

Combined oral contraceptives containing dienogest and estradiol valerate may carry a lower risk of venous and arterial thromboembolism compared to conventional preparations: Results from the extended INAS-SCORE study

Jürgen Dinger*, Sabine Möhner and Klaas Heinemann

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

Abstract

Objectives: The previously reported initial results of the International Active Surveillance study “Safety of Contraceptives: Role of Estrogens” (INAS-SCORE) suggested that the risks of venous and arterial thromboembolism (VTE/ATE) of ethinylestradiol-containing combined oral contraceptives (COCs) might be higher compared to preparations containing dienogest and estradiol valerate (DNG/EV). Follow-up of study participants was extended to substantiate these findings.

Study design: Prospective, non-interventional cohort study conducted in the United States and seven European countries with two main exposure groups and one exposure subgroup: new users of DNG/EV and other COC (oCOC), particularly levonorgestrel-containing COCs (LNG/EE). All self-reported clinical outcomes of interest were validated via attending physicians and relevant source documents. The extension phase of the study focused on VTE and ATE. Comprehensive follow-up procedures were implemented. Statistical analyses were based on Cox regression models.

Results: A total of 50,203 COC users were followed up for 2 to 7 years. Overall, 20.3% and 79.7% of these women used DNG/EV and oCOC (including 11.5% LNG/EE users), respectively. A low loss to follow-up of 3.3% was achieved. Based on 68 VTE and 32 ATE the primary analyses (European dataset) yielded the following adjusted hazard ratios and 95% confidence intervals: DNG/EV versus oCOC, 0.4 (0.2-0.9) and 0.1 (0.0-0.5), respectively; DNG/EV versus LNG/EE, 0.4 (0.2-1.1) and 0.1 (0.0-1.0), respectively.

Conclusion: Results showed a lower risk of VTE and ATE for DNG/EV versus oCOC, and a similar or lower risk for DNG/EV versus LNG/EE.

Abbreviations: ATE: Arterial Thromboembolic Event; BMI: Body Mass Index; CI: Confidence Interval; COC: Combined Oral Contraceptive; DNG: Dienogest; DNG/EV: Combined Oral Contraceptive containing Dienogest and Estradiol Valerate; EE: Ethinylestradiol; EV: Estradiol Valerate; HR: Hazard Ratio; LNG: Levonorgestrel; LNG/EE: Combined Oral Contraceptive containing Levonorgestrel and Ethinylestradiol; oCOC: Other COC; OHC: Other Hormonal Contraceptives; VTE: Venous Thromboembolic Event; WY: Woman-Years

Introduction

As combined oral contraceptives (COCs) have been further developed over the past decades, their ethinylestradiol (EE) content has been reduced based on the hypothesis that lower EE doses lead to a better safety profile and specifically to a lower incidence of venous thromboembolism (VTE) [1]. However, reducing the EE dose led to a less favorable control of bleeding. Although EE has been used in numerous COCs, efforts have also been made to use estradiol and estradiol valerate (EV), which have a lower impact on the hepatic system and subsequently on hemostatic parameters [2]. In 2016 we published initial cardiovascular safety data from the “International Active Surveillance Study on the Safety of Contraceptives and the Role of Estrogens” (INAS-SCORE) which compared the safety profile of a new EV-based COC with established COCs containing EE [3]. We concluded that the EV-containing COC is associated with similar or even lower cardiovascular risk compared to EE-containing COCs –

including combinations of levonorgestrel (LNG) and EE. The point estimates of the hazard ratios (HRs) for VTE, arterial thromboembolic events (ATE) and other serious cardiovascular events were favorable for the new COC. These results were plausible because of the substitution of EE but the statistical power at that time was not sufficient to exclude chance as a possible explanation. Triggered by requests from the European regulatory authorities it was decided to prolong the follow-up of study participants for another two years. The cardiovascular data presented here reflect the analysis of the final study database.

The EV-containing regimen consists of four phases within a 26-day timeframe, with each phase containing different doses of EV, either alone or in combination with different doses of dienogest (DNG): 1) 2 tablets with 3 mg EV; 2) 5 tablets with 2 mg DNG and 2 mg EV; 3) 17 tablets with 3 mg DNG and 2 mg EV; and 4) 2 tablets with 1 mg EV. This sequence is followed by two days of placebo tablets.

***Correspondence to:** Jürgen Dinger, ZEG-Berlin Center for Epidemiology and Health Research, Invalidenstrasse 115, 10115 Berlin, Germany, Tel: 1719745433; E-mail: j.dinger@zeg-berlin.de

Key words: VTE, ATE, combined oral contraceptives, prospective cohort study, dienogest, estradiol valerate, routine clinical practice

Received: January 25, 2020; **Accepted:** February 17, 2020; **Published:** February 20, 2020

Materials and methods

We conducted a prospective, non-interventional cohort study of more than 50,000 new COC users to assess the occurrence of rare or unexpected adverse outcomes possibly related to COC exposure. The methodology and initial results of the INAS-SCORE study have been reported previously [3]. Here we present the results of the extension of the follow-up phase of the study.

The planning, conduct, and evaluation of the study were supervised by an independent Safety Monitoring and Advisory Council, which endorsed all the conclusions presented in this publication. The primary ethical approvals of the study in Europe and the United States were granted by the physicians' association in Berlin, Germany ("Ethik-Kommission der Ärztekammer Berlin"), and the Western Institutional Review Board (WIRB) in Olympia, WA, USA. The study is listed in the public clinical trials registry of the US National Library of Medicine under the identifier NCT01009684.

Study objectives

The primary objective of the study was to assess the risks of short- and long-term use of DNG/EV, other combined oral contraceptives (oCOC) and LNG/EE in a study population that is representative for the actual users of the individual preparations.

