COCs containing dienogest and 30 μg ethinylestradiol may carry a higher VTE risk compared to corresponding preparations with levonorgestrel: A meta-analysis of four large cohort studies

Jürgen Dinger*

ZEG-Berlin Center for Epidemiology and Health Research, Invalidenstrasse 115, 10115 Berlin, Germany

Abstract

Background: The European Medicines Agency requested a meta-analysis of four large, multi-national cohort studies on hormonal contraceptives to clarify whether dienogest/ethinylestradiol-containing combined oral contraceptives (DNG/EE) carry a different risk of venous thromboembolic events (VTE) compared to levonorgestrel/ethinylestradiol-containing preparations (LNG/EE). The primary objective of the meta-analysis was to assess VTE risk in a study population that is representative for the actual users of the individual preparations.

Methods: All four studies were prospective, observational cohort studies. Cohorts consisted of new users of hormonal contraceptives: starters, switchers and restarters. Study participants were followed up for up to 10 years. The analysis was restricted to preparations containing 30 μg of ethinylestradiol. Primary risk measure: VTE hazard ratio (HR) in the European study population for DNG/EE versus LNG/EE adjusted for age, BMI, duration of current use, family history of VTE and data source.

Results: The analysis set included data from 228,122 users of hormonal contraceptives. The European study participants had used DNG/EE and LNG/EE (WY) for 38,708 and 45,359 woman years, respectively. The meta-analysis includes 102 VTEs: DNG/EE, 56 cases and 14.5 VTE/10,000 WY; LNG/EE, 46 cases and 10.1 VTE/10,000 WY. The primary analysis showed an adjusted HR for DNG/EE versus LNG/EE of 1.6 (95% confidence interval 1.1-2.3). Four alternative analyses showed similar results although only one of these analyses reached statistical significance.

Conclusion: DNG/EE is probably associated with a slightly higher risk of VTE compared to LNG/EE. However, some uncertainty regarding the validity of this result remains.

Abbreviations: BMI: Body Mass Index; CI: Confidence Interval; COC: Combined Oral Contraceptive; DNG: Dienogest; DNG/EE: Combined Oral Contraceptive containing Dienogest and 30 μg of Ethinylestradiol; EE: Ethinylestradiol; HR: Hazard Ratio; LNG: Levonorgestrel; LNG/EE: Combined Oral Contraceptive containing Levonorgestrel and 30 μg of Ethinylestradiol; VTE: Venous Thromboembolic Event; WY: Woman Years

Introduction

The safety of combined oral contraceptives (COCs) has improved over the years with the reduction in doses of estrogen and progestogen. However, concerns about COC safety have remained, peaking in the mid-1990s and early 2010s with discussion on whether COCs containing so-called "third" and "fourth" generation progestogens (desogestrel/gestodene and drospirenone, respectively) have a higher risk of cardiovascular side effects - especially venous thromboembolc events (VTE) - than older formulations [1-4].

Although the specific combination of 2 mg of dienogest (DNG) and 30 μg of ethinylestradiol (EE) has a substantial market share in Europe, published data on VTE risk are limited. It is currently scientifically unclear whether this specific combination (DNG/EE) is associated with a different risk of VTE compared to levonorgestrel (LNG)/EE-containing COCs which are often used as the reference standard for the VTE risk associated with combined hormonal contraceptives.

The Berlin Center for Epidemiology and Health Research conducted several large prospective cohort studies on the risk of VTE associated with the use of hormonal contraceptives. Four of these studies included a substantial number of women using DNG/EE or LNG/EE-containing COCs. The European Medicines Agency requested a meta-analysis of these four prospective cohort studies to clarify whether DNG/EE carries a different VTE risk compared to LNG/EE. Therefore, the data on DNG/EE and LNG/EE from the following four prospective cohort studies were combined: i) "Long-term Active Surveillance Study for Oral Contraceptives" (LASS) [5]; ii) "International Active Surveillance Study of Women Taking Oral Contraceptives" (INAS-OC) [6]; iii) "Transatlantic Active Surveillance on Cardiovascular Safety of Nuvaring" (TASC) [7]; and iv) "International Active Surveillance Study - Safety of Contraceptives: Role of Estrogens" (INAS-SCORE) [8].

