PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (see an example) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below. Some articles will have been accepted based in part or entirely on reviews undertaken for other BMJ Group journals. These will be reproduced where possible.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Exposure to oral hormonal contraceptives and risk of venous thromboembolism: a protocol for nested case-control studies using the QResearch and the CPRD databases
AUTHORS	Vinogradova, Yana; Coupland, Carol; Hippisley-Cox, Julia

VERSION 1 - REVIEW

REVIEWER	Susan Jick DSc Director Boston Collaborative Drug Surveillance Program and Professor of Epidemiology Boston University School of Public Health, USA I have published articles on the topic of hormonal contraceptives and venous thromboembolism
REVIEW RETURNED	23-Dec-2013

GENERAL COMMENTS	This manuscript describes a study protocol not a completed study, thus many of the points in the score sheet above are not applicable. I have gone into detail in the comment section below to describe concerns and limitations of the proposed study. This protocol by Vinogradova et all describes a study to evaluate newer hormonal contraceptives in relation to VTE compared to other hormonal contraceptive users. The stated objective of the study is to overcome prior study biases and limitations and to update information on risk for drospirenone and cyproterone containing OCs. (Note: though I would not necessarily call cyproterone containing-OCs “newer” since it was marketed in the late 1990’s, and they are no longer used in some countries such as France because of the established increased risk of VTE with these pills) The title of the study describes a study of hormonal contraceptives (HCs) in relation to VTE yet the focus of the study is on oral contraceptives with a brief nod toward other HCs. This title should more accurately reflect the exposures discussed in the protocol. Also, they authors suggest that the cyprroterone-containing OCs were mostly marketed to treat severe acne, hirsutism, and PCOS (last paragraph page 5). Do they mean in addition to its use as a contraceptive? If not, perhaps it should not be included in this protocol that describes VTE in HCs. The Introduction contains some errors and some inaccurate descriptions of the prior literature: Reference 7 is not a Danish study as reported. It is from the Netherlands. The authors cite reference 8 and say that according to that paper the studies in references 5-7 have “serious” methodological limitations. In fact, the paper by Shaprio and Dinger (ref 8), only discusses 2 papers and these are not among the papers referenced (5-7).

	Reference 8 was published in 2010 and so could not have discussed the papers by Lidegaard et al (refs 5 and 6) which were published in 2011 and 2012. The third reference (7) was a paper by Vlieg et al. This paper was not discussed in the Shapiro/Dinger paper either. I suggest that the authors of this protocol read the papers being critiqued by Shapiro and Dinger and consider the merits of the criticisms themselves. I would not accept the conclusions of Shapiro and DInger so readily. There is considerable disagreement in the epidemiology community on the proper methodology for the studies of hormonal contraceptives in relation to VTE and only a careful review of each study can properly reveal the true differences between them and identify which methodological differences would most likely explain the differences in the various study results. I would not accept the opinion expressed in this one review without careful consideration of the strengths and limitations of all the published studies. For References 10 and 11 the “main” limitations reported by the authors were that pregnant and postpartum women were not “identified”. I respectfully suggest that these were hardly “the main” limitations compared to many others that would have had a much larger impact on the study results. To start, the Gronich study (ref 10) compared current second generation OC use to current drospirenone use. Pregnant women would not be current users of OCs, and thus there would have been no need to “identify” them in the study; thus not doing so would not be even a minor limitation. The issue of postpartum use of combined OCs could have created minor bias but postpartum OC users would represent few women since relatively few women start combined OCs within a few months of giving birth. Thus this “main” limitation identified by the authors is likely not important. A more important limitation of both studies (10 and 11) is that they included women with other important proximate causes and risk factors for VTE such as recent surgery, lower limb injury and trauma, cardiovascular disease and cancer among others. Study ref 11 even included women with prior VTE. The small number of idiopathic cases in this study was likely the biggest limitation. The inclusion of non-idiopathic cases, which could have a big impact on the results, was never discussed by the protocol authors. I suggest that different case inclusion criteria are among the biggest differences between the studies referenced and that the inclusion of pregnant and postpartum women is a small portion of these women. The inclusion of fatal cases is also not likely to have a big impact on the study results; 1) because not many young women die from a VTE and 2) because it is unlikely that the deaths would occur disproportionately in users of one OC rather than another. The authors have identified many limitations that would have minor if any impact on the study results and have neglected to discuss more important limitations such as case inclusion/exclusion criteria and selection of the reference group. Each published study has different limitations, some major and some minor, but they should not all be so quickly dismissed. Further, the study methods described in this proposal will not in fact address the important limitations of former studies, and the authors have not successfully identified the factors that have led to the differences in the previously published study results. That is, this study includes many of the “serious” limitations of some of the former studies. These are discussed below.sji Methods How well has QResearch been validated? To my knowledge there is no way to link electronic patient data to original records so validation is not to the same scale as in CPRD. The reason I point this out is

	that there is opportunity for outcome misclassification from missing as well as incorrect diagnosis of VTE. If there are differences in results between the two databases this could be one explanation. It is important to acknowledge that the databases are not identical. The CPRD was developed with the express intention to collect data for research. This is not the case for EMIS data, which is the basis for the data in QResearch. A separate description of each database would be useful. The authors say this is a nested case-control study but it seems to include all women in the databases of the relevant age group. So I'm not sure why it is a nested study. They also say that use of a case-control design produces "unbiased estimate". This is only true if the case definition and control selection are appropriate as well as the selection of the exposure and referent. I do not know what the authors mean when they say that case-control studies produce unbiased estimates. Under the heading Cases and controls: the authors describe the study design. Perhaps the heading should be changed. They then go on to provide the definition of a case. Cases are defined as all eligible women with a diagnosis of VTE excluding only women with a prior VTE, women who are pregnant or within 3 months of a pregnancy, and women with anticoagulation therapy more than 6 weeks prior to the VTE date. This is a very loose definition of VTE and will likely include many non-cases. It is easy to require a prescription for an anticoagulant to confirm the VTE diagnosis and I would recommend adding this requirement to the case definition to avoid misclassification. In addition, particularly because the authors plan to include non-exposed cases in this study, it is important to exclude women with prior cardiovascular disease and other important risk factors for VTE such as cancer, renal disease, etc., from the study since oral contraceptives are contraindicated in these women. The authors indicate that they will use ONS mortality data to identify fatal cases of VTE. It should be noted that, at least in the CPRD, these data are only available for a portion of all practices (those in England). Practices in Wales, Scotland and Ireland are not linked to mortality data. This may be true for QResearch data too. The same is true for the HES data. Under control matching, the authors say that they will match controls to cases on age (please specify within 1 year, 2 years, etc.), and on calendar year. They then say in the next sentence that controls will be allocated the same index date as their matched case, so in fact they will be matched on the actual index day, not on calendar year. Interventions: Since this is an observational study, not a clinical trial I would suggest using the term "Exposure" rather than "intervention" since intervention implies that the researcher will have some control over the exposure allocation which they will not. It is not necessary to repeat the case definition here. Exposure = at least 1 Rx for a combined pill containing either drospirenone or cyproterone in the year prior to the index date. "Overall exposure" will be assessed to take into account switching. I am not sure what "overall exposure" means. If a woman switches from a drospirenone OC to a second generation OC how will her exposure be classified? The authors list under recency of use different windows of exposure including current, recent, remote, past and nonuse in the past year. This implies that there will be a non-exposed referent category but it is never actually specified what the referent will be. The authors also say on page 9, paragraph 2 that the main focus will be on drospirenone and cyproterone containing OCs versus the older HCs,