The main clinical outcomes of interest were serious cardiovascular events, in particular venous thromboembolism (VTE) such as deep venous thrombosis and pulmonary embolism, as well as arterial thromboembolic events (ATE) such as acute myocardial infarction and cerebrovascular accidents (CVA). In addition to VTE and ATE all other serious cardiovascular events (i.e., events that result in death, a life-threatening experience, inpatient hospitalization, persistent or significant disability/incapacity or require medical/surgical intervention to prevent one of these outcomes) were analyzed, including confirmed angina necessitating hospitalization, acute coronary syndromes, and congestive heart failure.

Study population

The study was conducted in the United States and seven European countries: Austria, France, Germany, Italy, Poland, Sweden, and the United Kingdom. Recruitment in Europe began in September 2009 and ended in October 2012. Because of the late market introduction of DNG/EV in the US, recruitment did not commence there until October 2010, and was completed in February 2013. Study participants in the US and Europe were followed until February 2015 and December 2016. The individual maximum follow-up therefore varied from 2 to 7 years.

Study participants were women who received a new prescription for a COC. Participating women could be starters (first-ever users of COCs), switchers (users who switched from one COC to another - without an intake break or with one of less than 4 weeks), or restarters (users who restarted a COC after an intake break of at least 4 weeks, i.e. at least one complete cycle). More specific inclusion or exclusion criteria were not introduced in keeping with the non-interference approach of the study design. Once enrolled, a subject could switch or discontinue use of the prescribed COC at any time. However, subjects continued to be followed whether or not they continued to use hormonal contraception, provided that they did not withdraw their consent.

The recruitment procedures for the study centres and participants were the same as those used in the EURAS-OC and INAS-OC studies, where they successfully yielded representative samples of typical COC

users with regard to age structure, socioeconomic and life-style factors, cardiovascular risk factors, the spectrum of prescribed OCs, and percentages of urban and rural COC users [4-9].

Data collection and quality control

Baseline data were recorded via self-administered questionnaires on participants' state of health, medical history including medication history and history of COC use, and potential prognostic factors for serious diseases, particularly cardiovascular disease. In addition, participants provided their addresses and phone numbers, as well as back-up contacts and contact information for their primary care physicians and/or gynaecologists. Baseline questionnaires were completed in the physicians' offices and checked by the physicians or their co-workers. Follow-up assessments for each woman were scheduled every 6 months for the first two years and annually thereafter. The follow-up questionnaires addressed the occurrence of adverse events - in particular serious adverse events and cardiovascular events. Reasons for discontinuing COC use or for switching to another hormonal contraceptive were requested if applicable.

A low loss to follow-up rate was essential for the validity of the study. In order to minimize loss to follow-up, a comprehensive four-level process was established [4]. The study protocol specified that the total loss to follow-up at the end of the study should be less than 10% of the study population.

For reported serious adverse events a group of medical doctors specializing in epidemiology, drug safety and internal medicine (medical reviewer group) contacted the study participants as well as the diagnosing or treating physicians to clarify and validate the information (including diagnosis, diagnostic procedures, exposure and treatment) received from participants [4]. Events that were substantiated by a diagnostic measure with high specificity (such as cerebral magnetic resonance imaging for cerebrovascular accidents) or by a clinical diagnosis supported by a diagnostic test with low specificity (such as D-dimer for VTE) were categorized as 'confirmed'. Events were categorized as 'not confirmed' if diagnostic measures did not substantiate the diagnosis reported by the participant, if a different medical condition was diagnosed by the attending physician, or if the participant did not contact a health professional to clarify her symptoms and no diagnostic measures were performed [4].

For the analysis, classification of all VTE and ATE was verified by independent blinded adjudication. All decisions made by the medical reviewer group were reassessed by three independent medical experts specializing in radiology and nuclear medicine, cardiology, as well as internal medicine and vascular diseases. Brand names, doses, regimens and compositions of the hormonal contraceptives used by the study participants were rendered anonymous for this process. The adjudicators performed the reviews independently of each other and without knowing the judgment of the other adjudicators or the medical reviewer group [4]. Events were classified as confirmed if that was the judgment of at least one adjudicator.

Evaluation

The analyses of the extended study were carried out in accordance with the statistical analysis plan, which was agreed upon with the European regulatory authorities and approved by the Safety Monitoring and Advisory Council prior to the first inferential analysis of the final database. The analysis plan stipulated that the primary analysis be based on the European arm of the study alone. This was due to concerns by the European regulatory authorities about the low

proportion of DNG/EV users in the United States (see results section). Cox regression analysis of the combined US and European data was performed for exploratory reasons only.

Inferential statistics were based on Cox proportional hazard models. Adjustment for potential confounding was based on an a priori defined expert model (primary model). For venous thromboembolism in the European dataset this model included age, body mass index (BMI), duration of current hormonal contraceptive use, and family history of VTE; for ATE it included age, BMI, smoking, treated hypertension, and a family history of fatal ATE. The prognostic factors were included as time-varying covariates in the statistical model. For the combined analysis of US and European datasets, geographic region was used as an additional covariate. Furthermore, a 'backward stepwise procedure' was chosen to generate an automated statistical model (secondary model). This procedure started with all available prognostic factors (e.g., age, BMI, duration of current hormonal contraceptive use, family history of thromboembolic events, starter, restarter and switcher status, estrogen dose of the OC preparation, concomitant medication, chronic disease, smoking, geographic region, educational level) included in the statistical model as covariates (like in a saturated model). All prognostic factors that had no relevant impact on the risk estimates were removed from the model in a stepwise procedure. The results of the primary and secondary models were nearly identical. Therefore, results reporting focuses on the primary model.

Three exposure groups were compared: users of new COCs containing DNG/EV, users of other COCs, and users of levonorgestrel-containing COCs. All VTE and ATE were always adjudicated for the hormonal contraceptive used by the respective participant at the time of the event. The exposure times of a comparison group are based on the sum of all periods for which the preparations of this group were used. If hormonal contraceptive use had been stopped during the three-month period prior to the VTE diagnosis, the event was adjudicated to the last hormonal contraceptive used before the event; if the last exposure was more than three months prior, the event was adjudicated to none of the exposure groups.