*Correspondence to:** Jürgen Dinger, ZEG-Berlin Center for Epidemiology and Health Research, Invalidenstrasse 115, 10115 Berlin, Germany, Tel: +49 30 9451010; E-mail: j.dinger@zeg-berlin.de

Materials and methods

Details of design and methodology of the four cohort studies are described elsewhere [1-4]. All studies were conducted in accordance with the ethical principles of the Declaration of Helsinki. The primary ethical approvals in Europe for all four studies were provided by the ethical committee of the physicians' association in Berlin, Germany. The study outlines were published at ClinicalTrials.gov prior to the recruitment phase of the individual studies. Each study was governed by an independent Safety Monitoring and Advisory Council to ensure its scientific independence.

The primary objective of the meta-analysis was to assess the risk of VTE associated with the short and long-term use of DNG/EE and LNG/EE in a study population that is representative for the actual users of the individual preparations. The secondary objective was to characterize the baseline risk of users of the two formulations (lifetime history of co-morbidity, prognostic factors for VTE, co-medication, socio-demographic and life-style data).

The LASS study was conducted in Europe only; the other three studies were transatlantic studies that included a large proportion of European women. Study participants were recruited via large networks of hormonal contraceptive prescribing health care professionals in the United States and a total of 12 European countries. DNG/EE-containing COCs are not available in the United States. Thus, only European data were used for the meta-analysis presented here. The methodology used in the four studies was almost identical. All studies were large, prospective, observational, active surveillance studies that focused on the risk of VTE associated with the use of hormonal contraceptives. Cohorts consisted of new users of COCs: starters, switchers and restarters. A 'non-interference' approach was used to provide standardized, comprehensive, reliable information under routine medical conditions: i.e., all patients who were new users of an OC were eligible for enrolment if they gave their informed consent, and the physicians' prescribing behaviour was not influenced by quotas for specific OCs. Study participants were followed up for up to 10 years. All outcomes of interest were captured by direct contacts between the investigator team and the study participants. Inclusion and exclusion criteria, methods of patient recruitment, follow-up and data documentation (including prognostic factors for VTE) were almost identical. The data from these studies could therefore be combined for the planned meta-analysis without any substantial methodological problems.

Overall, the analysis set included data from 228,122 users of hormonal contraceptives with a follow-up of 736,793 woman years (WY) of observation. The European study participants had used DNG/EE and LNG/EE (preparations with 30 µg EE only) for 38.708 WY and 45.359 WY, respectively. The proportions of starters, switchers and restarters in the DNG/EE user group were 21%, 31% and 48%. The corresponding values for LNG/EE were 26%, 27% and 47%.

A low "loss to follow-up rate" was essential for the validity of all four studies. In order to minimize loss to follow-up the same multi-faceted, four-level follow-up process was established in all studies. Level1 activities included mailing the follow-up questionnaire and in case of no response two reminder letters. If level1 activities did not lead to a response, multiple attempts were made to contact the woman, her friends, relatives, and gynaecologist/primary care physician by phone. In parallel to these level2 activities, searches in national and international telephone and address directories as well as social networks were started (level3 activities). If this was not successful, an official address search via the respective governmental administration was conducted (in some countries centralized, in others decentralized at community level). This level4 activity usually yielded information on a new address (or information that the respondent had moved abroad or died). Overall, the loss to follow-up rate was 3.3% or lower in each of the four studies.

In all four studies the same procedures were used for the validation of reported adverse events. All serious adverse events and particularly VTEs were validated via the diagnosing and/or treating physician. All VTEs were checked at the end of the studies by three independent medical experts specializing in radiology/nuclear medicine, cardiology, and internal medicine/phlebology. For the blinded adjudication process, the brand names, dose, regimen and composition of the OC(s) used by the reporting woman were rendered anonymous. The adjudicators conducted their reviews independently of each other and without knowing the judgement of the other adjudicators or the investigators.

