Risk of venous thromboembolism among users of oral contraceptives: a review of two recently published studies

Samuel Shapiro, Jürgen Dinger

Abstract

Background Two recent studies, a cohort study from Denmark, and a case-control study from The Netherlands, have reported increased risks of venous thromboembolism (VTE) among users of oral contraceptives (OCs) containing desogestrel, gestodene, drospirenone and cyproterone, relative to the use of levonorgestrel.

Critique In the Danish study the comparisons were not valid. (1) VTE risk is highest soon after commencement of OC use, and duration of use was underestimated for levonorgestrel users, but not for drospirenone users; for the remaining compounds duration was only slightly underestimated. The underestimation for levonorgestrel resulted in systematic overestimation of the relative risks for the compared OCs. (2) Duration was also incorrectly estimated: only the duration of current use, not duration of all episodes of use was relevant to VTE risk. (3) Confounding was not adequately controlled.

In The Netherlands study the comparisons were not valid. (1) The relative risk for drospirenone versus levonorgestrel was not statistically significant. (2) Extensive publicity had been given to the risk of VTE among users of desogestrel, gestodene, drospirenone and cyproterone: information bias and detection bias were therefore likely. (3) Inadequate allowance was made for duration of use. (4) The combination of two different control groups, both of them likely to have been biased, into a single category was not valid.

Conclusion The best evidence continues to suggest that the increased risk of VTE in OC users is a class effect, dependent on the estrogen dose and duration of use, and independent of the progestogen used.

Keywords combined oral contraceptives, progestogen, risk assessment, venous thromboembolism

Introduction

Since the 1960s it has been known that the use of combined estrogen/progestogen oral contraceptives (OCs) increases the risk of venous thromboembolism (VTE). An increased risk has not generally been identified for progestogen-only preparations, but has been associated with the estrogen dose. In 1995, based on evidence from three studies published in the same issue of the Lancet, the further claim was made that OCs containing the progestogens, desogestrel and gestodene (‘third-generation’ OCs), increase the risk of VTE more than do ‘second-generation’ OCs, as represented by products containing levonorgestrel. In particular, the World Health Organization’s publication created a widespread pill-scare in the media. Soon further studies, some supporting and some not supporting a higher risk for users of ‘third-generation’ OCs, were published. The most detailed study suggested that there was no difference,9 but the topic remains controversial.10,11

Now, based on evidence from two studies recently published in the same issue of the British Medical Journal, a cohort study in Denmark,12 and a case-control study in The Netherlands,13 it has again been claimed that relative to levonorgestrel-containing OCs, desogestrel- and gestodene-containing OCs do indeed carry a higher risk of VTE. In addition, the authors claimed that OCs containing drospirenone and cyproterone also carry a higher risk. The Netherlands case-control study also found an increased relative risk for lynestrenol and norgestimate. Here we consider the validity of those claims, and we focus particularly on drospirenone, as results on that progestogen have received the greatest public attention and OCs containing that progestogen are now generally the most commonly used. Most of the following arguments, however, apply also to the other progestogens.

Background

We commence with some background considerations. 1. While it has been known since the 1960s that OCs increase the risk of VTE, more recently it has also been established in multiple studies that the risk is greatest soon after OCs are first used and that with continued use it soon declines. However, even after prolonged use the risk of VTE still remains elevated, although its magnitude is not as high as initially. One suggested explanation for this phenomenon is that those women who are most susceptible to OC-induced VTE tend to develop that complication soon after starting OC use, and are, as it were, ‘weeded out’, so that less susceptible women survive (‘depletion of susceptibles’). This explanation is disputed. However, whatever the explanation may be, what is not disputed is that the risk of OC-induced VTE is highest initially, and that it then declines with continued use. It also appears likely that women who have previously used OCs, and who then recommence use, have a higher risk during the first months of recurrent use compared to long-term use. Sub-analyses of the EURAS study suggest that the increase in the early months is linked to the restart of OC use as women who switched their OC without a break had no or almost no increased risk during the first months compared to the following months. The important point, however, is that the statistical analysis has to account separately for each episode of OC use.