The analyses for VTE used two different datasets: all VTE and what are known as 'idiopathic' VTE. The latter dataset excludes cases with acute risk factors (such as pregnancy, delivery, trauma, immobilization, long-haul travel, surgery, and chemotherapy).

After reviewing the study results, the FDA expressed concerns that the regression method used might be biased "in the presence of time-dependent exposure and time-dependent confounders that are affected by exposure". Therefore, additional analyses of VTE and ATE were done in which: 1) only the baseline values of the covariates age, BMI, and family history of VTE were used for adjustment; and 2) analysis was restricted to the first treatment period after study entry (i.e. data that accrued after study participants switched preparations within the study were not included).

The market introduction of DNG/EV in the United States was less successful than expected by the manufacturer of the product. As a result, the proportion of US study participants who used DNG/EV was very small. Therefore, the European regulatory authorities requested that the primary analysis be based on the European arm of the study alone. To compensate for the loss of statistical power, the follow-up of European study participants was extended. This extension of the follow-up also has the advantage that the results are more representative for the cardiovascular risks associated with long-term use of COCs. The extension of the study was powered to detect a 2.0-fold and 0.5-fold

VTE risk for DNG/EV compared to oCOC within the European study arm. The power calculations were based on the long-term VTE incidence rates observed in the EURAS and INAS-OC studies [4,5].

Results

A total of 53,750 women were enrolled by 1,327 health-care professionals. Overall, 3,547 women (6.6%) were excluded because they: 1) were enrolled two or more times by one or more study centers (1.5%); 2) continued to use their previous hormonal contraceptive (1.5%); 3) did not start OC use after study entry (2.5%); or 4) declined to sign the informed consent form (1.1%). The remaining 50,203 quality-controlled computerized data sets from the women (one per woman) with baseline information were analyzed. A total of 30,098 (60.0%) and 20,105 women (40.0%) were recruited in Europe and the United States, respectively. In the combined European and US cohorts these 50,203 study participants were followed up for 145,224 woman-years (WY) of observation (mean value, 2.9 years per study participant): 112,638 WY (77.6%) in Europe and 32,586 WY (22.4%) in the United States). The later start of recruitment and the earlier stop of follow-up in the United States resulted in a shorter average follow-up in the United States compared to Europe: 1.6 years and 3.7 years, respectively.

At study entry, 10,191 women received a prescription for DNG/EV and 40,012 for oCOC. The latter included 5,796 users of LNG/EE (Table 1). At the end of the INAS-SCORE study, DNG/EV, oCOC, and LNG/EE had been used for 15,850 WY, 78,446 WY, and 13,078 WY, respectively. For 7,033 WY and 43,895 WY, study participants had switched to other hormonal contraceptives (e.g., patches, injections, vaginal rings) or had not used any hormonal contraceptive, respectively.

At the end of the study, 1,645 of the 50,203 women, or 3.3% (3.7% for DNG/EV, 3.2% for oCOC, and 3.3% for LNG/EE), were lost during the follow-up period (September 2009-December 2016). Overall, all loss to follow-up rates were low and balanced across exposure groups. The goal of a loss to follow-up rate of less than 10% was achieved for the total study population, for each of the exposure groups, and for the European and US populations.

For each of the main user groups (DNG/EV and oCOC) plus the LNG/EE subgroup, Table 2 shows the number of women with baseline information (N), the exposure, the corresponding proportion of exposure for each of these populations, and descriptive statistics for age, weight, and BMI. At study entry, 20.3% of women were prescribed DNG/EV, and 79.7% oCOC (LNG/EE, 11.5%). Mean age in the DNG/EV exposure group was 5.7 years higher than that in the COC and LNG/EE (sub)-groups; the 75th age percentile of DNG/EV exceeds the oCOC figure by about 10 years. The reasons for the age differences between the cohorts were discussed in our previous report [3].

Mean weight and mean BMI were similar for all COC (sub)-groups. This is true for both the European and the US study populations. However, weight and BMI were substantially higher in the United States than in Europe [3], and only few DNG/EV users were recruited in the United States. Accordingly, the combined transatlantic DNG/EV values were lower than the corresponding values of the other COC groups because the latter were influenced much more by the high US values.

Overall, 16,233 women (32.3% of the study population) were starters (first-time users) at study entry, 10,175 women (20.3%) were switchers, and 23,795 (47.4%) were restarters. No substantial differences were observed between the three (sub)-groups of DNG/EV, oCOC, and LNG/EE. The proportion of starters was slightly lower

Table 1. Number of women enrolled, excluded, and analysed.

Women	Nos.	(%) ^a	[%] ^b
A) who agreed to participate	53,750	---	[100.0]
B) excluded because of protocol violations ^c	3,547	---	[6.6]
C) analysed	50,203	(100.0)	[93.4]
Cohorts			
DNG/EV	10,191	(20.3)	[19.0]
oCOC	40,012	(79.7)	[74.4]
<i>of which</i> LNG/EE	5,796	(11.5)	[10.8]
Regions			
United States	20,105	(40.0)	[37.4]
Europe	30,098	(60.0)	[56.0]
European countries			
Austria	2,208	(4.4)	[4.1]
France	252	(0.5)	[0.5]
Germany	8,613	(17.2)	[16.0]
Italy	8,508	(16.9)	[15.8]
Poland	9,131	(18.2)	[17.0]
Sweden	1,111	(2.2)	[2.1]
UK	275	(0.5)	[0.5]
<i>Primary analysis [based on European data]</i>			
DNG/EV	9,791	(19.5)	[18.2]
oCOC	20,307	(40.4)	[37.8]
<i>of which</i> LNG/EE	3,736	(7.4)	[7.0]

^a: Percentage of women who agreed to participate; ^b: Percentage of women who were in the final analysis; ^c: Women who: 1) were enrolled two or more times by one or more study centers; or 2) continued their previous hormonal contraceptive; 3) never started OC use after study entry; or 4) declined to sign the informed consent form.