The primary risk measure was the VTE hazard ratio (HR) in the European study population for DNG/EE *versus* LNG/EE. In general, it is very difficult to interpret a relative risk of two or less in observational research [9,10]. Therefore, the author focused the analysis on excluding a twofold risk. Accordingly, the null hypothesis prior to the meta-analysis was: $HR_{VTE} > 2$ (i.e., the adjusted VTE hazard ratio for DNG/EE vs. LNG/EE is higher than 2). The alternative hypothesis was: $HR_{VTE} \leq 2$. However, regulatory authorities often request the exclusion of a 1.5-fold risk - even in non-experimental studies. The a priori power of the pooled analysis to exclude a twofold and 1.5-fold VTE risk for DNG/EE compared to LNG/EE was about 94% and 58%, respectively [11,12]. The corresponding values for the individual studies are shown in table 1. These power calculations confirm that the meta-analysis is sufficiently powered to show non-inferiority of DNG/EE compared to LNG/EE if a non-inferiority hazard ratio of 2 is accepted. If a non-inferiority hazard ratio of 1.5 is requested, the power of the pooled analysis is limited.

Inferential statistics were based on Cox proportional hazard models. Crude and adjusted HRs between the two cohorts of interest - DNG/EE and LNG/EE - were calculated. Four prognostic factors for VTE - age (continuous variable), BMI (continuous variable), current duration of use (continuous variable), and family history of VTE (binary variable) - were included as covariates in the Cox models. This selection was based on those factors that had consistently shown a substantial impact on VTE risk estimates in the regular statistical analyses of the individual studies. Furthermore, the data source (i.e. the study the women were participating in: LASS, INAS-OC, TASC or INAS-SCORE) was included in the Cox models for the pooled analysis.

To assess the robustness and validity of the primary statistical model, four alternative Cox models were used for sensitivity analyses: i) overall information on 20 potential confounders for VTE was available in all 4 studies (e.g., concomitant medication, smoking, acne), and the 20 potential confounders were included as covariates in a saturated Cox model; ii) starting with the 20 potential confounders for VTE a backward

Table 1. Power calculations for the individual studies and the pooled analysis

Study	Power to exclude a twofold risk	Power to exclude a 1.5-fold risk
LASS	84%	46%
INAS-OC	31%	16%
TASC	14%	9%
INAS-SCORE	31%	16%
Pooled analysis	94%	58%

stepwise procedure was used to reduce the number of covariates; all covariates that did not change the point estimate of the hazard ratio by more than 10% or that had no statistically significant impact (*p > 0.05*) were removed from the model in a stepwise procedure; iii) a Cox model with acne included in the primary model - in addition to the prognostic factors included in the primary model: this model was used to investigate the impact of the baseline differences between the two treatment groups; iv) a Cox model selected by the Akaike information criterion (AIC) as an estimator of the relative quality of different Cox models [13]: i.e. during a stepwise backward procedure (see model ii) the model with the lowest AIC value was chosen.

Regulatory authorities requested that the statistical analyses be conducted based on the "as treated" (AT) population as well as the "intention to treat" (ITT) population. For the AT analyses, data on outcomes of interest were assigned to the product actually used by the respective study participant at the time of the event. For the ITT analyses, all data from individual participants were assigned to the treatment they used at study entry, regardless of any switching (or stopping) or of any different (or no) product being used at the time of the event. For studies on efficacy, the "intention to treat" (ITT) approach is often preferred because it is conservative with respect to the superiority of a new treatment. For an analysis of drug safety, however, the ITT approach dilutes differences between treatments. Therefore, the "as treated" analysis was designated as the primary analysis for assessing the data.

The primary analysis was based on the pooled cohorts. As requested by regulatory authorities HRs were calculated per study as well as per user status (starter/switcher/restarter) for exploratory reasons. It should be noted that the total number of statistical tests is 25: 5 (4 studies plus pooled analysis) times 4 (complete cohorts, starters, switchers, restarters) plus 5 ITT analyses (complete cohorts in the pooled analysis and four studies). Accordingly, the likelihood of incorrectly rejecting a null hypothesis was substantial. Therefore, 24 out of the 25 tests were calculated for exploratory reasons only. Furthermore, a sub-analysis was only conducted if a minimum of three VTE was available for each of the two comparison groups.