	so this leaves the reader confused. If the Comparison will be use of any other OC, including combined OCs with levonorgestrel, norethisterone, norgestimate, desogestrel, and gestodene (as they indicate on page 9), the choice of reference group will bias the results toward the null. Many of these comparison OCs have been shown to increase the risk of VTE compared to second generation OCs, so including these in the referent will increase the baseline risk of VTE in the comparison group and yield lower relative risks for the drospirenone and cyproterone OCs in comparison. As I read it, duration of use will be evaluated without regard to whether the use was current at the index date. This will underestimate any true effect of duration since non-current use has been shown not to be associated with an increased risk of VTE. In the last paragraph of the section titled Intervention, the authors briefly mention that use of the contraceptive patch will be analyzed as will be other HCs (IUDs, rings, etc.). These all have different exposure considerations (how use is determined, duration and recency of use, etc.). This should either be described more completely or left out of the study. The choice of referent group for these HCs is also missing here. Confounding Factors: Many covariates will be included and evaluated as confounders. Family history of coagulation defects however will not be recorded in any complete or systematic way (at least not in the CPRD) so I do not think it will be feasible to include this covariate. The authors mention that they will “take into account” several other medical conditions if they occur in the 6 months prior to the index date: acute infection, surgery, and leg or hip fracture. It is notable that they have left these most important other proximate causes (surgery and limb injury) to the very last sentence of the Confounding section. It is also notable that major trauma is not noted here. These are by far the most common causes of VTE in relatively young and healthy women, which characterizes women taking hormonal contraceptives. Because these are such strong risk factors for VTE it is unlikely that there will be any differential effects of different hormonal contraceptives (HC) on the VTE risk in women who have these other proximate causes for their VTE. If this is the case, then the distribution of HC use in these cases should be the same as the distribution in the controls which would lead to a null effect comparing different HCs even if there are true differences in the risks of the HCs. I would suggest excluding cases with other proximate causes for their VTEs in the 3 (or 6) months prior to the VTE. At the very least, the analyses should be stratified according to idiopathic and non-idiopathic case status (cases where a proximate cause is present). It is not sufficient to conduct a sensitivity analysis at the end when there is a strong likelihood that this will be an important source of bias. Statistical analysis: The authors state that they will use conditional logistic regression to estimate odds ratios, but they never specify the referent group. I think it is important to be clear what this referent will be. They can use multiple referents but they need to be specified. The authors state that a sensitivity analysis will be conducted on the subgroup of patients with an anticoagulant code in the 6 weeks post VTE. As indicated above, I think all analyses should be restricted to these patients. Anticoagulation is a known and necessary treatment for VTE and it is almost inconceivable that a patient with a true VTE would not be treated with anticoagulation. To include non-treated patients would invite misclassification. I would suggest conducting analyses stratified by age to test for effect modification by age since increasing age is associated with

	increasing risk of VTE and decreasing use of OCs. Sensitivity analysis: Restricting to HES, ONS: As stated above it should be clear that a large proportion of patients will not be linked to HES or mortality data. Sample Size: Drospirenone and cyproterone OCs were never heavily used in the UK and thus the number of users in both the CPRD and QResearch will not be as great as for other OCs. Since these are the purported exposures of interest the number of users of these 2 OCs should be used in the power calculation, not the total number of HC users. Limitations: It is a limitation that some women receive contraception at family planning clinics and their HC use may be missed in these data. If the study were restricted to current HC users then this would not be a concern but since the authors plan to compare OC use to non-use there may be users included in the non-use category. It is not clear why the authors have chosen to have a non-exposed category (if in fact they plan to include the non-use exposure in their analyses). According to French RS, Mercer CH, Johnson AM, et al. (Use of contraceptive services in Britain: findings from the second National Survey of Sexual Attitudes and Lifestyles (Natsal-2). J Fam Plann Reprod Health Care 2009;35(1):9-14.), approximately 15% of women utilize clinics for contraception,

REVIEWER	Øjvind Lidegaard Rigshospitalet, University of Copenhagen Denmark The reviewer has been an expert witness in two US legal trials in 2011 and 2012, and has received honoraria for speeches in pharmacoepidemiological issues.
REVIEW RETURNED	26-Dec-2013

GENERAL COMMENTS	Vinogradova Y, Coupland C, Hippisley-Cox J. *Exposure to hormonal contraceptives and risk of venous thromboembolism: a protocol for nested case-control studies using the QResearch and the CPRD databases.* **Reviewers comments.** This protocol describes a new planned study. ***Abstract*** Introduction: *"Many studies have found an increased risk of venous thromboembolism associated with the use of hormonal contraceptives, but these have been against a background of evolving drug technology and all have been subject to various biases and methodological limitations."* Nowhere are the claimed biases and methodological limitations in these studies documented. *"This study will focus on newer hormonal contraceptives – drugs containing drospirenone and cyproterone".* While drospirenone has been 12 years on the market, oral contraceptives with cyproterone have now been on the market for several decades. Therefore, the focus of this study is not on newer hormonal contraceptives, but on an old combined pill and on one 12 years old combined pill. According to the protocol, all types of hormonal contraception are included. This should be stressed in the

abstract.
Methods and analysis.
Expected number of included cases is not indicated.
Exposure is defined as *"at least one prescription of hormonal contraceptives in the year before the index date."* This will severely underestimates the risk with use of hormonal contraception, as use longer back in time than few weeks are unlikely to increase the risk of VTE.

Strenghts and limitations

"No recall or selection biases." No one has delivered any evidence that previous studies were affected by selection bias.

"Possible uncertainty in diagnosis of venous thromboembolism."
Uncertainly in some diagnoses of venous thromboembolism is a clinical condition no study design can remove. What can be done, and what should be done is do define a priori which criteria a certain case has to fulfil to be considered as a case. Such criteria are not indicated.

"Underestimation of hormonal contraceptive use".
The fact that a significant proportion of women get their hormonal contraceptive product from contraception clinics and thereby will be misclassified as non-users will certainly underestimate the risk of VTE from use of hormonal contraceptives, but not necessarily the difference in risk between different product groups.

Key words

The following key words were suggested: "*bisphosphonates, neoplasms, case-control studies, osteoporosis/ drug therapy, risk factors."*
None of these key words are appropriate. Use instead: Oral contraceptives, hormonal contraception, progestogen only contraception, venous thromboembolism, PCOS.

Introduction

"These associations have no established biological explanations, however, beyond speculation that third-generation progestins may have potentiated oestrogenic effects on clotting factors[1]."
This is not correct. Several studies have demonstrated a higher increase in SHBG and an increase in activated protein C with those products implying the highest risk of VTE.

"Currently, commonly-used fourth-generation pills, specifically those containing drospirenone (introduced in 2002) rather than levonorgestrel, have also been shown to have associations with increased risk of VTE in three large studies based on a general female population with data from the Danish Registry[5-7], but these studies had serious methodological limitations including lack of adjustments for confounding factors, selection bias and confounding by indication[8]."
First, reference 7 was a Dutch study demonstrating the same increased risk of VTE with use of 4th generation pills as the two Danish studies (ref. 5 and 6). Secondly, the company sponsored authors of reference 8 did not claim the Danish studies to be influenced by selection bias or confounding by indication. Other critique points have been refuted by the authors of these studies

	(Lidegaard *Ø. Critique of a Danish cohort study on hormonal contraception and VTE.* J Fam Plann Reprod Health Care 2010; 36: 103-4.pdf and Lidegaard *Ø. Reply to Jürgen Dinger and Samuel Shapiro.* BMJ 2011; online December 12, 2011.) *"Two of those studies, using an Israeli database[10] and USA insurance data[9], identified 165 and 18 VTE cases on drospirenone and reported an increased risk compared with other hormonal contraceptives. The main limitation for both studies, however, was no identification of pregnant or post-partum cases at the time of the diagnosis"* The circumstance that these studies did not exclude pregnant and post-partum cases would have underestimated the risk of VTE in current users of drospirenone, and not the opposite. *"A German study[11]based on 26 drospirenone users showed no difference with the other hormonal compositions – the study included pregnant women but failed to adjust for this condition"* They also failed to exclude other predisposed women such as women with previous thrombosis, known thrombophilia, cancer etc, explaining why the only two studies not demonstrating a difference between 2nd and 3rd/4th generation pills (both by Dinger et al) did not do so. *Two other studies from the same researchers [12, 13] excluded pregnant and post-partum women from the analysis and showed increased VTE risk associated with contraceptives that contained drospirenone rather than levonorgestrel. The one based on CPRD data[13] showed a 3-fold increased risk, but the precision of estimates was limited as only 17 exposed cases were identified.* These authors included CPRD data covering the period 2002-2009 and found 61 idiopathic VTE cases. The proposed new study covering the period 2001-July 2013 is thus expected to include about 100 women with idiopathic VTE from the CPRD research database. *"The one[12] based on American insurance data was the only study of the five, which adjusted for menstrual disorders – one of the indications for hormonal contraceptives and, as shown in the paper, associated with increased risk of VTE. Although the study reported a 2.4-fold increased risk among drospirenone users, it was conducted only on participants who survived."* With a case-fatality rate of about 1% among young women with VTE the exclusion of fatal VTE in this study is very unlikely to have biased the results of this study. *"Another hormonal contraceptive containing cyproterone has been used mostly for treatment of women with severe acne, hirsutism and polycystic ovary syndrome (PCOS). It has been shown that VTE risk is higher for such groups of women[14] but, of the two studies of hormonal contraceptives and VTE risk which adjusted for PCOS, one analysed only twelve exposed cases[15] while the other did not include cyproterone-containing contraceptives in the analysis[14]."* Thus the results for drospirenone containing oral contraceptives in study [14] were adjusted for PCOS and still significantly higher than for 2nd generation products. *"Three large population-based studies, which did include*