2. As new OCs have been introduced over the years a widespread, although mistaken, perception has been that the most recently introduced products are the safest. Consequently there has been a general tendency to prescribe the most recently introduced OC to women thought to be at increased risk of VTE (e.g. because of obesity: a phenomenon known as ‘confounding by indication’).
addition, individual progestogens have specific pharmacological profiles (e.g. anti-mineralocorticoid and anti-androgenic effects of drospirenone and cyproterone, respectively), which may result in preferential prescribing of new progestogens to women with health conditions that are also associated with an increased risk of VTE. Indeed for the use of drospirenone-containing OCs preferential prescribing to obese women has been shown to be the case. For the same reason, among women who are already using OCs, women perceived to be at high risk of VTE are switched to the newest OCs in the (mistaken) belief that they are the safest or that their specific pharmacological effects may be beneficial for a specific medical condition. For example, women using an older levonorgestrel-containing OC may switch to a newer drospirenone- or cyproterone-containing OC because they have become obese, or because polycystic ovary syndrome has been diagnosed. In observational studies, confounding by indication can be reduced by separately comparing switchers with switchers and starters with starters. 

The issue of confounding by indication, and of other confounding, raises the question of what are the most relevant risk factors for VTE in a female population of fertile age, and in particular COC users. One of the most relevant risk factors is age. The risk of VTE increases four-fold in women between the ages of 15 and 40 years. Another important factor that has gained importance in recent decades is obesity. Many studies have shown that obese women [body mass index (BMI) ≥30] have approximately a four-fold increase in VTE risk during OC use compared to women of normal weight. In one study it has been shown that age and obesity are independent risk factors and that the combination of both factors results in an increased risk that is more than additive.

Genetic predisposition to VTE due to gene mutations such as factor V Leiden or prothrombin 20210A is also important. In everyday clinical practice, information on genetic mutations is usually not available, but a positive family history of VTE – which may reflect the presence of genetic risk factors – is an indicator of an increased risk.

Patients with cancer have a greatly increased risk of VTE. Haematological malignancies, lung cancer and gastrointestinal cancer are associated with particularly high risk estimates. The same holds for trauma, surgery and immobilisation. Long-haul travellers have also a high risk of VTE. In some studies, but not all, current smoking has been associated with a slight to moderate increase of risk.

Preferential prescribing of a specific OC to women with a high prevalence of one or more of these risk factors results in confounding, and any observational study that compares the risk of VTE among different OCs should have reliable information on such factors in order to adequately control them.

3. When an observational study is massive, involving millions of women and even more woman-years, any elevated risk, no matter how small, tends to be statistically significant. Significance, however, does not equate with causation. Even though an association may be significant it may nevertheless be due to bias or confounding – and in that circumstance such bias or confounding becomes ‘statistically significant’. In addition, when relative risks (RRs) are small, say 2.0 or less, the resolving power of the epidemiological microscope simply is insufficient to discriminate among bias, confounding and causation as alternative explanations. Given a small RR, causation can only be inferred, and then only tentatively, if a study is virtually perfect – which is seldom if ever the case.

Against this background we move next to a consideration of the risks associated with the use of cyproterone, desogestrel, drospirenone and gestodene, relative to the use of levonorgestrel, in the two studies.

The Danish cohort study
In Denmark all persons have a unique identity number; all prescriptions of medications, including OCs, are recorded in a national registry; all hospital discharge diagnoses are recorded in a disease registry; and information on education level is recorded in an education registry. Lidegaard et al. linked the registry data and carried out a cohort study among women aged 15–49 years from January 1995 to December 2005. The data comprised 10.4 million woman-years of observation and 3.3 million woman-years of OC exposure. Overall, 4213 incident cases of VTE were identified, 2045 of which occurred in current OC users. Because of the massive size of the study, all the RRs given below were statistically significant, and for simplicity the 95% confidence limits are omitted. Overall, relative to never having used OCs, the RR for current OC users, adjusted for age, calendar year and educational level, was 2.8, and the higher the estrogen dose, the higher was the risk. The RR also declined with increasing duration of use from 4.2 (<1 year) to 2.8 (>4 years).

Relative to the use of levonorgestrel-containing OCs combined with 30–40 µg ethinylestradiol, the RR adjusted for duration of use was 1.6 for drospirenone, 1.8 for desogestrel, 1.9 for gestodene and 1.9 for cyproterone.