Results

Overall, baseline characteristics were similar for the DNG/EE and LNG/EE users (Table 2). The similarities include age, weight, height, BMI, cardiovascular risk factors (e.g. family history of VTE), medical history and concomitant medication. A substantial difference was found for the prevalence of acne (30.0% and 9.1% for DNG/EE and LNG/EE, respectively). Slight to moderate differences were also found: i) a higher proportion of LNG/EE users had delivered a child or had been pregnant prior to study entry; ii) a higher proportion of DNG/EE users had switched OCs prior to study entry; and iii) DNG/EE users had a higher educational level.

Acne is associated with polycystic ovary syndrome (PCOS) which is associated with an approximately twofold risk of VTE [14-16]. Therefore, it was important that the sensitivity analyses included a Cox model that adjusted for acne (alternative model iii). A more detailed analysis of the age profile showed that the similar mean age of the two exposure groups is slightly misleading as the LNG/EE group included both, high proportions of teenagers and women age 30. Given the exponential increase of the VTE risk with age the LNG/EE users already had a slightly higher VTE risk prior to their enrolment. Therefore, it was to be expected that age adjusted VTE hazard ratios for DNG/EE versus LNG/EE would be slightly higher compared to crude hazard

Table 2. Baseline characteristics for the European users of DNG/EE* and LNG/EE*.

	DNG/EE	LNG/EE
Number of women	9.593 (100.0%)	13.309 (100.0%)
Patient characteristics		
Age at study entry (years)	25.1 (± 7.98)	25.5 (±8.52)
Age < 20	2.679 (27.9%)	3.868 (29.1%)
20 ≤ 30	4.444 (46.3%)	5.690 (42.8%)
30 ≤ 40	1.811 (18.9%)	2.605 (19.6%)
40 ≤ 50	634 (6.6%)	1.100 (8.3%)
50+	25 (0.3%)	46 (0.3%)
Weight at study entry (kg)	62.6 (±11.80)	63.5 (±12.43)
Height at study entry (cm)	167 (±6.2)	167 (±6.4)
BMI at study entry	22.4 (±3.95)	22.8 (±4.17)
Gynaecological history		
Age at menarche (years)	12.9 (±1.39)	13.0 (±1.40)
Gravidity (ever)	3.177 (33.1%)	5.317 (40.0%)
Parity (ever)	2.860 (29.8%)	4.838 (36.4%)
Number of live births	1.5 (±0.7)	1.6 (±0.8)
OC history		
OC user type: Starter	1.971 (20.5%)	3.194 (24.0%)
Switcher	3.016 (31.4%)	3.708 (27.9%)
Restarter	4.606 (48.0%)	6.407 (48.1%)
Ever switched OC	4.883 (50.9%)	5.284 (39.7%)
Duration of OC-use at study entry (years)	6.4 (± 6.08)	6.9 (± 6.79)
Acne		
Current acne	2.874 (30.0%)	1.207 (9.1%)
Cardiovascular risk factors		
High blood pressure	287 (3.0%)	420 (3.2%)
High cholesterol	204 (2.1%)	206 (1.5%)
Family history of ATE	131 (1.4%)	182 (1.4%)
Family history of VTE	567 (5.9%)	737 (5.5%)
Smoker	3.345 (34.9%)	5.245 (39.4%)
Heavy smoker (>15 cigarettes)	574 (6.0%)	903 (6.8%)
Medical history		
Diabetes	42 (0.4%)	10 (0.1%)
Cancer	46 (0.5%)	45 (0.3%)
Any surgery	3.042 (31.7%)	3.819 (28.7%)
Concomitant medication		
Regular use of medication	1.397 (14.6%)	1.784 (13.4%)
Education		
Higher than university entrance level	3.868 (40.3%)	4.691 (35.2%)