cyproterone [3, 5, 7], omitted to adjust for PCOS, so their results might be subject to confounding by indication."

Yes, PCOS is a potential confounder, but it is still unlikely that preferential prescribing of 3rd and 4th generation products to women with PCOS can explain the increased risk of VTE in users of these products.

"The proposed nested case control studies based on the female general population will investigate the association between the use of hormonal contraceptives – classified by type and use – and risk of VTE adjusted for PCOS and menstrual disorders as the major possible indications, other co-morbidities and concomitant drug exposure. It will concentrate on the most recent compositions and increase its power by combining the results obtained from the two of the largest electronic medical records databases, QResearch and CPRD analyses."

No information is given on how the authors will identify and verify the PCOS women and women with menstrual disorders. Menstrual disorders (e.g. dysmenorrhea) have never been demonstrated to be associated with VTE except if combined with PCOS.

METHODS AND ANALYSIS

Cases and controls

"*Eligible cases and controls will have at least two years of records prior to the index date."*

This is one third as compared with the latest Danish study which had at least 6 years records prior to the index date (ref. 6).

Interventions

"The observational period for assessing exposure for each patient will be defined as the last year before the index date.

No study has ever reported an increased risk of VTE after few weeks cessation of use. Considering all women with a prescription of hormonal contraceptives within one year before the index date as exposed women will therefore severely underestimate the risk of VTE with use of hormonal contraceptives.

"For the main analysis, the cases will include all patients with Read codes for VTE in their GP records or with death certificate codes."

So no diagnostic validation criteria are applied?

"A participant will be considered as exposed if they had at least one prescription for a composition containing drospiredone or cyproterone."

Later the protocol states that users of all 2nd, 3rd, 4th generation oral contraceptives as well as users of NuvaRing and patches are included. This is inconsistent.

"As some women might have switched between contraceptive types, overall exposure will be assessed. Exposure to other hormonal contraceptives of at least one prescription during the observation period will be included in the analysis."

This is a vague description of how to handle switchers. A more precise description is needed. E.g. a women experiencing a VTE 3 weeks after a switch from a 2nd to a 4th generation pill, to which group will such a woman be allocated?

	"The main focus will be on drospirenon- and cyprindiol-containing compositions. These drugs will be compared with the most-used compositions containing levonorgestrel, desogestrel, norgestimate and other progestogens (norethistiron and gestogene). Progestogen-only drugs are not expected to be associated with an increased risk of VTE[5] but will be kept in the analysis for comparison purposes." Non-oral hormonal contraceptive products are also included (according to later paragraphs). *"The effect of duration will be assessed by calculating the number of days prescribed within the previous year. If the gap between the end of one prescription and the start of the next is not more than 30 days, use will be considered as continuous and the duration of the prescriptions will be summed. If a gap between prescriptions is more than 30 days only the latest period of exposure will be considered. The duration will be categorised: no use in last year; 1-30 days; 31-90 days; 91-180 days; 181-365 days. A trend test will be performed using the actual number of days."* It is meaningless to categorise a women having used a pill for three months 9 months ago in the same category as a current user having used the pill for three months. The length of use analysis should be restricted to current users at the index date. *"Recency of use will be analysed by calculating the gap between estimated date for the last use and the index date, and categorising it by days before the index date as: the index date precedes or coincides with the date of last use (current use); last use between 1 and 30 days before the index date (recent use); last use between 31 and 90 days (remote use), last use between 91 and 365 days (past use); no use in last year. A trend test will be performed using the actual number of days."* No one will expect an increased risk of VTE in the groups of women having used hormonal contraception more than 30 days ago. *"New users of hormonal contraceptives within the last year will also be identified and the start time relative to the index date will be investigated in the analysis, categorised as: started within last 90 days, 91 to 180 days, 181 to 365 days, previous user (started more than 365 days ago), and no use in the last year."* Thus new users may have used oral contraceptives for more than 10 years, if that use was prior to the latest year? *"As an association of increased risk with VTE in transdermal versus oral contraceptive users has been found[19], the route of administration will also be analysed – oral or long-acting reversible contraceptives (IU devices, IU systems, injectable contraceptives, patches, rings and implants)."* This information is in contrast to earlier indications. Why not simply state that the risk of VTE In users of all types of hormonal contraceptives are included and categorised according to oestrogen dose, progestogen type, and route of administration? **Confounding factors** *"All analyses will include a priori confounders established as risk factors for VTE[20] and measured at the closest date before the index date, these are: body mass index (BMI) (continuous variable);*

	smoking status (current smoker: light 1 to 9 cigarettes/day, medium 10 to 19, heavy 20 or more; ex smoker; non smoker); alcohol consumption; ethnicity (White, Black, Asian, Other)[21]." BMI has in none of the studies on oral contraceptives and VTE been found to be confounder. In the study of Parkin et al (ref. 13) analysing the CPRD data, BMI was not a confounder. So far these variables have been found to be risk factors, but not confounders. *"As there is a large group of women taking hormonal contraceptives for treatment of polycystic ovary syndrome, acne, hirsutism and menstrual disorders, these conditions will also be included as important confounders because of associations with increased risk of VTE[14]."* It's fine to include PCOS, hirsutism and acne as potential confounders, but so far nobody has demonstrated these variables to be *important* confounders. *"Other potential confounders will be included if they change the odds ratio for the exposure variables by more than 10%. The list of additional potential confounders will contain socio-economic status (Townsend score in fifths) and co-morbidities associated with increased risk of VTE[22]: cancer; congestive cardiac failure; varicose veins; cardiovascular disease; rheumatoid arthritis, chronic renal disease; asthma, chronic obstructive pulmonary disease; Crohn's or ulcerative colitis; a family history of VTE and coagulation disturbances (Leiden factor V, protein C and S deficiencies)[23]. A number of medical conditions will also be taken into account if recorded in the last 6 months prior to the index date: acute infections; surgery; hospitalisation; and leg or hip fracture [22, 24]."* How will the authors get information about family disposition? Adjustment for each of these many conditions is not likely to change the risk estimates by 10% or more. But adjustment for all these potential confounders at the same time could change the estimates substantially. **Statistical analysis** Seems to be appropriate. *"A 1% level of statistical significance will be used to allow for multiple comparisons. Stata v 12 will be used for all the analyses."* Due to the expected relatively small sample of cases, a five percent significance level will be more appropriate. **Sample size calculation** *"All eligible cases from QResearch and CPRD will be used. According to Office for National Statistics, combined contraceptives are used by 25% of women aged 15 to 49 in the UK.[29] In order to obtain 90% power to detect a clinically important odds ratio of 1.5 at significance levels of 1% for an exposure prevalence of 25%, 584 cases and 2920 controls will be required. For an individual drug with exposure of 5%, 2115 cases and 10575 controls will be needed.* The expected number of included cases should be calculated. **DISCUSSION** *"This is an observational study based on routinely collected data from large primary care research databases and will have the strengths and limitations common to all such studies. It will be larger*