Table 2. User cohorts (USA and Europe combined): Number of women, exposure, and descriptive statistics on age, weight, and BMI at study entry. *: Exposure includes 50,928 WY for women who stopped hormonal contraceptive use after study entry or switched to non-COC hormonal contraceptives.

	DNG/EV	oCOC		Total
		All	LNG/EE	
N [%]	N = 10,191 [20.3]	N = 40,012 [79.7]	N = 5,796 [11.5]	N = 50,203 [100.0]
WY [%]	15,850 [10.9]	78,446 [54.0]	13,078 [9.5]	145,224 [100.0]*
Age, mean [SD]	31.7 [10.0]	26.0 [7.9]	26.0 [8.4]	27.1 [8.7]
Age, minimum	11	11	12	11
Age, percentile 5	17	16	16	16
Age, percentile 25	23	20	19	20
Age, median	31	24	24	25
Age, percentile 75	40	30	31	32
Age, percentile 95	48	42	43	44
Age, maximum	59	58	55	59
Weight, mean [SD]	62.7 [12.2]	66.6 [16.4]	66.1 [15.2]	65.8 [15.7]
Weight, minimum	30	30	37	30
Weight, percentile 5	48	48	48	48
Weight, percentile 25	55	55	55	55
Weight, median	60	63	63	62
Weight, percentile 75	69	73	73	72
Weight, percentile 95	85	99	96	96
Weight, maximum	172	191	173	191
BMI, mean [SD]	23.0 [4.2]	24.5 [5.9]	24.2 [5.4]	24.2 [5.6]
BMI, minimum	12.7	10.6	14.2	10.6
BMI, percentile 5	18.0	18.1	18.2	18.1
BMI, percentile 25	20.2	20.4	20.5	20.4
BMI, median	22.2	23.0	23.0	22.8
BMI, percentile 75	24.8	26.9	26.6	26.4
BMI, percentile 95	31.1	36.2	35.1	35.3
BMI, maximum	60.9	71.9	61.6	71.9

for oCOC (32%) and LNG/EE (30%) compared to the new DNG/EV (34.0%). No notable differences between the COC (sub)-cohorts were seen regarding gynecological history, including age at menarche, number of live births, and age at first delivery.

The distribution of prognostic factors for cardiovascular outcomes of interest was described in our previous report [3]. After adjusting for the different age profiles of the user populations, no substantial differences were seen between the exposure groups. Overall, the oCOC and LNG/EE exposure groups showed typical characteristics of US and European COC user populations regarding age structure, socioeconomic and life-style factors, and cardiovascular risk factors [4-9]. The most important difference between the exposure groups was the substantially higher age of DNG/EV users compared to oCOC and LNG/EE. Given this difference, the risk of serious cardiovascular events was a priori higher for DNG/EV users compared to those of other COCs.

Venous thromboembolic events

Table 3 shows the number of VTE, point estimates, and 95% confidence intervals (95% CI) for the exposure groups. A total of 98 VTE were observed, with a lower incidence rate in the DNG/EV group compared to the other COC groups: DNG/EV 11 cases for 6.9 VTE per 10,000 WY, oCOC 69 cases for 8.8 VTE per 10,000 WY, and LNG/EE 13 VTE for 9.9 VTE per 10,000 WY. The incidence rate in the ‘no use’ cohort (15 cases for 3.4 VTE per 10,000 WY) was substantially lower compared to the COC (sub)-cohorts. The results for those women who switched after recruitment to other hormonal contraceptives (OHC: i.e. injections, implants, levonorgestrel-releasing IUDs, or contraceptive patches) are too sparse (3 VTE) for any meaningful analysis. For 31 of the 98 VTE cases (32%), a pulmonary embolism was observed (DNG/EV cohort: 5 cases; oCOC cohort: 21 cases; LNG/EE sub-cohort: 5 cases; OHC cohort: no case; ‘no use’ cohort: 5 cases).

The VTE risk for COC users was approximately 2.5 times higher than that for non-users. Exclusion of 33 US and European VTE cases associated with acute risk factors for VTE resulted in an overall incidence in COC users of 6.0 per 10,000 WY for so-called ‘idiopathic’ VTE. The VTE risk for COC users without acute risk factors was 3.8 times higher than that for non-pregnant non-users without acute risk factors.

The results of the Cox regression analyses are shown in Table 4. The crude hazard ratio (HR_{crude}) for DNG/EV versus oCOC was 0.8 (95% CI, 0.4-1.5). The corresponding adjusted hazard ratio (HR_{adj.}) was 0.5 (95% CI, 0.2-0.8). The effect of the adjustment reflects primarily the differences in the age profile of the two cohorts. Alternative analyses using a backwards stepwise procedure for the selection of prognostic factors yielded almost identical results. A comparison of the DNG/EV and LNG/EE groups showed similar results: the crude and adjusted VTE hazard ratios were 0.7 (95% CI, 0.3-1.6) and 0.4 (95% CI, 0.2-1.1), respectively.

No VTE were observed in US users of DNG/EV. Therefore, no meaningful analysis of the US VTE data is possible. Number of VTE, point estimates, and 95% confidence intervals for the (sub)-cohorts are presented in Table 3.

A total of 68 VTE were observed in the European study population (primary analysis) with a lower incidence rate in the DNG/EV cohort compared to the other COC (sub)-cohorts (Table 3): DNG/EV 11 cases for 7.1 VTE per 10,000 WY, oCOC 44 cases for 8.1 VTE per 10,000 WY, and LNG/EE 9 VTE for 8.8 VTE per 10,000 WY. The incidence

Table 3. Venous thromboembolic events: Number, incidence, and 95% confidence intervals per exposure group * : Incidence rates are given in events/10,000 WY.