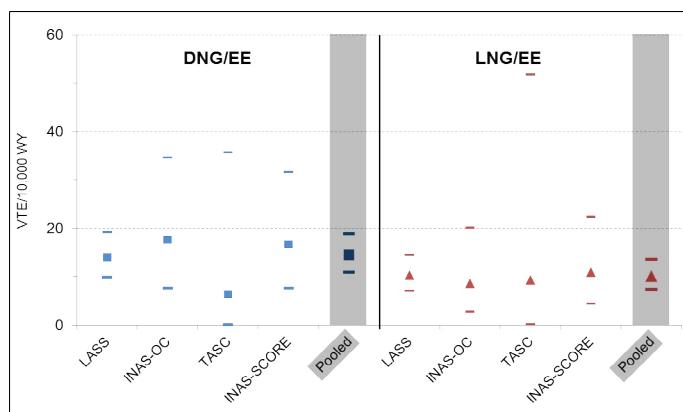*preparations with 30 µg ethinylestradiol only

ratios. Given the opposing effects of adjustment for acne and age, it was not expected that adjustment for the differences discussed above would result in a substantial change of the unadjusted VTE hazard ratio.

The meta-analysis is based on 102 VTEs: 56 and 46 VTEs occurred in the DNG/EE and LNG/EE exposure groups, respectively. Only preparations with the same EE-content (30 µg) were used for this comparison. The VTE incidence rates were higher for DNG/EE compared to LNG/EE (14.5 vs. 10.1 VTE/10,000 WY). The corresponding overall incidence rate ratio was 1.4; the 95% confidence interval (CI) included unity: 1.0 - 2.1. The results for the individual studies are shown in figure 1. Breaking down VTE into deep venous thrombosis (DVT) and pulmonary embolism (PE) showed similar results: DVT, incidence rate ratio 1.4 (95% CI, 0.9 - 2.2); PE, incidence rate ratio 1.4 (95% CI 0.6 - 3.1).

The primary analysis (pooled dataset for all users of DNG/EE or LNG/EE) resulted in an adjusted HR of 1.6 (95% CI, 1.1 - 2.3). The result
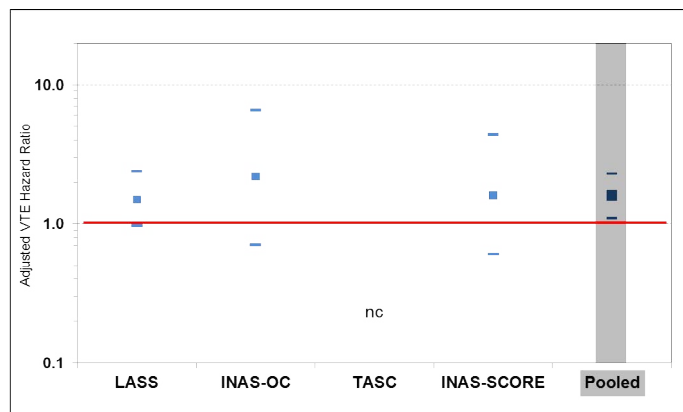
of the ITT analysis was similar: 1.5 (95% CI, 1.0 - 2.4). Also, the results for starters, switchers and restarters showed similar results: starters, 1.6 (95% CI, 0.6 - 4.6); switchers, 1.9 (1.0 - 3.4); and restarters 1.3 (95% CI, 0.7 - 2.5). A comparison of the individual study results shows consistent results (Figure 2). The huge LASS dataset had the strongest impact on the overall results. However, the data are sufficiently consistent across studies to justify combining of all four study databases.

The results of the four alternative Cox models are shown in table 3. The point estimates of the VTE hazard ratio were only slightly lower compared to the estimate from the primary model. The lower limits of the 95% confidence interval were close to one in all cases.. Formal statistical significance for DNG/EE versus LNG/EE was only reached for the primary model and alternative model ii (backward stepwise procedure).

