Comparative risks of venous thromboembolism among users of oral contraceptives containing drospirenone and levonorgestrel

Klaas Heinemann, Lothar A J Heinemann

Background
Two case-control studies, both with the involvement of the Boston Collaborative Drug Surveillance Program, have recently been published in the British Medical Journal (Jick et al.1 and Parkin et al.2) in which increased risks of venous thromboembolism (VTE) among users of drospirenone (DRSP)-containing oral contraceptives (OCs) compared with levonorgestrel (LNG)-containing OCs were reported. Both studies used electronic medical databases [USA: managed care plans (Pharmetrics)1; UK: general practices (GPRD)2].

Risk assessments of VTE in a predominantly young and healthy population identified in electronic databases always face similar methodological problems concerning the validity and completeness of the recorded diagnoses, the selection of comparison groups, and the completeness and validity of the information on potential confounders.3–5 These problems are reminiscent of the issues experienced during the ‘third vs second generation’ pill crisis. Recently in Austria, where the impact of the crisis in the mid-1990s on OC use was negligible, a study was conducted in which the risk of VTE for ‘third vs second generation pills’ was investigated. The methods were the same as those used 10 years earlier. However, over the intervening 10 years the characteristics of the third- and second-generation pill user populations had grown similar, and – importantly – the elevated risks reported in the mid-1990s were not found.6

In the two database studies now reported it is unlikely that the characteristics of the populations of DRSP- and LNG-containing OC users were sufficiently similar to permit valid comparisons, for the reasons given below.

Under-ascertainment of VTE incidence
There was major under-ascertainment of the incidence of VTE, as has also been the case in previous database studies. Among OC users it is now well established that the incidence is of the order of 9–10 per 10 000 woman-years (WY), and in some studies it is as much as 13.5 per 10 000 WY.7 By contrast, in database studies the incidence in OC users has ranged from 0.7 to 3.8 per 10 000 WY. As has been the case in previous database studies, the two new studies were again characterised by major under-ascertainment of cases of VTE.

In the study of Jick et al.,1 among OC users the overall incidence of idiopathic plus non-idiopathic cases of VTE was 5.2 per 10 000 WY, or about half of what has been documented in well-designed cohort studies.8 For DRSP users the incidence was not mentioned, but it can be estimated to have been about 7.9 per 10 000 WY – close to the overall incidence among all OC users in studies not derived from databases. For LNG users the incidence was 3.2 per 10 000 WY, and two- to three-fold lower than in studies using patient-reported outcomes with 100% validation of cases. The latter rate is no different from the baseline incidence of VTE in non-users of OCs in cohort studies.8

The incidence rates of idiopathic VTE reported in Parkin et al.2 were even lower than those reported by Jick et al. Parkin et al. did not mention the number of non-idiopathic cases that they excluded; however, the rules for exclusion were broadly the same to those used by Jick et al. If it is assumed that about 39% of the cases were defined as idiopathic, as in Jick et al., the
total incidence of idiopathic plus non-idiopathic VTE in the DRSP-exposed women would have been about \(5.9\) per \(10,000\) \(WY\), and in the LNG-exposed 2.3 per \(10,000\) \(WY\). Those incidence rates were not only lower than in the abovementioned cohort study,\(^9\) they were also much lower than in a recently published database study from Denmark.\(^9\)

**Diagnostic bias**

The substantial under-ascertainment of VTE cases in the Jick et al. study made bias virtually inevitable. Over 90% of cases of VTE cause no or only mild, unspecific symptoms,\(^9\) and the most common presentation, a swollen leg, may or may not be VTE. In OC users awareness that a swollen leg could be VTE, which then is followed by the decision to carry out diagnostic procedures, can readily be selectively influenced by negative publicity about a certain OC. In the USA considerable publicity has been given to the allegation that DRSP-containing OCs cause more VTE than do other OCs. That publicity has been reinforced by litigation. Awareness of the possibility of VTE, followed by diagnostic work-up, is likely to have been greater in users of DRSP-containing OCs than in users of other OCs, and the major under-ascertainment of cases maximised the opportunity for this to occur.

As in the USA, in the UK publicity has been given to the allegation that DRSP is more thrombogenic than other progestogens. Again, given the marked under-ascertainment of VTE cases, it was virtually inevitable that the data would be biased by a tendency to selectively diagnose VTE more commonly in DRSP-containing OC users than in users of other OCs.

Not only was there under-ascertainment of VTE incidence in the Parkin study, but there were also substantially different odds ratios for the risks of deep venous thrombosis (DVT) and pulmonary embolism (PE) reported: for DVT the odds ratio for DRSP vs LNG was four-fold higher than the corresponding odds ratio for PE. Yet the incidence of PE should directly have correlated with the incidence of DVT. The non-correspondence therefore constitutes quantitative evidence of substantial diagnostic bias: among women who presented with mild or non-specific symptoms it is likely that physicians more commonly initiated diagnostic procedures among DRSP users compared to LNG users.