Dataset	DNG/EV		oCOC		LNG/EE		OHC		No use		Total n
	n	Incidence* & 95% CI	n	Incidence* & 95% CI	n	Incidence* & 95% CI	n	Incidence* & 95% CI	n	Incidence* & 95% CI	
USA and Europe											
All VTE	11	6.9 (3.5-12.4)	69	8.8 (6.8-11.1)	13	9.9 (5.3-17.0)	3	4.3 (0.9-12.5)	15	3.4 (1.9-5.6)	98
'Idiopathic' VTE	7	4.4 (1.8-9.1)	50	6.4 (4.7-8.4)	11	8.4 (4.2-15.0)	1	1.4 (0.0-7.9)	7	1.6 (0.6-3.3)	65
USA											
All VTE	0	0.0 (0.0-81.9)	25	10.4 (6.8-15.4)	4	14.3 (3.9-36.7)	0	0.0 (0.0-16.6)	5	7.7 (2.5-18.0)	30
'Idiopathic' VTE	0	0.0 (0.0-81.9)	16	6.7 (3.8-10.8)	3	10.8 (2.2-31.4)	0	0.0 (0.0-16.6)	2	3.1 (0.4-11.2)	18
Europe											
All VTE (primary analysis)	11	7.1 (3.5-12.7)	44	8.1 (5.9-10.9)	9	8.8 (4.0-16.6)	3	5.7 (1.2-16.8)	10	2.7 (1.3-4.9)	68
'Idiopathic' VTE	7	4.5 (1.8-9.3)	34	6.2 (4.3-8.7)	8	7.8 (3.4-15.3)	1	1.9 (0.0-10.7)	5	1.3 (0.4-3.1)	47

Table 4. Cox regression analyses of VTE: Crude and adjusted hazard ratios and corresponding 95% confidence intervals. ^a adjusted for age, BMI, and family history of VTE at the time of VTE diagnosis, as well as geographic region and current duration of use; ^b adjusted for age, BMI, and family history of VTE at the time of VTE diagnosis, as well as current duration of use; ^c adjusted for age, BMI, and family history of VTE at baseline, as well as current duration of use; ^d : p=0.06; ^e : p=0.05.

Dataset	Comparator	HR (DNG/EV vs. comparators)			
		Crude Estimate	95% CI	Adjusted Estimate	95% CI
USA and Europe	oCOC	0.8	0.4-1.5	0.5 ^a	0.2-0.8
	LNG/EE	0.7	0.3-1.6	0.4 ^a	0.2-1.1
Europe					
Primary analysis	oCOC	0.9	0.4-1.7	0.4 ^b	0.2-0.9
	LNG/EE	0.8	0.3-2.0	0.4 ^b	0.2-1.1
'Idiopathic' VTE only	oCOC	0.7	0.3-1.6	0.4 ^b	0.2-0.9
	LNG/EE	0.6	0.2-1.7	0.4 ^b	0.1-0.9
Time-independent covariates	oCOC	0.9	0.4-1.7	0.4 ^c	0.2-0.8
	LNG/EE	0.8	0.3-2.0	0.4 ^c	0.2-1.0 ^d
Time-independent covariates; first treatment period only	oCOC	0.9	0.4-1.7	0.4 ^e	0.2-0.8
	LNG/EE	0.8	0.3-2.0	0.4 ^e	0.1-1.0 ^e

rate in the 'no use' group (10 cases for 2.7 VTE per 10,000 WY) was substantially lower compared to the COC exposure groups.

In Europe the HR_{crude} for DNG/EV versus oCOC was 0.9 with a 95% confidence interval of 0.4 to 1.7. The HR_{adj.} was 0.4 (Table 4) with an upper 95% confidence limit of 0.9. The effect of the adjustment reflects primarily the differences in the age profile of the two cohorts. Alternative analyses using a backwards stepwise procedure for the selection of prognostic factors yielded almost identical results. A comparison of the DNG/EV cohort with the LNG/EE sub-cohort showed similar point estimates with wider confidence intervals: the crude and adjusted VTE hazard ratios were 0.8 (95% CI, 0.3-2.0) and 0.4 (95% CI, 0.2-1.1), respectively. Restriction of the analysis to 'idiopathic' VTE only yielded adjusted HRs of 0.4 (95% CI, 0.2-0.9) and 0.4 (95% CI, 0.1-0.9) for DNG/EV vs. oCOC and LNG/EE, respectively. Furthermore, the results of the alternative calculations requested by the FDA (see section 2.4) varied only marginally compared to the primary analysis (Table 4). This is also true for sensitivity analyses using a saturated Cox regression model that included 14 potential prognostic factors and a Cox regression model that resulted from a backwards stepwise procedure. Moreover, the difference between DNG/EV and oCOC remained statistically significant after excluding preparations containing so-called third generation (gestodene, desogestrel) and fourth generation (drospirenone) progestins.

Arterial thromboembolic events

The number of ATE more than doubled during the extension phase of the study. A total of 39 ATE were observed in the study (Table 5): 11 AMIs, 18 ischemic strokes, 7 TIAs, and 3 thromboses of a peripheral artery. The ATEs break down among the (sub)-cohorts as follows:

DNG/EV 1 case, oCOC 29 cases, LNG/EE 3 cases, OHC 0 cases, and 'no use' 9 cases. This corresponds to ATE incidence rates of 0.6 ATE/10,000 WY for the DNG/EV cohort, and of 3.7, 2.3, 0.0, and 2.1 for the oCOC, LNG/EE, OHC, and 'no use' (sub)-cohorts, respectively. The incidence rates for oCOC and LNG/EE were slightly higher than the incidence assumptions used for the power calculations. In contrast the incidence rates for DNG/EV were lower than expected.