Critique
Levonorgestrel-containing OCs have been in use since the 1970s, and follow-up in the Danish study only commenced in 1995. Inevitably, therefore, since a large proportion of levonorgestrel use commenced before 1995, and since the RR for short-duration use relative to never-use would have been the highest, the overall incidence of VTE among users of levonorgestrel was underestimated. Despite that underestimation, however, the identification of an overall increase in the risk of VTE risk for current OC users, and of an estrogen dose-response gradient, was qualitatively valid, probably because the associations were sufficiently strong to override this limitation. However, quantification of the RR estimates was unreliable.

Left censorship and duration of use
For the comparison of drospirenone with levonorgestrel the data were demonstrably invalid, both qualitatively and quantitatively. Levonorgestrel-containing OCs were in use for decades prior to 1995, whereas drospirenone has been in use only since 2000 and in Denmark only since 2001. Inevitably, therefore, because follow-up only commenced in 1995, the Danish study incorrectly estimated the duration of use of levonorgestrel (ie, there was ‘left censorship’), whereas the duration of use of drospirenone was not underestimated. Consequently, since the incidence of VTE would have been highest among levonorgestrel users soon after use commenced, and since such early use was incompletely recorded, the RR for the use of drospirenone relative to the use of levonorgestrel was overestimated.

The failure in the Danish study to take the duration of levonorgestrel use fully into account is not conjecture but can be seen in the study results. Relative to never-use, the RR for the use of levonorgestrel combined with estrogen in a dose of 30–40 µg was 1.9 for <1 year of use, 2.2 for 1–4 years of use, and 1.9 for >4 years of use (Table 2 in the Danish study); that is, the RR was no higher in the shortest duration category than in the longer duration categories. By
contrast, for the use of drospirenone relative to never-use of any OC, the corresponding estimates were 7.9, 2.7 and 3.3. Based on the drospirenone data it is likely that the relative risk of levonorgestrel in the first year of use was underestimated by a factor of about 3. According to Table 1 of the publication, 22% of the OC exposure was in the first year of use. Extrapolation of a probable three-fold risk in the first year to the overall incidence rate would result in an estimate of about 7.9 VTEs per 10 000 woman-years. This extrapolated incidence rate is similar to the incidence rate for drospirenone, gestodene, desogestrel and cyproterone, respectively (i.e. 7.8, 7.0, 6.5 and 7.1).

For desogestrel, gestodene and cyproterone a higher risk of VTE with short-duration OC use was again consistently present, but at a lower level compared to drospirenone. For these progestogens, left censorship again gave rise to a partial underestimation of duration of use as they were in use before 1995. However, compared to levonorgestrel, duration of use is less affected because market introduction was at a later point in time. Even more important, the market shares of these progestogens were more or less stable between 1995 and 2005 or they showed a strong growth over time (cyproterone). By contrast, the use of levonorgestrel declined.

In short, in the comparison of drospirenone and the other newer progestogens with levonorgestrel, the absence of reliable data on short duration use of levonorgestrel provides a plausible explanation as to why the incidence rate for levonorgestrel was lower than for other newer progestogens, drospirenone in particular.

Not only were the duration data invalid because of left censorship, but they were further invalidated because ‘length of use’ was defined as the total duration of use of prescriptions, “with periods of non-use subtracted (our own emphasis) if they occurred between periods of use”. Multiple studies have shown that the risk of VTE declines to baseline almost as soon as OC use is stopped. As stated above, an adequate statistical analysis has to take into account each separate episode of OC use. The duration of previous use was irrelevant, and what should have been assessed was the duration of current OC use only. Consequently, the use of levonorgestrel as the reference category against which to compare drospirenone, desogestrel, gestodene and cyproterone was not valid.

In an exchange of e-mail correspondence on the British Medical Journal website, Lidegaard argued that the failure to compare short-term users with short-term users and long-term users with long-term users would not have accounted for the findings, because in quantitative terms ‘attrition of susceptibles’ could not explain the differences in VTE risk. That reasoning misses the point. Whatever the reason, relative to never-use the incidence of VTE among levonorgestrel users was underestimated in the Danish data. Consequently, the use of levonorgestrel as reference category against which to compare drospirenone and cyproterone was not valid.

