that knowledge of the exposure influenced the diagnosis of VTE, and its treatment with anticoagulants, and the markedly different ratios make it likely that there was substantial bias.

It is also striking that the numbers in the published re-analysis⁴ are not the same as the numbers reported to the EMA. In Table 3 of the publication⁴ the ratios of confirmed to unconfirmed diagnoses for DRSP- and LNG-containing preparations were almost identical (DRSP 2.8 and LNG 2.8). Yet in Table 10 of the EMA report, although the number of VTEs was the same, the ratios differed markedly: 1.5 for DRSP and 0.4 for LNG. The differences indicate the potential for substantial bias in making the diagnosis of VTE, conditional on the specific COC used; and the differences between the data in the publication and in the EMA report illustrate the lack of adequate quality assurance. Overall the definition of VTE was unreliable.

(3) To ensure valid comparisons of different COCs precise information on the timing and duration of use for each of the compared cohorts was essential. It has been well established in multiple studies^{14–18} that the use of progestogen/estrogen combinations (for hormone replacement therapy as well as oral contraception) is associated with a high increase in the risk of VTE during the first months after starting or restarting use, and that with continued use the risk declines. For a valid comparison of risk across different COC groups exact information on starting and stopping dates was needed. However, this information was not available in the Danish registry. Having a prescription filled did not mean that the product was actually taken; alternatively, some women may have commenced use weeks or months after the filling the prescription because they first wanted to use up their old prescriptions. Hence, the duration of exposure at the time of VTE cannot be verified based solely on the registry data.

In the original publication¹ only DRSP-containing COCs showed a substantial ‘early use effect’, whereas LNG-containing preparations did not. In addition, the overall ‘early use effect’ for COCs was much less pronounced in the original analysis, as compared with what has been published in the literature, and in an earlier case-control study,¹⁹ also based on VTE

diagnoses from the Danish health register. That study, however, did not rely on registry-recorded COC use, but on information provided by the patients themselves.

In the re-analysis the investigators stipulated rules based on several assumptions, in order to minimise the impact of the lack of precise exposure information. Using these rules a moderate ‘early use effect’ could be demonstrated for LNG, but surprisingly, at the same time the ‘early use effect’, previously present in DRSP users,¹ had now disappeared. This striking change raises doubts about the validity of the assumptions made in the re-analysis. The strong likelihood is that the rules and imputations were not a useful substitute for the missing information. It is doubtful whether an analysis that fails to show the typical duration-related pattern of VTE risk associated with COC use can be considered to be a reliable source for identifying differences in risks of low magnitude, and close to the resolving power of the ‘epidemiological microscope’ discrimination among bias, confounding and causation is virtually impossible.^{20,21}

Risk profiles of COC user cohorts

In clinical practice age, body mass index (BMI) and family history of VTE are important determinants of VTE risk;^{18,22} it is likely that they are also determinants of COC use, as well as determinants of the particular COC used. In the Danish database there was no information on BMI or family history of VTE. The principal investigator acknowledged the important role of these factors in general, and pointed out, correctly, that a risk factor is not necessarily a confounder.¹⁵ However, his statement that so far no study has demonstrated “any confounding influence from BMI”¹⁸ is not correct.

The European Active Surveillance (EURAS) study^{17,23} showed that obesity was more common among users of DRSP-containing COCs compared to users of LNG-containing COCs. Overall, the prevalence of obesity among DRSP users was 1.6 and 1.8-fold higher compared to users of LNG and other progestogens, respectively. This is not surprising as weight gain is a common concern among users of oral contraceptives, and the antimineralocorticoid effect of DRSP counteracts water retention induced by EE. The EURAS analyses also identified more than additive effects of age and BMI (interaction). Overall, adjustment for age, BMI, duration of current use, family history of VTE and the interaction between age and BMI reduced the VTE relative risk by 27%, as compared with an analysis that adjusted only for age.

In the Danish study the age differences between the cohorts were more pronounced than in the EURAS study:¹⁷ fully 74.2% of the DRSP users were 15–29 years of age, whereas the majority (57.4%) of the LNG users were 30–49 years old. This striking difference adds to the concern that adjustment for confounding,

based on data with missing information on major confounders, cannot compensate for differences between these obviously different user populations. It is conceivable, even likely, that adjustment for BMI, age/BMI interaction, family history of VTE, and duration of current COC use (with exact data) would have had an even stronger impact on the risk estimates than in the EURAS study.

Finally, the main emphasis in the re-analysis, as submitted to the EMA, and as published,⁴ was placed on comparisons of users of DRSP who could not have commenced use before 2001, with users of LNG who could have commenced use as early as 1994, or earlier, before the Danish registry existed. In a properly designed cohort study contemporaneous time intervals need to be compared in order to minimise bias and confounding, and the analyses based on non-contemporaneous intervals were not valid.

Conclusions

To sum up: the basic information needed for the comparison of VTE risk across cohorts, which includes the number of events, overall exposure and time pattern of exposure, was not valid. In addition, the re-analysis lacked transparency. The manner in which the analyses were conducted was inadequate, as were the measures to ensure adequate quality control. Adjustment for major potential confounders was not possible because of missing data.

The study population was massive, involving millions of women, and the reported confidence intervals (CIs) around the relative risk estimates were narrow. However, confidence limits only allow for random variation. They do not allow for systematic errors due to bias or confounding. If bias or confounding is present, as in the Danish analysis,¹ and re-analysis,⁴ it can readily overwhelm any statistical variation,²⁴ and the CIs were misleading.