	and will have greater statistical power than previous studies." According to the information from the previous CPRD study and the information in this protocol, 2 x 100 cases with idiopathic VTE are expected, perhaps the double or 400 including also non-idiopathic VTE. In comparison the Danish studies included more than 4000 women with VTE. *"As the data on prescriptions and potential confounding variables are routinely and prospectively collected and recorded before the index date, the study will be free from recall bias, and, as all eligible cases and randomly selected controls will be included, this will ensure there is no selection bias."* It is a little surprising that the authors indicate that no selection bias could be in effect in their design, as a comparative Danish design is claimed to imply selection bias despite also including all eligible cases in the whole country, and without missing exposure information. *The limitations of the study will include possible uncertainty in VTE diagnosis. Lawrenson et al looked specifically at VTE validation and found that 84% of the diagnoses were supported by hospitalisation or death certificate.[31] Any misclassification (assuming it is non-differential between cases and controls) will result in underestimation of associations with hormonal contraceptives, shifting the odds ratios toward unity."* This circumstance is important, and confirmed by Danish data, where un-confirmed VTE diagnoses were increased by 65% and confirmed by 100 % comparing 4th with 2nd generation products. **Conclusion** The planned study will add new data to the issue about the risk of venous thromboembolism in users of different types of hormonal contraception, although including partly previously published data from the CPRD. The authors generally over-emphasize potential problems in previous studies, and underestimate the potential methodological problems in their own design. The protocol should be revised according to the suggestions made, so that contradicting passages are brought in mutual accordance. It is unlikely, that this new study will provide results differing substantially from recent large-scale studies from other countries, and from studies conducted with the same data source (CPRD). Information about sponsors should be declared.

VERSION 1 – AUTHOR RESPONSE

Reviewer Name Susan Jick DSc

This protocol by Vinogradova et all describes a study to evaluate newer hormonal contraceptives in relation to VTE compared to other hormonal contraceptive users. The stated objective of the study is to overcome prior study biases and limitations and to update information on risk for drospirenone and cyproterone containing OCs. (Note: though I would not necessarily call cyproterone containing-OCs "newer" since it was marketed in the late 1990's, and they are no longer used in some countries such as France because of the established increased risk of VTE with these pills).

We have changed our analysis to include all commonly-prescribed drugs and removed the terminology of 'newer drugs'.

The title of the study describes a study of hormonal contraceptives (HCs) in relation to VTE yet the focus of the study is on oral contraceptives with a brief nod toward other HCs. This title should more accurately reflect the exposures discussed in the protocol. Also, they authors suggest that the cyprroterone-containing OCs were mostly marketed to treat severe acne, hirsutism, and PCOS (last paragraph page 5). Do they mean in addition to its use as a contraceptive? If not, perhaps it should not be included in this protocol that describes VTE in HCs.

We have added the word "oral" to the title to reflect our focus. We have justified the inclusion of cyproterone in Methods Exposure.

The Introduction contains some errors and some inaccurate descriptions of the prior literature: Reference 7 is not a Danish study as reported. It is from the Netherlands.

This is fixed.

The authors cite reference 8 and say that according to that paper the studies in references 5-7 have "serious" methodological limitations. In fact, the paper by Shaprio and Dinger (ref 8), only discusses 2 papers and these are not among the papers referenced (5-7). Reference 8 was published in 2010 and so could not have discussed the papers by Lidegaard et al (refs 5 and 6) which were published in 2011 and 2012. The third reference (7) was a paper by Vlieg et al. This paper was not discussed in the Shapiro/Dinger paper either.

I suggest that the authors of this protocol read the papers being critiqued by Shapiro and Dinger and consider the merits of the criticisms themselves. I would not accept the conclusions of Shapiro and DInger so readily. There is considerable disagreement in the epidemiology community on the proper methodology for the studies of hormonal contraceptives in relation to VTE and only a careful review of each study can properly reveal the true differences between them and identify which methodological differences would most likely explain the differences in the various study results. I would not accept the opinion expressed in this one review without careful consideration of the strengths and limitations of all the published studies.

Reference to this paper has been removed.

For References 10 and 11 the "main" limitations reported by the authors were that pregnant and postpartum women were not "identified". I respectfully suggest that these were hardly "the main" limitations compared to many others that would have had a much larger impact on the study results. To start, the Gronich study (ref 10) compared current second generation OC use to current drospirenone use. Pregnant women would not be current users of OCs, and thus there would have been no need to "identify" them in the study; thus not doing so would not be even a minor limitation.

The issue of postpartum use of combined OCs could have created minor bias but postpartum OC users would represent few women since relatively few women start combined OCs within a few months of giving birth. Thus this “main” limitation identified by the authors is likely not important. A more important limitation of both studies (10 and 11) is that they included women with other important proximate causes and risk factors for VTE such as recent surgery, lower limb injury and trauma, cardiovascular disease and cancer among others. Study ref 11 even included women with prior VTE. The small number of idiopathic cases in this study was likely the biggest limitation. The inclusion of non-idiopathic cases, which could have a big impact on the results, was never discussed by the protocol authors. I suggest that different case inclusion criteria are among the biggest differences between the studies referenced and that the inclusion of pregnant and postpartum women is a small portion of these women.

Our introduction now reflects difference in earlier study designs rather than a critique of approaches.

The inclusion of fatal cases is also not likely to have a big impact on the study results; 1) because not many young women die from a VTE and 2) because it is unlikely that the deaths would occur disproportionately in users of one OC rather than another. The authors have identified many limitations that would have minor if any impact on the study results and have neglected to discuss more important limitations such as case inclusion/exclusion criteria and selection of the reference group. Each published study has different limitations, some major and some minor, but they should not all be so quickly dismissed. Further, the study methods described in this proposal will not in fact address the important limitations of former studies, and the authors have not successfully identified the factors that have led to the differences in the previously published study results. That is, this study includes many of the “serious” limitations of some of the former studies. These are discussed below.sj

Our introduction now reflects difference in earlier study designs rather than a critique of approaches.

Methods How well has QResearch been validated? To my knowledge there is no way to link electronic patient data to original records so validation is not to the same scale as in CPRD. The reason I point this out is that there is opportunity for outcome misclassification from missing as well as incorrect diagnosis of VTE. If there are differences in results between the two databases this could be one explanation. It is important to acknowledge that the databases are not identical. The CPRD was developed with the express intention to collect data for research. This is not the case for EMIS data, which is the basis for the data in QResearch. A separate description of each database would be useful.

We have added details of a few publications which demonstrated the similarity of the databases.

The authors say this is a nested case-control study but it seems to include all women in the databases of the relevant age group. So I’m not sure why it is a nested study. They also say that use of a case-control design produces “unbiased estimate”. This is only true if the case definition and control selection are appropriate as well as the selection of the exposure and referent. I do not know what the authors mean when they say that case-control studies produce unbiased estimates.

We are identifying a cohort of patients and the case-control study which will be used for the analyses is nested within this cohort. We have changed our wording to clarify this.

Under the heading Cases and controls: the authors describe the study design. Perhaps the heading should be changed.

We have changed headings according to these comments.

They then go on to provide the definition of a case. Cases are defined as all eligible women with a diagnosis of VTE excluding only women with a prior VTE, women who are pregnant or within 3 months of a pregnancy, and women with anticoagulation therapy more than 6 weeks prior to the VTE date. This is a very loose definition of VTE and will likely include many non-cases. It is easy to require a prescription for an anticoagulant to confirm the VTE diagnosis and I would recommend adding this requirement to the case definition to avoid misclassification. In addition, particularly because the authors plan to include non-exposed cases in this study, it is important to exclude women with prior cardiovascular disease and other important risk factors for VTE such as cancer, renal disease, etc., from the study since oral contraceptives are contraindicated in these women.

We have added a table of the READ VTE codes. The sensitivity analysis with cases supported by thrombolytic prescriptions should address this issue. We have defined idiopathic and non-idiopathic cases and will run separate sensitivity analyses on these groups.

The authors indicate that they will use ONS mortality data to identify fatal cases of VTE. It should be noted that, at least in the CPRD, these data are only available for a portion of all practices (those in England). Practices in Wales, Scotland and Ireland are not linked to mortality data. This may be true for QResearch data too. The same is true for the HES data.