* preparations with 30 µg ethinylestradiol only

Figure 1. VTE incidence rates and 95% CI for the individual studies and the pooled dataset: DNG/EE* versus LNG/EE*

nc not calculated because the minimum number of 3 VTE per exposure group was not reached

Figure 2. VTE hazard ratios and 95% CI for the individual studies and the pooled dataset: DNG/EE versus LNG/EE

Table 3. VTE hazard ratios: Comparisons between the primary Cox model and 4 alternative Cox models

Cox Model	Hazard Ratio	95% Confidence Interval		
Primary model	1.57	1.07	-	2.30
Alternative models				
i) Saturated model	1.47	0.98	-	2.20
ii) Backward stepwise	1.54	1.05	-	2.26
iii) Primary with acne included	1.44	0.97	-	2.13
iv) AIC selected model*	1.44	0.97	-	2.13

*identical with alternative model iii because the AIC selection process ended in a model that was identical with model iii.

Discussion and conclusion

The incidence rate for VTE was higher for DNG/EE compared to LNG/EE and the primary statistical analysis yielded an adjusted VTE hazard ratio for the comparison of DNG/EE versus LNG/EE of 1.6. The corresponding 95% CI did not include unity and suggested a slightly higher VTE risk for DNG/EE compared to LNG/EE. This result is supported in principle by the results of four alternative analyses although only one of these analyses reached statistical significance.

In non-experimental studies like LASS, INAS-OC, TASC and INAS-SCORE the possibility of bias and residual confounding can never be entirely eliminated, and the ability to infer causation is correspondingly limited [17]. Valid information on potential sources of confounding, and sophisticated statistical and epidemiologic methodology help to reduce the impact of bias and residual confounding [18]. However, the difficulty remains unresolved when all that exists is a weak association [19,20]. Relative risk estimates that are close to unity may not allow differentiation between causation, bias and confounding [21,22]. In general, it is very difficult to interpret a relative risk of two or less in observational research [9,10].

Selection and misclassification bias were probably not a major issue in any of the four cohort studies because i) their participants are representative for adult COC users [23]; and ii) reliable information on exposure and duration of OC use was available. Furthermore, the low loss to follow-up rates of 3.3% or less in all studies is noteworthy. In theory, a disproportionately high percentage of VTE could have occurred in those patients who were lost to follow-up, because VTEs could be the reason for the break in contact with the investigators. An advantage of the design of the four included studies, however, is that the investigator teams had direct contact with the participants; contact was not lost if the women changed their gynaecologists (e.g. due to change of residence or dissatisfaction with treatment).

In contrast, it was impossible to exclude diagnostic bias. Clinical symptoms of VTE cover the spectrum from a complete absence or unspecific, slight symptoms to dramatic, acute, life-threatening symptoms [24-26]. A high awareness of potential cardiovascular risks of combined oral contraceptive use might have led to more diagnostic procedures and therefore to more detected VTEs. However, this is a general consideration and there is no evidence that diagnostic bias influenced the results of this meta-analysis. Another issue is the fact that information on specific gene mutations was only available for VTE cases but not for the vast majority of study participants. This limitation was mitigated by information on family history of VTE which has a higher predictive value for VTE compared to gene mutations [27].

The studies included in this meta-analysis combine several methodological strengths that are substantial for the validity of the results such as: i) prospective, comparative cohort design; ii) availability of important confounder information (e.g. BMI and family history of VTE); iii) validation of outcomes of interest and exposure for the relevant cases; iv) comprehensive follow-up procedure and very low loss to follow-up to minimize underreporting; v) independent, blinded adjudication of VTE cases; vi) study population representative for oral contraceptive users under routine clinical conditions; vii) quite different statistical approaches resulting in similar risk estimates and 95% CIs support the validity and robustness of the primary statistical model; and viii) supervision by an independent Safety Monitoring and Advisory Council as well as scientific independence from the study funder.

The validity of synthetic meta-analysis as applied to non-experimental (observational) studies has been challenged [28-30]. Reasons why this approach is questioned include issues such as: variation in quality among studies; variation in methodology including variable definitions of exposure and outcome; and variable precision in the recording, measurement, and control of confounding factors. These issues do not apply to this meta-analysis. However, the possibility that a series of studies may tend to share the same biases and sources of confounding cannot be rejected for the studies included in this meta-analysis. The latter consideration is of special relevance when it comes to considering small associations [19].