The 11 AMIs break down among the (sub)-groups as follows: DNG/EV no case, oCOC 8 cases, LNG/EE 2 cases, OHC no case, and 'no use' 3 cases. This corresponds to AMI incidence rates of 0.0 ATE/10,000 WY for the DNG/EV cohort, and of 1.0, 1.5, 0.0, and 0.7 for the oCOC, LNG/EE, OHC, and 'no use' cohorts, respectively. Overall, 18 cases of ischemic strokes and 7 cases of TIA occurred: DNG/EV 1 and 0 case, oCOC 14 and 4 case(s), LNG/EE 1 and 0 case, OHC 0 and 0 case, and 'no use' 3 and 3 cases. This corresponds to the following incidence rates: DNG/EV, 0.6 and 0.0 events/10,000 WY; oCOC, 1.8 and 0.5 events/10,000 WY; LNG/EE, 0.8 and 0.0 events/10,000 WY; OHC, 0.0 and 0.0 events/10,000 WY; 'no use', 0.7 and 0.7 events/10,000 WY.

The results for Europe alone are similar to the overall ATE results (Table 5). The US data were too sparse for any meaningful comparison between the (sub)-groups (Table 5).

The statistical analysis plan stipulated that hazard ratios were only to be calculated if a minimum of 5 confirmed events were available in each of the comparison groups. This requirement was not fulfilled for two of the three COC (sub)-groups (DNG/EV and LNG/EE). However, during the extension phase of the study the FDA requested calculation of ATE hazard ratios using alternative Cox regression models. The results of these calculations varied only marginally (Figure 1). All

Table 5. Arterial thromboembolic events: Number, incidence, and 95% confidence intervals per exposure group. AMI: Acute Myocardial Infarction; TIA: Transient Ischemic Attack; *: incidence rates are given in events/10,000 WY.

Category	DNG/EV		oCOC		LNG/EE		OHC		No use		Total n
	n	Incidence* & 95% CI	n	Incidence* & 95% CI	n	Incidence* & 95% CI	n	Incidence* & 95% CI	n	Incidence* & 95% CI	
Complete dataset All ATE <i>of which</i>	1	0.6 (0.0-3.5)	29	3.7 (2.5-5.3)	3	2.3 (0.5-6.7)	0	0.0 (0.0-4.3)	9	2.1 (0.9-3.9)	39
AMI	0	0.0 (0.0-1.9)	8	1.0 (0.4-2.0)	2	1.5 (0.2-5.5)	0	0.0 (0.0-4.3)	3	0.7 (0.1-2.0)	11
Ischemic stroke	1	0.6 (0.0-3.5)	14	1.8 (1.0-3.0)	1	0.8 (0.0-4.3)	0	0.0 (0.0-4.3)	3	0.7 (0.1-2.0)	18
TIA	0	0.0 (0.0-1.9)	4	0.5 (0.1-1.3)	0	0.0 (0.0-2.3)	0	0.0 (0.0-4.3)	3	0.7 (0.1-2.0)	7
Peripheral ATE	0	0.0 (0.0-1.9)	3	0.4 (0.1-1.1)	0	0.0 (0.0-2.3)	0	0.0 (0.0-4.3)	0	0.0 (0.0-0.7)	3
European data All ATE <i>of which</i>	1	0.6 (0.0-3.6)	23	4.2 (2.7-6.3)	3	2.9 (0.6-8.5)	0	0.0 (0.0-5.7)	8	2.1 (0.9-4.2)	32
AMI	0	0.0 (0.0-1.9)	5	0.9 (0.3-2.1)	2	1.9 (0.2-7.0)	0	0.0 (0.0-5.7)	3	0.8 (0.2-2.3)	8
Ischemic stroke	1	0.6 (0.0-3.6)	12	2.2 (1.1-3.9)	1	1.0 (0.0-5.4)	0	0.0 (0.0-5.7)	2	0.5 (0.1-1.9)	15
TIA	0	0.0 (0.0-1.9)	3	0.6 (0.1-1.6)	0	0.0 (0.0-2.9)	0	0.0 (0.0-5.7)	3	0.8 (0.2-2.3)	6
Peripheral ATE	0	0.0 (0.0-1.9)	3	0.6 (0.1-1.6)	0	0.0 (0.0-2.9)	0	0.0 (0.0-5.7)	0	0.0 (0.0-0.8)	3
US data All ATE <i>of which</i>	0	0.0 (0.0-82.0)	6	2.5 (0.9-5.5)	0	0.0 (0.0-10.7)	0	0.0 (0.0-16.6)	1	1.5 (0.0-8.6)	7
AMI	0	0.0 (0.0-82.0)	3	1.3 (0.3-3.7)	0	0.0 (0.0-10.7)	0	0.0 (0.0-16.6)	0	0.0 (0.0-4.6)	3
Ischemic stroke	0	0.0 (0.0-82.0)	2	0.8 (0.1-3.0)	0	0.0 (0.0-10.7)	0	0.0 (0.0-16.6)	1	1.5 (0.0-8.6)	3
TIA	0	0.0 (0.0-82.0)	1	0.4 (0.1-3.0)	0	0.0 (0.0-10.7)	0	0.0 (0.0-16.6)	0	0.0 (0.0-4.6)	1
Peripheral ATE	0	0.0 (0.0-82.0)	0	0.0 (0.0-1.3)	0	0.0 (0.0-10.7)	0	0.0 (0.0-16.6)	0	0.0 (0.0-4.6)	0