As the number of additional cases identified through ONS Mortality data might be very low and with only partial coverage of CPRD patients, we have decided to remove this link from the study. HES data will be used only for sensitivity analysis and only on practices which are linked to HES.

Under control matching, the authors say that they will match controls to cases on age (please specify within 1 year, 2 years, etc.), and on calendar year. They then say in the next sentence that controls will be allocated the same index date as their matched case, so in fact they will be matched on the actual index day, not on calendar year.

We have removed the confusing term “on calendar year” and added that controls will be matched on the same year of birth.

Interventions: Since this is an observational study, not a clinical trial I would suggest using the term “Exposure” rather than “intervention” since intervention implies that the researcher will have some control over the exposure allocation which they will not.

We have changed this.

It is not necessary to repeat the case definition here.

This has been removed.

Exposure = at least 1 Rx for a combined pill containing either drospirenone or cyproterone in the year prior to the index date. “Overall exposure” will be assessed to take into account switching. I am not sure what “overall exposure” means. If a woman switches from a drospirenone OC to a second generation OC how will her exposure be classified?

We have removed this and added that we will look at commonly-prescribed types of oral contraceptives. We have now defined how we intend to treat switching in the last paragraph of Exposure in Methods.

The authors list under recency of use different windows of exposure including current, recent, remote, past and nonuse in the past year. This implies that there will be a non-exposed referent category but it

is never actually specified what the referent will be. The authors also say on page 9, paragraph 2 that the main focus will be on drospirenone and cyproterone containing OCs versus the older HCs, so this leaves the reader confused.

We have added that all our analyses will have a reference category “no use in the last year”.

If the Comparison will be use of any other OC, including combined OCs with levonorgestrel, norethisterone, norgestimate, desogestrel, and gestodene (as they indicate on page 9), the choice of reference group will bias the results toward the null. Many of these comparison OCs have been shown to increase the risk of VTE compared to second generation OCs, so including these in the referent will increase the baseline risk of VTE in the comparison group and yield lower relative risks for the drospirenone and cyproterone OCs in comparison.

Our comparison group is “no use in the last year”.

As I read it, duration of use will be evaluated without regard to whether the use was current at the index date. This will underestimate any true effect of duration since non-current use has been shown not to be associated with an increased risk of VTE.

We will evaluate duration of use only for current users.

In the last paragraph of the section titled Intervention, the authors briefly mention that use of the contraceptive patch will be analyzed as will be other HCs (IUDs, rings, etc.). These all have different exposure considerations (how use is determined, duration and recency of use, etc.). This should either be described more completely or left out of the study. The choice of referent group for these HCs is also missing here.

We have removed this.

Confounding Factors: Many covariates will be included and evaluated as confounders. Family history of coagulation defects however will not be recorded in any complete or systematic way (at least not in the CPRD) so I do not think it will be feasible to include this covariate.

We have reworded this.

The authors mention that they will “take into account” several other medical conditions if they occur in the 6 months prior to the index date: acute infection, surgery, and leg or hip fracture. It is notable that they have left these most important other proximate causes (surgery and limb injury) to the very last sentence of the Confounding section. It is also notable that major trauma is not noted here.

Major trauma is likely to be followed by hospital admission. We are unable to distinguish admissions because of VTE and admissions because of other events – a great number of events associated with hospital admission have READ codes such as “Emergency hospital admission” or “Other hospital admission NOS”.

These are by far the most common causes of VTE in relatively young and healthy women, which characterizes women taking hormonal contraceptives. Because these are such strong risk factors for VTE it is unlikely that there will be any differential effects of different hormonal contraceptives (HC) on the VTE risk in women who have these other proximate causes for their VTE. If this is the case, then the distribution of HC use in these cases should be the same as the distribution in the controls which would lead to a null effect comparing different HCs even if there are true differences in the risks of the HCs. I would suggest excluding cases with other proximate causes for their VTEs in the 3 (or 6)

months prior to the VTE. At the very least, the analyses should be stratified according to idiopathic and non-idiopathic case status (cases where a proximate cause is present). It is not sufficient to conduct a sensitivity analysis at the end when there is a strong likelihood that this will be an important source of bias.

We have added sensitivity analyses for both idiopathic and non-idiopathic cases.

Statistical analysis: The authors state that they will use conditional logistic regression to estimate odds ratios, but they never specify the referent group. I think it is important to be clear what this referent will be. They can use multiple referents but they need to be specified.

We have clarified this.

The authors state that a sensitivity analysis will be conducted on the subgroup of patients with an anticoagulant code in the 6 weeks post VTE. As indicated above, I think all analyses should be restricted to these patients. Anticoagulation is a known and necessary treatment for VTE and it is almost inconceivable that a patient with a true VTE would not be treated with anticoagulation. To include non-treated patients would invite misclassification.

We have justified in the reworded introduction our decision to run the main analysis on the whole sample and the sensitivity analysis on the confirmed cases and their controls.

I would suggest conducting analyses stratified by age to test for effect modification by age since increasing age is associated with increasing risk of VTE and decreasing use of OCs.

We intend to do this by looking at older and younger patients, and will test for effect modification by age.

Sensitivity analysis: Restricting to HES, ONS: As stated above it should be clear that a large proportion of patients will not be linked to HES or mortality data.

We have added this.

Sample Size: Drospirenone and cyproterone OCs were never heavily used in the UK and thus the number of users in both the CPRD and QResearch will not be as great as for other OCs. Since these are the purported exposures of interest the number of users of these 2 OCs should be used in the power calculation, not the total number of HC users.

We have added more information in the Sample size section.

Limitations: It is a limitation that some women receive contraception at family planning clinics and their HC use may be missed in these data. If the study were restricted to current HC users then this would not be a concern but since the authors plan to compare OC use to non-use there may be users included in the non-use category. It is not clear why the authors have chosen to have a non-exposed category (if in fact they plan to include the non-use exposure in their analyses). According to French RS, Mercer CH, Johnson AM, et al. (Use of contraceptive services in Britain: findings from the second National Survey of Sexual Attitudes and Lifestyles (Natsal-2). J Fam Plann Reprod Health Care 2009;35(1):9-14.), approximately 15% of women utilize clinics for contraception,

We have added this useful information to the text.

++++++++++++++++

Reviewer Name Øjvind Lidegaard

Abstract
Introduction: “Many studies have found an increased risk of venous thromboembolism associated with the use of hormonal contraceptives, but these have been against a background of evolving drug technology and all have been subject to various biases and methodological limitations.”
Nowhere are the claimed biases and methodological limitations in these studies documented.

We have removed this sentence.

“This study will focus on newer hormonal contraceptives – drugs containing drospirenone and cyproterone”.
While drospirenone has been 12 years on the market, oral contraceptives with cyproterone have now been on the market for several decades. Therefore, the focus of this study is not on newer hormonal contraceptives, but on an old combined pill and on one 12 years old combined pill. According to the protocol, all types of hormonal contraception are included. This should be stressed in the abstract.

We have changed this.

Methods and analysis.
Expected number of included cases is not indicated.
Exposure is defined as “at least one prescription of hormonal contraceptives in the year before the index date.” This will severely underestimates the risk with use of hormonal contraception, as use longer back in time than few weeks are unlikely to increase the risk of VTE.

We are going to analyse the associations with recency of use and duration, and these will be the main results of the study.

Strenghts and limitations
“No recall or selection biases.” No one has delivered any evidence that previous studies were affected by selection bias.

We have removed this.

“Possible uncertainty in diagnosis of venous thromboembolism.”
Uncertainly in some diagnoses of venous thromboembolism is a clinical condition no study design can remove. What can be done, and what should be done is do define a priori which criteria a certain case has to fulfil to be considered as a case. Such criteria are not indicated.

We have added a list of READ codes for VTE.

“Underestimation of hormonal contraceptive use”.
The fact that a significant proportion of women get their hormonal contraceptive product from contraception clinics and thereby will be misclassified as non-users will certainly underestimate the risk of VTE from use of hormonal contraceptives, but not necessarily the difference in risk between different product groups.