In the author's judgment, the results of this meta-analysis are valid within the general limitations of observational research and meta-analysis of observational studies. The impact of residual confounding and bias could be limited to an extent that would allow causal interpretation of statistically significant results with hazard ratios that are equal to or higher than 2. Statistically significant results with hazard ratios of 1.5 to 2 can be cautiously interpreted as risk. However, some uncertainty remains in these cases.

The thresholds of 1.5 and 2 also reflect the limitations of the statistical power of this meta-analysis. The analysis was sufficiently powered to detect a twofold risk of VTE but had limited power to detect smaller risks. Nevertheless, the primary statistical analysis yielded a statistically significant increased risk of VTE for DNG/EE compared to LNG/EE. The adjusted hazard ratio of 1.6 is clearly below 2, and therefore some uncertainty regarding a causal interpretation of this result remains. The fact that the alternative analysis that adjusted for acne - the only baseline characteristic with a substantial difference between the exposure groups - showed no statistically significant difference adds to this uncertainty. It is also conceivable that adjustment for acne did not completely adjust for the likely differences in the prevalence of PCOS, and that some residual confounding remained unadjusted. However, it should be noted that primary and alternative analyses showed quantitatively similar risk estimates and the lower limits of the 95% CIs were always close to one.

Given the methodological strengths of the individual studies, the similarity of their study designs, and the quantitative consistency of the analysis results, the investigator considers it likely - but not definitively proven - that in 30 μg EE preparations DNG carries a higher risk of VTE compared to LNG.

Authorship

The meta-analysis was planned, conducted and reported by the author based on requests of the European Medicines Agency.

Acknowledgements

Funding information

Competing interests

References

1. World Health Organization Collaborative Study of Cardiovascular Disease and Steroid Hormone Contraception (195) Effect of different progestogens in low oestrogen oral contraceptives on venous thromboembolic disease. *Lancet* 346: 1582-1588.

2. Spitzer WO, Lewis MA, Heinemann LA, Thorogood M, MacRae KD (1996) Third generation oral contraceptives and risk of venous thromboembolic disorders: an international case-control study. Transnational Research Group on Oral Contraceptives and the Health of Young Women. *BMJ* 312: 83-88. [Crossref]

3. Lidegaard Ø, Løkkegaard E, Svendsen AL, Agger C (2009) Hormonal contraception and risk of venous thromboembolism: national follow-up study. *BMJ* 339: b2890. [Crossref]

4. Jick SS, Hernandez RK (2011) Risk of non-fatal venous thromboembolism in women using oral contraceptives containing drospirenone compared with women using oral contraceptives containing levonorgestrel: case-control study using United States claims data. *BMJ* 340: d2151. [Crossref]

5. Dinger J, Möhner S, Heinemann K (2016) Cardiovascular risks associated with the use of drospirenone-containing combined oral contraceptives. *Contraception* 93: 378-385. [Crossref]

6. Dinger J, Bardenheuer K, Heinemann K (2014) Cardiovascular and general safety of a 24-day regimen of drospirenone-containing combined oral contraceptives: final results from the International Active Surveillance Study of Women Taking Oral Contraceptives. *Contraception* 89: 253-263. [Crossref]

7. Dinger J, Möhner S, Heinemann K (2013) Cardiovascular Risk Associated with the Use of an Etonogestrel-Containing Vaginal Ring. *Obstet Gynecol* 122: 800-808. [Crossref]

8. Dinger J, Do Minh T, Heinemann K (2016) Impact of Estrogen Type on Cardiovascular Safety of Combined Oral Contraceptives. *Contraception* 94: 328-339. [Crossref]

9. Taubes G (1995) Epidemiology faces its limits. *Science* 269: 164-169. [Crossref]

10. Hill AB (1965) The environment and disease: association or causation? *Proc R Soc Med* 58: 295-300. [Crossref]

11. Chow SC, Shao J, Wang H (2008) Sample Size Calculations in Clinical Research. (2nd Edtn), Chapman & Hall/CRC. pp: 174-78.