**Validation of VTE diagnosis**

Jick et al. did not validate the diagnosis of VTE, and they claimed that potential misclassification of recorded VTE diagnoses was likely to be non-differential among users of DRSP-containing and other OCs. However, no data were provided to support that claim. For the reasons given above it is likely that special attention was paid to DRSP-containing OCs. Moreover, it is well established that when a class of drugs (i.e. OCs) is known to increase the risk of a well-established side effect (i.e. VTE), there is a selective tendency to more commonly suspect and more commonly diagnose that side effect in users of the most recently introduced compound.\(^11\)–\(^13\) And still further, DRSP-containing OCs were introduced with a special label that emphasised the risk of VTE. That emphasis could further have augmented a selective tendency to suspect and diagnose VTE in DRSP-exposed women.

In the study of Parkin et al. over one third of cases were not validated. When only the validated cases were analysed, no statistically significant result was evident. In the light of the missing validation data (Jick et al.) or partially missing data (Parkin et al.), the under-ascertainment of actual cases of VTE might have been even more profound than estimated above. Problems with the validation of reported outcomes in database studies, and the limitations to the utility data, have been widely discussed.\(^1\)–\(^3\) In the present instance the under-ascertainment of cases, coupled with the failure to validate them entirely in the one study\(^1\) and partly in the other,\(^2\) is good reason to question whether the results can be interpreted at all.

**Duration of use effect**

Neither Jick et al. nor Parkin et al. adequately evaluated the effects of the duration of use. It is well established that the risk of VTE among OC users is greatest in the initial 3 months of use, after which there is a rapid decline to a lower but still elevated level of risk (Figure 1).\(^5\)\(^14\)–\(^17\) This pattern was not evident in the study of Jick et al.,\(^1\) and cases of VTE exposed to OCs for <3 months were under-represented. Thus it is likely that under-ascertainment of VTE in the initial months of OC use was even more marked than later on. The failure to identify an early use effect is sufficient, on its own, to invalidate the findings in the study of Jick et al. Parkin et al. adjusted for confounding by duration of use, but they did not provide the data needed to evaluate whether there was an early use effect in their study.

**Exclusion of ‘non-idiopathic’ VTE cases**

Jick et al. and Parkin et al. have claimed that their findings were valid because they excluded so-called ‘non-idiopathic’ cases of VTE, whereas studies reporting no difference between DRSP vs other OCs were not valid because they included predisposed cases. That claim is demonstrably incorrect: the occurrence of VTE in obese women, or in women with a family history of VTE, or in women with other predisposing factors, sometimes unknown (e.g. Factor V Leiden) can hardly be designated as ‘idiopathic’. Moreover, in the study of Jick et al. obesity, hypertension, atherosclerosis, hyperlipidaemia/hypercholesterolaemia, asthma, emergency room visits, and physician visits were all more common among cases of VTE than among controls. Without any question, general ill health predisposes to VTE, and the notion that an entity given the label ‘idiopathic VTE’ can be defined, and be reliably identified, is illusory. Furthermore, the combination of two or more risk
factors – such as Factor V Leiden mutation and OC use – leads to an over-additive risk. Therefore, the analysis of ‘idiopathic’ VTE underestimates the risk under real-life conditions. If the objective is to determine the risk of VTE in the OC-using population at large, it is more valid to record potential confounders as fully as possible and to make proper allowance for them in the analysis.

Lack of information on important prognostic factors
As with database studies in general, in both of the recently reported studies information on important risk factors for VTE was missing. In the study of Jick et al. no information was available on body mass index (BMI) (all that was recorded was the presence or absence of a diagnosis of obesity, which was not even defined) (Figure 2); there was also no information on a family history of VTE. Active surveillance studies have shown that both prognostic factors can be confounders. In the study of Parkin et al. there was no information on a family history of VTE, and statistical imputation was used to allow for missing data on BMI. The validity of imputation is based on the assumption of randomness, and for BMI that assumption may not have been tenable, since a high BMI has been shown to be a determinant of DRSP use.

Conclusions
It needs to be stressed that database studies such as the two studies considered here have been beset with
many more problems, and with less methodological transparency, than have been cohort and case-control studies specifically designed to compare VTE risks among users of different OCs, carried out in the population at large.19 Several recent studies have not identified differences in the risk of VTE for DRSP vs other OCs, including OCs containing LNG.8,20,21

The two recent studies suffered from the same methodological limitations as previously published studies based on electronic databases. It is therefore hardly surprising that the results replicate previous findings from studies that have employed this methodology. By contrast, studies designed specifically to evaluate VTE risk among users of different OCs have proven to be more valid, and more informative. Discussions on how to improve the methodology of electronic database studies are urgently needed.22

In an editorial accompanying the studies of Jick et al., and Parkin et al., Dunn has claimed that they accord with the earlier studies23 he claimed that those studies were ‘remarkably concordant’. In fact, as has been pointed out,4 the two earlier studies were not concordant. On the contrary, they were statistically incompatible. We conclude that the findings in both of the recent studies are difficult to interpret because of major under-ascertainment of VTE cases, the likelihood of biased diagnosis of VTE conditional on DRSP use, failure to identify the markedly elevated risk of VTE soon after commencement of use – or to make detailed allowance for it, and failure to adequately control potential confounding.

Competing interests Lothar Heinemann was an investigator for several studies mentioned in this commentary. He has also acted as a consultant for manufacturers of oral contraceptives in the past, but has not been actively involved in such consultancy work since 2006. Klaas Heinemann is an investigator on several post-authorisation studies that are supported by unconditional grants from manufacturers of hormonal contraceptives.

Provenance and peer review Commissioned; externally peer reviewed.

References