We agree and have mentioned this limitation in the Discussion.

Key words

The following key words were suggested: “bisphosphonates, neoplasms, case-control studies, osteoporosis/ drug therapy, risk factors.”
None of these key words are appropriate. Use instead: Oral contraceptives, hormonal contraception, progestogen only contraception, venous thromboembolism, PCOS.

We have changed these.

Introduction
“These associations have no established biological explanations, however, beyond speculation that third-generation progestins may have potentiated oestrogenic effects on clotting factors[1].”
This is not correct. Several studies have demonstrated a higher increase in SHBG and an increase in activated protein C with those products implying the highest risk of VTE.

We have removed this.

Other comments below relating to the Introduction are not addressed in detail as we have completely rewritten this section to take account of these by describing their differences in design rather than undertaking a critique of these.

“Currently, commonly-used fourth-generation pills, specifically those containing drospirenone (introduced in 2002) rather than levonorgestrel, have also been shown to have associations with increased risk of VTE in three large studies based on a general female population with data from the Danish Registry[5-7], but these studies had serious methodological limitations including lack of adjustments for confounding factors, selection bias and confounding by indication[8].”
First, reference 7 was a Dutch study demonstrating the same increased risk of VTE with use of 4th generation pills as the two Danish studies (ref. 5 and 6). Secondly, the company sponsored authors of reference 8 did not claim the Danish studies to be influenced by selection bias or confounding by indication. Other critique points have been refuted by the authors of these studies (Lidegaard Ø. Critique of a Danish cohort study on hormonal contraception and VTE. J Fam Plann Reprod Health Care
2010; 36: 103-4.pdf and Lidegaard Ø. Reply to Jürgen Dinger and Samuel Shapiro. BMJ 2011; online December 12,
2011.)

“Two of those studies, using an Israeli database[10] and USA insurance data[9], identified 165 and 18 VTE cases on drospirenone and reported an increased risk compared with other hormonal contraceptives. The main limitation for both studies, however, was no identification of pregnant or post-partum cases at the time of the diagnosis”
The circumstance that these studies did not exclude pregnant and post-partum cases would have underestimated the risk of VTE in current users of drospirenone, and not the opposite.

“A German study[11]based on 26 drospirenone users showed no difference with the other hormonal compositions – the study included pregnant women but failed to adjust for this condition”
They also failed to exclude other predisposed women such as women with previous thrombosis, known thrombophilia, cancer etc, explaining why the only two studies not demonstrating a difference between 2nd and 3rd/4th generation pills (both by Dinger et al) did not do so.

Two other studies from the same researchers [12, 13] excluded pregnant and post-partum women from the analysis and showed increased VTE risk associated with contraceptives that contained drospirenone rather than levonorgestrel. The one based on CPRD data[13] showed a 3-fold increased risk, but the precision of estimates was limited as only 17 exposed cases were identified.

These authors included CPRD data covering the period 2002-2009 and found 61 idiopathic VTE cases. The proposed new study covering the period 2001-July 2013 is thus expected to include about 100 women with idiopathic VTE from the CPRD research database.

“The one[12] based on American insurance data was the only study of the five, which adjusted for menstrual disorders – one of the indications for hormonal contraceptives and, as shown in the paper, associated with increased risk of VTE. Although the study reported a 2.4-fold increased risk among drospirenone users, it was conducted only on participants who survived.”
With a case-fatality rate of about 1% among young women with VTE the exclusion of fatal VTE in this study is very unlikely to have biased the results of this study.

“Another hormonal contraceptive containing cyproterone has been used mostly for treatment of women with severe acne, hirsutism and polycystic ovary syndrome (PCOS). It has been shown that VTE risk is higher for such groups of women[14] but, of the two studies of hormonal contraceptives and VTE risk which adjusted for PCOS, one analysed only twelve exposed cases[15] while the other did not include cyproterone-containing contraceptives in the analysis[14].”
Thus the results for drospirenone containing oral contraceptives in study [14] were adjusted for PCOS and still significantly higher than for 2nd generation products.

“Three large population-based studies, which did include cyproterone [3, 5, 7], omitted to adjust for PCOS, so their results might be subject to confounding by indication.”
Yes, PCOS is a potential confounder, but it is still unlikely that preferential prescribing of 3rd and 4th generation products to women with PCOS can explain the increased risk of VTE in users of these products.

“The proposed nested case control studies based on the female general population will investigate the association between the use of hormonal contraceptives – classified by type and use – and risk of VTE adjusted for PCOS and menstrual disorders as the major possible indications, other comorbidities
and concomitant drug exposure. It will concentrate on the most recent compositions
and increase its power by combining the results obtained from the two of the largest electronic medical records databases, QResearch and CPRD analyses.”
No information is given on how the authors will identify and verify the PCOS women and women with menstrual disorders. Menstrual disorders (e.g. dysmenorrhea) have never been demonstrated to be associated with VTE except if combined with PCOS.

METHODS AND ANALYSIS
Cases and controls
“Eligible cases and controls will have at least two years of records prior to the index date.”
This is one third as compared with the latest Danish study which had at least 6 years records prior to the index date (ref. 6).

We have changed the restriction on the amount of medical information to “at least 1 year of medical records”. This is because the risk of VTE decreases within a few months after starting a contraceptive drug and disappears after stopping it within 3 months.

Interventions
“The observational period for assessing exposure for each patient will be defined as the last year before the index date.
No study has ever reported an increased risk of VTE after few weeks cessation of use.
Considering all women with a prescription of hormonal contraceptives within one year before the

index date as exposed women will therefore severely underestimate the risk of VTE with use of hormonal contraceptives.

We are also looking at the effect of duration and time since the last use.

“For the main analysis, the cases will include all patients with Read codes for VTE in their GP records or with death certificate codes.”
So no diagnostic validation criteria are applied?

We have justified this approach in the rewritten Introduction and we shall also run a sensitivity analysis restricted to cases with anticoagulant therapy.

“A participant will be considered as exposed if they had at least one prescription for a composition containing drospiredone or cyproterone.”
Later the protocol states that users of all 2nd, 3rd, 4th generation oral contraceptives as well as users of NuvaRing and patches are included. This is inconsistent.

We have removed this.

“As some women might have switched between contraceptive types, overall exposure will be assessed. Exposure to other hormonal contraceptives of at least one prescription during the observation period will be included in the analysis.”
This is a vague description of how to handle switchers. A more precise description is needed. E.g. a women experiencing a VTE 3 weeks after a switch from a 2nd to a 4th generation pill, to which group will such a woman be allocated?

We have clarified how we are going to treat switchers in Exposure in Methods.

“The main focus will be on drospirenon- and cyprindiol-containing compositions. These drugs will be compared with the most-used compositions containing levonorgestrel, desogestrel, norgestimate and other progestogens (norethistiron and gestogene). Progestogen-only drugs are not expected to be associated with an increased risk of VTE[5] but will be kept in the analysis for comparison purposes.”
Non-oral hormonal contraceptive products are also included (according to later paragraphs).

We will keep them as a category in the analysis but will not perform a detailed analysis on them.

“The effect of duration will be assessed by calculating the number of days prescribed within the previous year. If the gap between the end of one prescription and the start of the next is not more than 30 days, use will be considered as continuous and the duration of the prescriptions will be summed. If a gap between prescriptions is more than 30 days only the latest period of exposure will be considered. The duration will be categorised: no use in last year; 1-30 days; 31-90 days; 91-180 days; 181-365 days. A trend test will be performed using the actual number of days.”
It is meaningless to categorise a women having used a pill for three months 9 months ago in the same category as a current user having used the pill for three months. The length of use analysis should be restricted to current users at the index date.

We have changed this.

“Recency of use will be analysed by calculating the gap between estimated date for the last use and the index date, and categorising it by days before the index date as: the index date precedes or coincides with the date of last use (current use); last use between 1 and 30 days before the

index date (recent use); last use between 31 and 90 days (remote use), last use between 91 and 365 days (past use); no use in last year. A trend test will be performed using the actual number of days."
No one will expect an increased risk of VTE in the groups of women having used hormonal contraception more than 30 days ago.

We have removed this from the analysis.