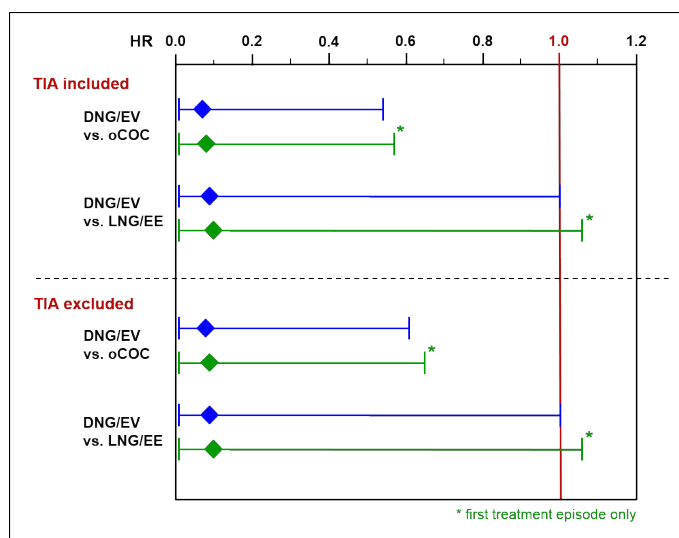


Figure 1. Cox regression analyses for ATE - adjusted hazard ratios and 95% confidence intervals for DNG/EV vs. oCOC and LNG/EE

comparisons for DNG/EV versus oCOC (ATE with TIAs included and ATE with TIAs excluded) resulted in very low adjusted HRs (approximately 0.1) and were statistically significant (p-values between 0.01 and 0.02). The difference between DNG/EV and oCOC remained statistically significant if the exclusion of preparations was restricted to gestodene, desogestrel, and drospirenone. The comparisons for DNG/EV versus LNG/EE showed similarly low HRs (approximately 0.1), but unity was included in the 95% confidence intervals.

Discussion and conclusion

The incidence rates for VTE and ATE were substantially lower for DNG/EV compared to oCOC and LNG/EE. The statistical analysis of the European dataset yielded adjusted hazard ratios of 0.4 for VTE and 0.1 for ATE for the comparison of DNG/EV versus oCOC. The corresponding 95% confidence intervals did not include unity and suggested superiority of DNG/EV. Comparisons versus LNG/EE yielded point estimates for the adjusted hazard ratios of 0.4 and 0.1

for VTE and ATE (i.e., similar to the comparison of DNG/EV versus oCOC) but the 95% confidence intervals included unity (except for 'idiopathic' VTE). The fact that the point estimates for the HRs were similar for DNG/EV versus oCOC and DNG/EV versus LNG/EE suggests that potential differences between the cohorts were not 'diluted' by the inclusion of COCs that are potentially associated with an increased risk of VTE. In addition, sensitivity analyses that included all other available prognostic factors for VTE, as well as the additional analyses requested by the US and European regulatory authorities, confirmed the results of the primary analyses.

INAS-SCORE was a non-experimental epidemiological study. In such studies the possibility of bias and residual confounding can never be entirely eliminated, and the ability to infer causation is correspondingly limited [10]. Valid information on potential sources of confounding and sophisticated statistical and epidemiologic methodology help to reduce the impact of bias and residual confounding [11]. Valid information on many relevant prognostic factors were available for INAS-SCORE and the statistical analyses incorporated many suggestions by the scientific advisory council as well as the European and US regulatory authorities. However, differentiating between causation, bias and residual confounding remains challenging when all that exists is a weak association [12,13]. Relative risk estimates that are close to unity may not allow this [14,15]. In general, strict causal interpretations of relative risks of two or less are difficult in observational research [16,17].

With regard to VTE, the HRs for DNG/EV vs. oCOC and for DNG/EV vs. LNG/EE were slightly lower than 0.5 (or higher than two for oCOC or LNG/EE vs. DNG/EV). Therefore, the results of INAS-SCORE should be considered valid if the assumption holds that reliable information on the most relevant confounders was available in this study, and that the influence of bias was limited to the extent possible for a large, non-experimental study. In fact, the INAS-SCORE study combines several methodological strengths that support the validity of its results: 1) a prospective, comparative cohort design; 2) availability of important confounder information (e.g., BMI and family history of cardiovascular outcomes); 3) validation of outcomes of interest and exposure for the relevant cases; 4) comprehensive long-term follow-up and very low loss to follow-up to minimize underreporting; 5)

independent, blinded adjudication of critical outcomes; 6) relevant statistical analyses (e.g., stratified analyses by geographic region, user status, and exposure period; comparison of isochronous, new user cohorts; sensitivity analyses on the impact of i) the adjudication process, ii) outcome definition, iii) selection of prognostic factors, iv) choice of statistical model, and v) choice of comparator cohort); 7) study population with baseline characteristics similar to OC users under routine clinical conditions; 8) reproducibility of the typical time pattern of VTE risk; and 9) supervision by an independent Safety Monitoring and Advisory Council as well as scientific independence from the study funder.

Selection and misclassification bias were probably not major issues in the INAS-SCORE study because VTE of both in- and out-patients were included in the analyses, the demographic characteristics of the participants are representative for adult COC users [6-8], and precise information on the exposure and the outcomes of interest were available. In addition, reliable information on duration of current use was available. Furthermore, the low loss to follow-up rate of 3.3% is noteworthy. In theory, a disproportionately high percentage of VTE and ATE could have occurred in those patients who were lost to follow-up, because these events could be the reason for the break in contact with the investigators. An advantage of the INAS-SCORE study design, however, is that the investigator team 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).

However, the investigator acknowledges that it was impossible to exclude diagnostic bias. This is true for all non-experimental studies on VTE. Clinical symptoms of VTE cover the spectrum from a complete absence or unspecific, slight symptoms to dramatic, acute, life-threatening symptoms [18-20]. A high awareness of potential cardiovascular risks of combined oral contraceptive use might have led to more diagnostic procedures and therefore to more detected VTE. It is conceivable that this potential bias led to an overestimate of the relative risk for new COCs like DNG/EV. However, of relevance here is that an overestimate of the VTE risk associated with the use of DNG/EV is more likely than an underestimate. Another consideration is the fact that due to the non-interventional character of the study, information on specific gene mutations was only available for VTE cases but not for the vast majority of study participants. This limitation is mitigated by information on family history of VTE which has an even higher predictive value for VTE compared to gene mutations [21].