Confounding
In the comparison of drospirenone and levonorgestrel there were also additional uncontrolled sources of confounding. Obesity (BMI) increases the risk of VTE, yet that factor was not allowed for, even though it has also been shown to be a determinant of drospirenone use.15 Lidegaard argued that failure to control BMI had no material effect.12 That argument was not tenable because all the RRs were small (<2.0), and adjustment even for a minor degree of confounding could readily have rendered the association non-significant. Furthermore, obesity and excessive levels of endogenous androgenic hormones are two of the principal features of the polycystic ovary syndrome, which in turn may lead to preferential prescribing of cyproterone to women with that condition.36

Other evidence also indicates that failure to control confounding by BMI was a major defect in the Danish study. Two secular trends in Denmark during the study period (1995–2005) were ignored. (1) During that time, prescriptions of levonorgestrel-containing OCs declined substantially, while prescriptions of drospirenone increased steadily following market introduction in 2001 and (2) the prevalence of obesity in Danish women increased dramatically.37 These trends would inevitably have resulted in spuriously elevated RR estimates for drospirenone, relative to levonorgestrel. As the prescriptions of desogestrel, gestodene and cyproterone were stable or increased during the study period, these trends also negatively influenced the risk estimates for these progestogens.

Additional factors, such as a family history of VTE, could also have resulted in the preferential prescription of the most recently introduced OCs, and thus also have given rise to confounding by indication. The authors argued that “after new studies were published in the 1990s this preferential prescribing stopped and was not apparent during 1994–1998”.31 This observation, however, does not prove that cessation of preferential prescribing of ‘third-generation OCs’ was permanent. Nor does it apply to the potential for a selective tendency to prescribe drospirenone or other newer progestogens to women perceived to be at a higher risk of VTE in the mistaken belief that the newest OCs were the safest.

The behavior of one potential confounder was unexplained. The risk of VTE increased with decreasing level of education. Women who only completed primary school had the highest risk, while all other women had a lower risk (RR 0.5). In many studies in which education has been assessed, an increasing level of education has been associated with a higher frequency of diagnostic procedures, and consequently a higher incidence of diagnosed VTE, not the reverse. It may be that Danish women were different in this regard, but this is doubtful. This finding raises the possibility that there may have been selection bias in the diagnosis of VTE, with a tendency for the least educated women to be selectively diagnosed.

Other issues
In an earlier analysis of the Danish registries Lidegaard et al. found a RR of 0.7 for the comparison of cyproterone with levonorgestrel.39 In the most recent analysis12 the RR for that compound (1.9) was the highest. This inconsistency was unexplained, and it raises additional concerns about the validity of the results.

VTE information in the Danish registries is based on hospital diagnosis and is not validated. The authors suggested that “about 10% [of diagnoses were] uncertain”. Therefore, it is conceivable that the media attention given to drospirenone, triggered by publicity given to two fatal cases of VTE in Denmark and the Netherlands in 2002,39 led to biased hospital referral for diagnosis and treatment. Among OC users who developed VTE, and who recovered, there may have been a selective tendency to (correctly) diagnose VTE that would otherwise have gone undiagnosed.

Finally, relative to the use of levonorgestrel, the RRs for drospirenone, desogestrel, gestodene and cyproterone were all well below 2.0. As pointed out above, except when a study is virtually perfect, which is seldom if ever the case, for such small risk increments it is impossible to discriminate among bias, confounding and causation as alternative explanations.
Conclusions on the Danish cohort study
We conclude that the overall increases in the risk of VTE among current OC users, as well as the dose–response gradient according to the dose of estrogen, were qualitatively valid in the Danish study, probably because the associations were large enough to override the limitations. However, the comparison of drospirenone and the other newer progestogens with levonorgestrel was not valid, and the study has not demonstrated that those progestogens carry a higher risk of VTE. Left censorship, together with the failure to estimate duration of use for current OC users only, and obvious sources of uncontrolled confounding invalidated the comparisons with levonorgestrel.

The Netherlands case-control study
In this study 1524 cases of incident VTE below the age of 50 years, and identified at six anticoagulation clinics between March 1999 and September 2004, were compared with 1524 control women. 40.5% of whom were partners of male cases in the study database. Additional controls (59.5% of the total) were recruited by random digit dialling, after January 2002. In the analysis the two groups of controls were pooled.