"New users of hormonal contraceptives within the last year will also be identified and the start time relative to the index date will be investigated in the analysis, categorised as: started within last 90 days, 91 to 180 days, 181 to 365 days, previous user (started more than 365 days ago), and no use in the last year."
Thus new users may have used oral contraceptives for more than 10 years, if that use was prior to the latest year?

We have removed this from the analysis.

"As an association of increased risk with VTE in transdermal versus oral contraceptive users has been found[19], the route of administration will also be analysed – oral or long-acting reversible contraceptives (IU devices, IU systems, injectable contraceptives, patches, rings and implants)."
This information is in contrast to earlier indications. Why not simply state that the risk of VTE In users of all types of hormonal contraceptives are included and categorised according to oestrogen dose, progestogen type, and route of administration?

We are keeping women with non-oral contraceptive prescriptions but will not perform a detailed analysis of them.

Confounding factors
"All analyses will include a priori confounders established as risk factors for VTE[20] and measured at the closest date before the index date, these are: body mass index (BMI) (continuous variable); smoking status (current smoker: light 1 to 9 cigarettes/day, medium 10 to 19, heavy 20 or more; ex smoker; non smoker); alcohol consumption; ethnicity (White, Black, Asian, Other)[21]."
BMI has in none of the studies on oral contraceptives and VTE been found to be confounder. In the study of Parkin et al (ref. 13) analysing the CPRD data, BMI was not a confounder. So far these variables have been found to be risk factors, but not confounders
.
We are keeping body mass index and smoking in the analysis as they are the risk factors for VTE added the references (Gronich 2011 and Pomp 2008) showing that use of contraceptives might be different according to BMI level and smoking status.

"As there is a large group of women taking hormonal contraceptives for treatment of polycystic ovary syndrome, acne, hirsutism and menstrual disorders, these conditions will also be included as important confounders because of associations with increased risk of VTE[14]."
It's fine to include PCOS, hirsutism and acne as potential confounders, but so far nobody has demonstrated these variables to be important confounders.

We have changed the wording and will consider these variables as possible confounders.

"Other potential confounders will be included if they change the odds ratio for the exposure variables by more than 10%. The list of additional potential confounders will contain socioeconomic status (Townsend score in fifths) and co-morbidities associated with increased risk of VTE[22]: cancer; congestive cardiac failure; varicose veins; cardiovascular disease; rheumatoid

arthritis, chronic renal disease; asthma, chronic obstructive pulmonary disease; Crohn's or ulcerative colitis; a family history of VTE and coagulation disturbances (Leiden factor V, protein C and S deficiencies)[23]. A number of medical conditions will also be taken into account if recorded in the last 6 months prior to the index date: acute infections; surgery; hospitalisation; and leg or hip fracture [22, 24]."
How will the authors get information about family disposition?
Adjustment for each of these many conditions is not likely to change the risk estimates by 10% or more. But adjustment for all these potential confounders at the same time could change the estimates substantially.

We have changed this.

Statistical analysis
Seems to be appropriate.
"A 1% level of statistical significance will be used to allow for multiple comparisons. Stata v 12 will be used for all the analyses."
Due to the expected relatively small sample of cases, a five percent significance level will be more appropriate.

We are reporting 95% confidence intervals so our results will be comparable with other studies. We will report also p-values but will make our conclusions on the basis of a 1% statistically significant level as there will be many comparisons to do.

Sample size calculation
"All eligible cases from QResearch and CPRD will be used. According to Office for National Statistics, combined contraceptives are used by 25% of women aged 15 to 49 in the UK.[29] In order to obtain 90% power to detect a clinically important odds ratio of 1.5 at significance levels of 1% for an exposure prevalence of 25%, 584 cases and 2920 controls will be required. For an individual drug with exposure of 5%, 2115 cases and 10575 controls will be needed.
The expected number of included cases should be calculated.

We have added this.

DISCUSSION
"This is an observational study based on routinely collected data from large primary care research databases and will have the strengths and limitations common to all such studies. It will be larger and will have greater statistical power than previous studies."
According to the information from the previous CPRD study and the information in this protocol, 2 x 100 cases with idiopathic VTE are expected, perhaps the double or 400 including also nonidiopathic VTE. In comparison the Danish studies included more than 4000 women with VTE.

We have changed this.

"As the data on prescriptions and potential confounding variables are routinely and prospectively collected and recorded before the index date, the study will be free from recall bias, and, as all eligible cases and randomly selected controls will be included, this will ensure there is no selection bias."
It is a little surprising that the authors indicate that no selection bias could be in effect in their design, as a comparative Danish design is claimed to imply selection bias despite also including all eligible cases in the whole country, and without missing exposure information.

We have reworded this.

The limitations of the study will include possible uncertainty in VTE diagnosis. Lawrenson et al looked specifically at VTE validation and found that 84% of the diagnoses were supported by hospitalisation or death certificate.[31] Any misclassification (assuming it is non-differential between cases and controls) will result in underestimation of associations with hormonal contraceptives, shifting the odds ratios toward unity."
This circumstance is important, and confirmed by Danish data, where un-confirmed VTE diagnoses were increased by 65% and confirmed by 100 % comparing 4th with 2nd generation products.

We have added a sentence about differential attentions to different types of contraceptives.

Conclusion
The planned study will add new data to the issue about the risk of venous thromboembolism in users of different types of hormonal contraception, although including partly previously published data from the CPRD.
The authors generally over-emphasize potential problems in previous studies, and underestimate the potential methodological problems in their own design. The protocol should be revised according to the suggestions made, so that contradicting passages are brought in mutual accordance.
It is unlikely, that this new study will provide results differing substantially from recent large-scale studies from other countries, and from studies conducted with the same data source (CPRD).
Information about sponsors should be declared
Information about sponsors can be found in Funding section at the end of the paper. We do not have any specific funding for this project.

VERSION 2 – REVIEW

REVIEWER	Susan Jick Boston University School of Public Health United States
REVIEW RETURNED	20-Feb-2014

GENERAL COMMENTS	The authors have done an excellent job of responding to the reviewer comments and have made major revisions to their protocol. They have addressed most of the issues very well though there are still a few minor points remaining. At the top of the submission under the bullet for Strengths the authors have listed: Possibility of investigating the effect of recency of use. It has been well demonstrated that current use is the relevant exposure in the OC VTE relation. I do not think this needs to be revisited. In any case I would not list it as a notable strength. Two key methodological differences noted by the authors: 1) case validation 2) case definition including inclusion of non-idiopathic cases. The authors have not mentioned differences in exposure definition and reference group selection as a major difference. I think this under-emphasizes the importance of reference group selection and its impact on the interpretation of the varied published study results. In the discussion of switching in the Methods section the authors still do not adequately describe how they will handle switchers. May I suggest that they consider the subject exposed to the OC last received prior to the index date and then include a variable for whether or not they were a switcher in the model. This could be the

	simplest and most reasonable way to handle switching.

REVIEWER	Øjvind Lidegaard Dept. of Obstetrics and Gynaecology, Rigshospitalet, University of Copenhagen The reviewer has received honoraria for speeches in pharmacoepidemiological issues.
REVIEW RETURNED	22-Feb-2014

GENERAL COMMENTS	The study protocol has certainly been improved. Few remaining problems are identified however: **Abstract** This study will focus on common oral hormonal contraceptives, including rarer compositions with drospirenone and cyproterone. It is contradictory to focus on common products and then to include rarer compositions. Combined pills with drospirenone have been the best selling pill through several years. Therefore it is inappropriate to include it in "rarer compositions". I suggest that the authors Limitations are those related to a prescription-based study: Lack of information on risk factors such as air travel Air travel is certainly a risk factor for VTE. It is unlikely, however, to be a confounder, because that would demand users of different hormonal contraceptives to have differential flight habits, which is unlikely (when adjusted for eventual age differences). Therefore this limitation is unlikely to be a real limitation. **Introduction** In English synthetic progesterons are called progestogens. In USA the same products are termed progestins. I think the authors should use the English version. Effects on vascular risk factors, including VTE from third generation contraceptive use, have, however, been contradictory, with some increased risks reported contrary to the aims of the changed formulation[4]. The introduction of less androgenic combined oral contraceptives (3rd and 4th generation pills) were potentially expected to decrease arterial thrombotic complications, but not venous complications, because the former are due to arteriosclerosis while the latter are due to changes in the coagulation system. That's just for the orientation of the authors. Standardized criteria for diagnostic categories include four levels of verification: 1. imaging tests and subsequent therapy (i: definite VTE); 2. Dopler ultrasound or impedance plethysmography with subsequent therapy (ii: probable VTE) 3. and without therapy (iii: possible VTE), and 4. 'typical symptoms' with confirming tests or therapy (iv: potential

VTE)[5].