We conclude that the best evidence continues to suggest that the increased risk of VTE among COC users is a class effect. In the Danish data an analysis confined to women who used COCs for the first time from 2001 onward did not support any differential effects of progestogens. Surprisingly, this information was neither presented nor discussed in the published re-analysis.⁴ Any potential differences, if they exist at all, are probably beyond the resolving power of the 'epidemiological microscope'.

Competing interests Jürgen Dinger was previously an employee of Schering until 2004. He presently consults, and in the past has consulted, with manufacturers of products discussed in this article. Samuel Shapiro presently consults, and in the past has consulted, with manufacturers of products discussed in this article.

Provenance and peer review Commissioned; externally peer reviewed.

References

- 1 Lidegaard Ø, Løkkegaard E, Svendsen AL, *et al.* Hormonal contraception and risk of venous thromboembolism: national follow-up study. *BMJ* 2009;339:b2890.
- 2 Shapiro S, Dinger J. Risk of venous thromboembolism among users of oral contraceptives: a review of two recently published studies. *J Fam Plann Reprod Health Care* 2010;36:33–38.
- 3 Reid RL, Westhoff C, Mansour D, *et al.* Oral contraceptives and venous thromboembolism consensus opinion from an international workshop held in Berlin, Germany in December 2009. *J Fam Plann Reprod Health Care* 2010;36:117–122.
- 4 Lidegaard O, Nielsen LH, Skovlund CW, *et al.* Risk of venous thromboembolism from use of oral contraceptives containing different progestogens and oestrogen doses: Danish cohort study, 2001–9. *BMJ* 2011;343:d6423.
- 5 Szarewski A, Mansour D. Letter to the editor. Risk of VTE with COCs containing different progestogens. <http://www.bmj.com/content/343/bmj.d6423?tab=responses> [accessed 13 November 2011].
- 6 Shapiro S. Letter to the editor. Re: Lidegaard *et al.* Risk of venous thromboembolism from use of oral contraceptives containing different progestogens and oestrogen doses: Danish cohort study 2001–9. <http://www.bmj.com/content/343/bmj.d6423?tab=responses> [accessed 13 November 2011].
- 7 Pauls AR. Letter to the editor. Re: Risk of venous thromboembolism from use of oral contraceptives containing different progestogens and oestrogen doses: Danish cohort study, 2001–9. <http://www.bmj.com/content/343/bmj.d6423?tab=responses> [accessed 13 November 2011].
- 8 Lidegaard Ø. Authors' reply to Anne Szarewski *et al.* <http://www.bmj.com/content/343/bmj.d6423?tab=responses> [accessed 13 November 2011].
- 9 Lidegaard Ø. Authors' reply to Professor Shapiro. <http://www.bmj.com/content/343/bmj.d6423?tab=responses> [accessed 13 November 2011].
- 10 Spitzer WO. The 1995 pill scare revisited: anatomy of a non-epidemic. *Hum Reprod* 1997;12:2347–2357.
- 11 Severinsen MT, Kristensen SR, Overvad K, *et al.* Venous thromboembolism discharge diagnoses in the Danish National Patient Registry should be used with caution. *J Clin Epidemiol* 2010;63:223–228.
- 12 Shapiro S, Dinger J. Risk of VTE among users of oral contraceptives. *J Fam Plann Reprod Health Care* 2010;36:103.
- 13 Grimes DA. Epidemiologic research using administrative databases: garbage in, garbage out. *Obstet Gynecol* 2010;116:1018–1019.
- 14 Suissa S, Blais L, Spitzer WO, *et al.* First-time use of newer oral contraceptives and the risk of venous thromboembolism. *Contraception* 1997;56:141–146.
- 15 Suissa S, Spitzer WO, Rainville B, *et al.* Recurrent use of newer oral contraceptives and the risk of venous thromboembolism. *Hum Reprod* 2000;15:817–821.
- 16 Writing Group for the Women's Health Initiative. Risks and benefits of estrogen plus progestin in healthy postmenopausal women. *JAMA* 2003;288:321–333.
- 17 Dinger JC, Heinemann LA, Kühl-Habich D. The safety of a drospirenone-containing oral contraceptive: final results from the European Active Surveillance Study on oral contraceptives based on 142,475 women-years of observation. *Contraception* 2007;75:344–354.
- 18 Dinger J, Thai DM, Moehner S. The risk of venous thromboembolism in OC users: time patterns after initiation of treatment. *Pharmacoepidemiol Drug Saf* 2010;19:S214–S215.

- 19 **Lidegaard Ø, Edström B, Kreiner S.** Oral contraceptives and venous thromboembolism: a five-year national case-control study. *Contraception* 2002;65:187–196.
- 20 **Shapiro S.** Causation, bias and confounding: a hitchhiker's guide to the epidemiological galaxy. Part 2. Principles of causality in epidemiological research: confounding, effect modification and strength of association. *J Fam Plann Reprod Health Care* 2008;34:185–190.
- 21 **Shapiro S.** Causation, bias and confounding: a hitchhiker's guide to the epidemiological galaxy. Part 3: principles of causality in epidemiological research: statistical stability, dose- and duration-response effects, internal and external consistency, analogy and biological plausibility. *J Fam Plann Reprod Health Care* 2008;34:261–264.
- 22 **Bezemer ID, van der Meer FJ, Eikenboom JC, et al.** The value of family history as a risk indicator for venous thrombosis. *Arch Intern Med* 2009;169:610–615.
- 23 **Heinemann LA, Dinger J.** Safety of a new oral contraceptive containing drospirenone. *Drug Saf* 2004;27:1001–1018.
- 24 **Taubes G.** Epidemiology faces its limits. *Science* 1995;269:164–169.