Histories of OC use for cases of VTE and their partners were obtained within a few weeks after the index date. This information was verified ‘at least three months after the discontinuation of anticoagulation therapy’ when the patients and their partners were invited to the anticoagulation clinic for a blood sample and an interview. For the controls identified by random digit dialling, the questionnaire was sent after their agreement to participate.

Overall, the risk of VTE among current users of any OC was significantly increased by some five-fold, relative to non-use, and the higher the estrogen dose, the higher was the risk. In addition, the risk of VTE was highest in the first months of use. Relative to the use of levonorgestrel, the RR for drospirenone was 1.7 [95% CI 0.7–3.9]; for desogestrel it was 2.0 [95% CI 1.4–2.8]; for gestodene it was 1.6 [95% CI 1.0–2.4]; and for cyproterone it was 2.0 [95% CI 1.3–3.0]. These risk estimates were adjusted for ‘age and period of inclusion’ (legend of Table 3).

Among women who had used OCs for more than 2 years the risk of VTE for desogestrel “remained increased compared with levonorgestrel,” and was similar to the overall risk – for gestodene users the odds ratio was 1.5 (95% CI 0.9–2.60) and for desogestrel 1.9 (95% CI 1.3–2.0)”. Among women who used the “newest types of [OCs]” for 3 months or less the odds ratios were 1.9 for drospirenone (95% CI 0.2–21.3) and 1.6 (95% CI 0.3–9.9) for cyproterone.

Critique
The comparisons among the different OCs did not establish that the use of drospirenone, desogestrel, gestodene or cyproterone was associated with a higher risk of VTE compared to levonorgestrel for the reasons detailed below.

Lack of statistical significance
1. The RR for drospirenone versus levonorgestrel was not statistically significant, and it could readily have been due to chance. In addition, relative to never-use of any OC the 95% confidence intervals (CIs) were extremely wide (2.9–13.7), indicating that the data were sparse and fragile. Specifically for that comparison, this study was uninformative.

2. The RRs for the use of drospirenone and cyproterone versus levonorgestrel for 3 months or less were also not statistically significant, and could also have been due to chance. Moreover, although numbers were not given, the exceptionally wide CIs indicated that there were virtually no data. The authors acknowledged that the CIs were wide, but they nevertheless claimed that “the results support an increased risk for both drospirenone … and … cyproterone”. That claim was not justified.

3. Adequate allowance was made for the known difference in VTE risk for long-term users and short-term users – a difference that was confirmed in this study. As mentioned above, levonorgestrel was introduced three decades before drospirenone was introduced, and it is likely that a substantially greater proportion of the levonorgestrel users were long-term users compared to drospirenone or cyproterone users, thus invalidating the comparisons.

The investigators attempted to allow for duration of use by comparing the use of gestodene and desogestrel with the use of levonorgestrel among women who used these OCs for >2 years – and for drospirenone and cyproterone, relative to levonorgestrel, they confined the comparisons to <3 months of use. Those comparisons did not constitute adequate allowance for duration of use. In addition, in the latter comparisons the exceptionally wide CIs indicated that there were virtually no data.

Information bias
In The Netherlands widespread publicity had been given to the allegation that ‘third-generation’ OCs increase the risk of VTE more than do ‘second-generation’ OCs. The same allegations were made regarding drospirenone and cyproterone. Thus exposed cases could selectively have been referred to the anticoagulation clinics. Most cases of VTE recover. Women with deep vein thrombosis who recovered, and whose symptoms were minimal (eg, a swollen leg) could selectively have been investigated (eg, Doppler ultrasound), and hence correctly have been diagnosed if they used ‘third-generation’ OCs, drospirenone or cyproterone. Publicity was not given to levonorgestrel; and among women who developed OC-induced VTE, and who recovered, the diagnosis could more commonly have been missed when they used that compound.

4. The cases were interviewed months after the onset of VTE, and after publicity had been given to the claim that ‘third-generation’ OCs, drospirenone or cyproterone increase the risk of VTE. The long time lapses makes it likely that the reporting of the use of such OCs was biased.

Selection of controls
5. The selection of 40.5% of the controls in the study was based on ‘reverse sex matching’. For the analysis of OCs the female partners of the male cases in the study database were used as controls. It is likely that those partners were not representative of the Dutch female population.