This is not a very logical or consistent classification. I suggest the following:

1. Positive imaging tests e.g. Positive Doppler ultrasound or impedance plethysmography and subsequent therapy (i: definite VTE);
2. Uncertain imaging tests with succeeding therapy (ii: probable VTE)
3. Positive imaging tests without succeeding therapy or
   Negative imaging tests but with succeeding therapy (iii: possible VTE)
4. 'typical symptoms' without confirming tests and without therapy (iv: potential VTE).

An Austrian study distinguished between confirmed and not confirmed cases, concentrating on cases with definite and probable VTE for the main analysis and performing additional analysis on the sample including possible and potential VTE cases, which produced statistically identical results to their main analysis.[6] ….. A Danish study based on national health care databases used anticoagulation prescriptions for verification and produced a stratified analysis of confirmed and non-confirmed diagnoses demonstrating a twofold to threefold higher risk associated with VTE in the confirmed group.[8]

One may stratify in two ways:

1) Analyses restricted to confirmed cases and 2) analyses restricted to non-confirmed cases.

Or you may

1) Include all confirmed cases and 2) All cases (confirmed and non-confirmed).

If you do the first thing (Danish study) you find significant differences. If you do the latter (Austrian study) you find small often non-significant differences.

The important thing here is that there is no contradiction between the results from the Danish and the Austrian studies; they have just used to different stratification strategies.

Further, although estimated VTE risk appeared increased, the authors of another meta-analysis[20] suggest that publication bias might shift such estimates towards increased risk.

It is unlikely that a good study demonstrating no difference in risk between different products should not be published. Contrary, those few (methodologically weak) studies demonstrating no differences according to progestogen type have been exposed several times more frequent at congresses, conferences etc. than larger and better studies demonstrating such a difference.

Methods

Page 10 line 6

"…….considered as exposed….."

	suggest "…..considered as ever exposed….." Page 10 line 8 "gestogene" should be spelled "gestodene" Table 1: Recurrent thrombosis (G4011 and G801G) should not be included, as women with previous thrombosis are excluded according to method section.

VERSION 2 – AUTHOR RESPONSE

Reviewer Name Susan Jick

The authors have done an excellent job of responding to the reviewer comments and have made major revisions to their protocol. They have addressed most of the issues very well though there are still a few minor points remaining.
At the top of the submission under the bullet for Strengths the authors have listed: Possibility of investigating the effect of recency of use. It has been well demonstrated that current use is the relevant exposure in the OC VTE relation. I do not think this needs to be revisited. In any case I would not list it as a notable strength.

We have removed this and added another strength regarding comparability with other studies.

Two key methodological differences noted by the authors: 1) case validation 2) case definition including inclusion of non-idiopathic cases. The authors have not mentioned differences in exposure definition and reference group selection as a major difference. I think this under-emphasizes the importance of reference group selection and its impact on the interpretation of the varied published study results.

We have added a paragraph describing the differences in definition of the exposure.

In the discussion of switching in the Methods section the authors still do not adequately describe how they will handle switchers. May I suggest that they consider the subject exposed to the OC last received prior to the index date and then include a variable for whether or not they were a switcher in the model. This could be the simplest and most reasonable way to handle switching.

We have changed the definition of the current exposure considering only the last received combined oral contraceptive and adding a variable for identifying switchers in the last month.

++++++++++++++++

Reviewer Name Øjvind Lidegaard

Reviewer's comments
The study protocol has certainly been improved. Few remaining problems are identified however:
Abstract
This study will focus on common oral hormonal contraceptives, including rarer compositions with drospirenone and cyproterone.
It is contradictory to focus on common products and then to include rarer compositions. Combined pills with drospirenone have been the best selling pill through several years. Therefore it is inappropriate to include it in "rarer compositions". I suggest that the authors

We regret that we have not the benefit of seeing your suggestion(!), but we have removed “rarer compositions with drospirenone” leaving only cyproterone with justification for inclusion.

Limitations are those related to a prescription-based study: Lack of information on risk factors such as air travel
Air travel is certainly a risk factor for VTE. It is unlikely, however, to be a confounder, because that would demand users of different hormonal contraceptives to have differential flight habits, which is unlikely (when adjusted for eventual age differences). Therefore this limitation is unlikely to be a real limitation.

We have removed this limitation.

Introduction
In English synthetic progesterons are called progestogens. In USA the same products are termed progestins. I think the authors should use the English version.

We have replaced progestins by progestogens.

Effects on vascular risk factors, including VTE from third generation contraceptive use, have, however, been contradictory, with some increased risks reported contrary to the aims of the changed formulation[4].
The introduction of less androgenic combined oral contraceptives (3rd and 4th generation pills) were potentially expected to decrease arterial thrombotic complications, but not venous complications, because the former are due to arteriosclerosis while the latter are due to changes in the coagulation system. That’s just for the orientation of the authors.

We have reworded the sentence to remove any misleading implication.

Standardized criteria for diagnostic categories include four levels of verification:
1. imaging tests and subsequent therapy (i: definite VTE);
2. Dopler ultrasound or impedance plethysmography with subsequent therapy (ii: probable VTE)
3. and without therapy (iii: possible VTE), and
4. ‘typical symptoms’ with confirming tests or therapy (iv: potential VTE)[5].
This is not a very logical or consistent classification. I suggest the following:
1. Positive imaging tests e.g. Positive Doppler ultrasound or impedance plethysmography and subsequent therapy (i: definite VTE);
2. Uncertain imaging tests with succeeding therapy (ii: probable VTE)
3. Positive imaging tests without succeeding therapy or
Negative imaging tests but with succeeding therapy (iii: possible VTE)
4. ‘typical symptoms’ without confirming tests and without therapy (iv: potential VTE).

We have adopted the suggested improved classification.

An Austrian study distinguished between confirmed and not confirmed cases, concentrating on cases with definite and probable VTE for the main analysis and performing additional analysis on the sample including possible and potential VTE cases, which produced statistically identical results to their main analysis.[6] ….. A Danish study based on national health care databases used anticoagulation prescriptions for verification and produced a stratified analysis of confirmed and non-confirmed diagnoses demonstrating a twofold to threefold higher risk associated with VTE in the confirmed group.[8]
One may stratify in two ways:
1) Analyses restricted to confirmed cases and 2) analyses restricted to non-confirmed cases.

Or you may
1) Include all confirmed cases and 2) All cases (confirmed and non-confirmed).
If you do the first thing (Danish study) you find significant differences. If you do the latter (Austrian study) you find small often non-significant differences.
The important thing here is that there is no contradiction between the results from the Danish and the Austrian studies; they have just used to different stratification strategies.

We have added an extra sentence to highlight that apparent contradictions may be the result only of strategic analytic choices.

Further, although estimated VTE risk appeared increased, the authors of another meta-analysis[20] suggest that publication bias might shift such estimates towards increased risk.
It is unlikely that a good study demonstrating no difference in risk between different products should not be published. Contrary, those few (methodologically weak) studies demonstrating no differences according to progestogen type have been exposed several times more frequent at congresses, conferences etc. than larger and better studies demonstrating such a difference.

We have removed this sentence as redundant.

Methods
Page 10 line 6
"…….considered as exposed….."
suggest "7..considered as ever exposed7.."

We have changed this as suggested.

Page 10 line 8
"gestogene" should be spelled "gestodene"

We have corrected this.

Table 1: Recurrent thrombosis (G4011 and G801G) should not be included, as women with previous thrombosis are excluded according to method section
We have removed these codes as redundant.