An additional potential issue discussed with regulatory authorities concerned the validation and blinded adjudication process. Many epidemiological studies on the safety of hormonal contraceptives suffer from the limited validity of information on VTE. For example, Severinsen's review of medical records [22] revealed that VTE diagnoses in the Danish registries were incorrect in 25% of cases diagnosed in hospital wards, and in 69% of cases diagnosed in emergency departments. It is therefore of utmost importance to validate the diagnoses of VTE and other cardiovascular outcomes. In the INAS-SCORE study, the validity of VTE and ATE diagnoses was ensured by a comprehensive procedure that included three blinded, independent external physicians specialized in the relevant medical areas. This blinded adjudication procedure was established to minimize classification bias in a non-experimental setting. The validation process established in INAS-SCORE can be considered a strength compared to many other epidemiological studies.

Overall, no major methodological shortcomings were identified that could have influenced the VTE and ATE hazard ratios substantially. A moderate influence of bias and residual confounding cannot be ruled out, but the most important limitation of the study – the difficulty of ruling out diagnostic bias – would tend to overestimate the VTE and ATE risk of new COCs. The adjusted hazard ratios for the risk of VTE and ATE are sufficiently different from unity (0.4 and 0.1) to suggest superiority of DNG/EV compared to the conglomerate of other COCs. This is probably also true for the comparison of DNG/EV versus LNG/EE. In any case, the authors consider DNG/EV to be at least as safe as levonorgestrel-containing COCs regarding the risk of VTE and ATE. There are data indicating that DNG is not safer than established progestins [23]. Therefore, the favorable safety profile of DNG/EV can be attributed to estradiol valerate.

Authorship

The study was planned, supervised and reported by J. Dinger. S. Möhner and K. Heinemann were responsible for data management and study management during the study, respectively.

Acknowledgements

The authors' special thanks are due to Anja Bauerfeind and Thai Do Minh for the statistical analyses, and to Marlene Schoofs for editorial support in preparing the manuscript. The authors would like to express their appreciation to the members of the independent Safety Monitoring and Advisory Council for their constructive criticism and unfailingly fair scientific discussion. The authors would also like to highlight the contributions of numerous colleagues who were responsible for the field work in the individual countries. They clarified data inconsistencies and missing data, validated patient-reported adverse events with patience, care and tenacity, and their untiring commitment enabled a remarkably low loss to follow-up rate.

Funding information

The study was supported by an unconditional grant from Bayer AG, Germany.

Competing Interests

The study was requested by the European health authorities. The costs of the study were covered by the manufacturer of DNG/EV, Bayer AG, Germany. The funder had no access to the source data and did not participate in designing the study or analysing the data. The study was supervised by an independent Safety Monitoring and Advisory Council with full authority over the study (including study protocol, protocol amendments, data analysis, and study cessation). The authors have no other financial or non-financial competing interests.

References

1. Dinger J, Shapiro S (2012) Combined oral contraceptives, venous thromboembolism, and the problem of interpreting large but incomplete datasets. *J Fam Plann Reprod Health Care* 38: 2-6. [[Crossref](#)]
2. Wiegratz I, Lee JH, Kutschera E, Winkler UH, Kuhl H (2004) Effect of four oral contraceptives on hemostatic parameters. *Contraception* 70: 97-106. [[Crossref](#)]
3. Dinger J, Do Minh T, Heinemann K (2016) Impact of Estrogen Type on Cardiovascular Safety of Combined Oral Contraceptives. *Contraception* 94: 328-339. [[Crossref](#)]
4. Dinger JC, Heinemann LAJ, Kühl-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 women-years of observation. *Contraception* 75: 344-354. [[Crossref](#)]

5. 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](#)]
6. Potthoff P, Heinemann LAJ, Güther B (2004) A household panel as a tool for cost-effective health-related population surveys: validity of the Healthcare Access Panel. *Ger Med Sci* 2: Doc05. [[Crossref](#)]
7. Güther B (2005) Access panel – quality aspects and experience, Surveys in public health – developments and perspectives in research and political counseling. Sankt Augustin: Asgard Verlag, pp: 91-102.
8. Potthoff P, Güther B (2005) Access panel in the context of health research. *Forum Public Health* 13: 16.
9. Dinger J, Möhner S, Heinemann K (2016) Cardiovascular risks associated with the use of drospirenone-containing combined oral contraceptives. *Contraception* 93: 378-385. [[Crossref](#)]
10. 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](#)]
11. Rothman KJ, Poole C (1988) A strengthening programme for weak associations. *Int J Epidemiol* 17: 955-959. [[Crossref](#)]
12. 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](#)]
13. Shapiro S (2000) Bias in the evaluation of low-magnitude associations: an empirical perspective. *Am J Epidemiol* 2000;151:939-945. [[Crossref](#)]
14. 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](#)]
15. 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](#)]
16. Taubes G (1995) Epidemiology faces its limits. *Science* 269: 164-169. [[Crossref](#)]
17. Hill AB (1965) The environment and disease: association or causation? *Proc R Soc Med* 58: 295-300. [[Crossref](#)]
18. 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](#)]
19. Belcaro C, Geroulakos G, Nicolaides AN, Myers KA, Winford M (2001) Venous thromboembolism from air travel: the LONFLIT Study. *Angiology* 52: 369-374. [[Crossref](#)]
20. Schwarz T, Siegert G, Oettler W, Halbritter K, Beyer J, et al. (2003) Venous thrombosis following long-haul flights. *Arch Intern Med* 2003;163: 2759-2764. [[Crossref](#)]
21. 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](#)]
22. Severinsen MT, Kristensen SR, Overvad K, Dethlefsen C, Tjønneland A, et al. (2010) Venous thromboembolism discharge diagnoses in the Danish National Patient Registry should be used with caution. *J Clin Epidemiol* 63: 223-238. [[Crossref](#)]
23. Dinger J (2020) 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. *Front Womens Health* 5: 175.