6. Incomplete data were given on the attrition of potentially eligible controls identified by random digit dialling: 2788 out of 4350 (64%) eligible random controls returned the questionnaire. That attrition was further augmented by failure to pick up the telephone, failure to respond to a message on an answering machine, and failure to provide information on potential eligibility. Thus the true proportion of eligible women who did not respond was unknown, but it was undoubtedly much higher than 36%. With the publicity given in The Netherlands to the putative harmfulness of ‘third-generation’ OCs, and later on of drospirenone and cyproterone, willingness to participate could have been biased.

7. The combination of the two biased control groups into a single category was not valid.

Conclusions on The Netherlands case-control study
Despite the multiple defects in this study, we conclude that the overall increases in the risk of VTE among current OC users, the dose–response effect for estrogen, and the highest risks for short-term OC use were valid. However, the
comparisons of desogestrel, gestodene, norgestimate and cyproterone with levonorgestrel were not valid. The comparison of norgestimate with levonorgestrel was not statistically significant; the combination of two control groups derived from different sources was not valid; differences between short-duration and long-duration users were not taken into account adequately; information and selection bias were likely; and all the ORs were small (2.0 or less). Especially for the short-term use of drospirenone and cyproterone, the exceptionally wide CIs indicated that there were virtually no data.

Other studies
Findings from two large cohort studies specifically designed to evaluate the risk of VTE among women using drospirenone have been reported. In a cohort study of almost 60 000 European women there was no evidence of an increased risk of VTE among users of drospirenone or the other new progestogens, relative to users of levonorgestrel.\textsuperscript{15} That study will not be reviewed in detail here, but it should be noted that in that study starters were compared with starters and switchers with switchers and that long-term users were excluded. In addition, in the design and in the analysis full provision was made for control of confounding by multiple factors, including duration of OC use, obesity and family history of VTE. In another large study of almost 67 000 women carried out using an automated claims database in the USA (also not reviewed here in detail) there was again no evidence of an increased risk of VTE among drospirenone users relative to the use of OC containing other progestogens, including levonorgestrel (hazard ratio 1.0).\textsuperscript{41} In that study propensity score matching was used to achieve a close balance on the patient characteristics available from the automated claims database. A follow-up validation study suggested an even distribution of relevant confounders such as BMI and family history of VTE.

Both studies have been criticised as being industry-sponsored and underpowered.\textsuperscript{13} With regard to industry sponsorship, the studies were mandated by the regulatory authorities; the protocols and statistical analysis plans were reviewed and approved by them; the investigators were independent; cases identified in both studies were assessed without knowledge of exposure status, and they were reported to the regulatory authorities; and the studies were carried out under the supervision and authority of independent advisory boards that monitored the studies and approved publication of the findings. As for the studies being underpowered, the 95\% CIs were narrow, and both studies were statistically robust.

Overall conclusions
In an editorial published together with the Danish and The Netherlands studies, Dunn asserted that “the remarkable concordance of the main results of these studies supports their validity”.\textsuperscript{42} In fact, the findings were significantly discordant: relative to never-use of OCs the overall RR for current OC use in the Danish study was 2.8 (95\% CI 2.7–3.0), while in The Netherlands study the OR was 5.0 (95\% CI 4.2–5.8).

We conclude that the most parsimonious interpretation of the evidence that exists to date continues to suggest that the increased risk of VTE among OC users is a class effect, dependent on the estrogen dose; and as best can be judged, given the limits of observational research, the risk appears to be independent of the specific progestogen component.

Statements on funding and competing interests
Funding None identified.

Competing interests The authors presently consult, and in the past have consulted, with manufacturers of products discussed in this article. At the request of Bayer Schering, Samuel Shapiro wrote a critique for presentation to regulatory authorities. Parts of that critique are repeated here verbatim or in modified form. Jürgen Dinger was actively involved in the conduct of the EURAS, a post-authorisation safety study for European health authorities, which was funded by the manufacturer of products discussed in this article (Bayer Schering Pharma AG).

References
19 Pomp ER, le Cessie S, Rosendaal FR, Doggen CJ. Risk of venous thrombosis: obesity and its joint effect with oral...
22 Trussell J, Guthrie KA, Schwarz EB. Much ado about little: obesity, combined hormonal contraceptive use and venous thrombosis. Contraception 2008; 77: 143–146.