12. Schoenfeld DA (1983) Sample Size Formula for the Proportional-Hazards Regression Model. *Biometrics* 39: 499-503. [Crossref]

13. Akaike H (1974) A new look at the statistical model identification. *IEEE Transactions on Automatic Control* 19: 716-723.

14. Bird ST, Hartzema AG, Brophy JM, Etminan M, Delaney JAC (2013) Risk of venous thromboembolism in women with polycystic ovary syndrome: a population-based matched cohort analysis. *CMAJ* 185: 115-120. [Crossref]

15. Okoroh EM, Hooper WC, Atrash HK, Yusuf HR, Boulet SL (2012) Is polycystic ovary syndrome another risk factor for venous thromboembolism) United States, 2003-2008. *Am J Oststet Gynecol* 207: 377. e1-377.e8. [Crossref]

16. Seaman HE, de Vries CS, Farmer RD (2003) Differences in the use of combined oral contraceptives amongst women with acne and without acne. *Hum Reprod* 18: 515-521. [Crossref]

17. Susser M (1991) What is a cause and how do we know one? A grammar for pragmatic epidemiology. *Am J Epidemiol* 133: 635-648. [Crossref]

18. Rothman KJ, Poole C (1988) A strengthening programme for weak associations. *Int J Epidemiol* 17: 955-959. [Crossref]

19. Shapiro S (2000) Bias in the evaluation of low-magnitude associations: an empirical perspective. *Am J Epidemiol* 151: 939-945. [Crossref]

20. Khoury MJ. James LM. Flanders WD. Erickson JD (1992) Interpretation of recurring weak associations obtained from epidemiologic studies of suspected human teratogens. *Teratology* 46: 69-77. [Crossref]

21. Shapiro S (2008) Causation, bias and confounding: a hitchhiker's guide to the epidemiological galaxy. Part 2. Principles of causality in epidemiological research: confounding, effect modification and strength of association. *J Fam Plann Reprod Health Care* 34: 185-190. [Crossref]

22. Shapiro S (2008) Causation, bias and confounding: a hitchhiker's guide to the epidemiological galaxy. Part 3: principles of causality in epidemiological research: statistical stability, dose- and duration response effects, internal and external consistency, analogy and biological plausibility. *J Fam Plann Reprod Health Care* 34: 261-264. [Crossref]

23. Dinger JC, Heinemann AJ, Kuehl-Habich D (2007) The safety of a drospirenone-containing oral contraceptive: final results from the European Active Surveillance study on oral contraceptives based on 142,475 woman-years of observation. *Contraception* 75: 344-354. [Crossref]

24. Scurr JH, Machin SJ, Bailey-King S, Mackie IJ, McDonald S, et al. (2001) Frequency and prevention of symptomless deep-vein thrombosis in long-haul flights: a randomised trial. *Lancet* 357: 1485-1489. [Crossref]

25. Belcaro C, Geroulakos G, Nicolaides AN, Myers KA, Winford M (2001) Venous thromboembolism from air travel: the LONFLIT Study. *Angiology* 52: 369-374. [Crossref]

26. Schwarz T, Siegert G, Oettler W, Halbritter K, Beyer J, et al. (2003) Venous thrombosis following long-haul flights. *Arch Intern Med* 163: 2759-2764. [Crossref]

27. Bezemer ID, van der Meer FJ, Eikenboom JC, Rosendaal FR, Doggen CJ (2009) The value of family history as a risk indicator for venous thrombosis. *Arch Intern Med* 169: 610-615. [Crossref]

28. Greenland S (1994) Can meta-analysis be salvaged? *Am J Epidemiol* 40: 783-787. [Crossref]

29. Shapiro S (1994) Meta-analysis/Shmeta-analysis. *Am J Epidemiol* 40: 771-778. [Crossref]

30. Shapiro S (1997) Is meta-analysis a valid approach to the evaluation of small effects in observational studies? *J Clin Epidemiol* 50: 223-239. [Crossref]