CHAPTER 6

Discriminant analysis: a method to integrally investigate differences in effects on haemostasis variables by combined hormonal contraceptives

Not published to this date

J. Burggraaf[1], M Rad[1], ML de Kam[1], SO Skouby[2], J Jespersen[3], U Winkler[4], AF Cohen[1], C Kluft[1,5]

1 Centre for Human Drug Research, Leiden
2 Unit of Reproductive Medicine, University of Copenhagen, Denmark
3 Unit for Thrombosis Research, Institute of Public Health, University of Southern Denmark
4 Klinikum Wetzlar, Wetzlar, Germany,
5 Good Biomarker Sciences, The Netherlands

ABSTRACT

OBJECTIVES Differences in steroid composition between combined hormonal contraceptives (CHCs) is linked to differences in changes in a number of haemostatic variables as well as the oestrogenicity marker, i.e. sex hormone binding globulin (SHBG). Changes in haemostasis variables are numerous and combinations can drive the translation into a clinical active difference within the balance of the haemostatic system. To sort out which changes and combination of changes in variables might be most relevant we contrasted the two combined oral contraceptives (COCs) containing 30 µg ethinyl estradiol (EE) combined with 150 µg levonorgestrel (LNG), and 30 µg EE combined with150 µg desogestrel (DSG) representing the two extremes of oestrogenicity/androgenicity ratio, also known as second- and third-generation COCs. While prospective work (active surveillance studies) failed to find differential risks, retrospective epidemiological work has consistently shown an increased risk of venous thromboembolism (VTE) with DSG as compared to LNG containing COC with a relative risk of about two.

METHODS We analysed twenty three haemostasis variables and SHBG, and included additionally five other COCs with n= 59-81 per group, excluding factor V Leiden and prothrombin G20210A mutation carriers.

RESULTS AND CONCLUSIONS The discriminant analysis for changes from pre-treatment levels yielded a function (Z) discriminating best among the total of seven COCs and between the two extremes: Z= 2.4 x Δlog SHBG + 3.2 x Δlog FVIIC – 3.2 x Δlog freePS – 2.0 x Δlog PLG – 0.5 x Δlog PCag

This algorithm results in classification of LNG/EE 150/30 and DSG/EE 150/30 with an accuracy between 92 and 97%. Leaving out SHBG, results in closely similar accuracy, between 89 and 96%. This suggests that SHBG, not known for any function in haemostasis, only has a non-causal association. The other COCs show a gradual shift between Z-function of the two at the extremes, challenging the dichotomous classification in second- and third-generation CHCs, suggesting to use the Z-function as rather a continuous variable. Remarkably, APC resistance is not included in this Z-function. This is attributed to its broad range and overlap in data before and during treatment. Also factor VII is included attributed to its very strong and significant rise during CHC use, reaching levels above the normal range. Having performed DA on this set of data does not guarantee its completeness. Other relevant variables may have been left out. The variables in the Z-function draw special attention for their role in the risk of VTE, to be validated in clinical studies.

INTRODUCTION

Mechanism(s) of steroid-induced venous thromboembolism (VTE) are still largely elusive. This may relate to the commonly used approach to focus on a single factor or variables and coagulation tests. This approach can provide plausible, but by definition limited understanding of multi-factorial conditions triggering the development of thrombosis. Therefore other methods or clusters of analysis should be applied to provide more useful information to further our understanding of the effects of combined hormonal contraceptives (CHCs) on the haemostatic balance. One innovative method could be use of discriminant analysis (DA) [1-3]. The value of DA is that, given the right set of variables, it might be able to identify combinations of variables or patterns that could differentiate various levels of risks, related to various extents of hormone-induced changes rather than focusing on a single variable. We previously reported the application of DA to identify combinations of variables that distinguished two CHCs with a low and a high oestrogenicity/androgenicity ratio [4]. That analysis showed that differential effects of CHCs on hepatically synthesized factors cause an altered balance between coagulation and fibrinolysis. More specifically, the change in the haemostatic balance in that data set was noted most prominently for the altered balance by increased factor VII total and plasminogen levels; a combination not captured by any single (coagulation) test. It was also found that SHBG was differentially influenced by the two CHCs. This notion is in keeping with the hypothesis that CHCs differ regarding effect on hepatic protein synthesis, but it is unlikely to provide an explanation for any difference in VTE risk, in terms of elevated levels of SHBG.

In particular due to the fact that a biologically plausible mechanism by which SHBG contributes to increased risk has yet to be identified. It should be mentioned that the link between compositions of COCs and VTE risk is anno 2015 a matter of unresolved dispute about interpretations of the types of clinical studies that have been performed. The large cohort retrospective registration studies show a difference between progestins [5;6], while the prospective, active surveillance studies do not show this difference [7;8].

The approach to analysis of markers in the previous paper was based on three principles. First, we included all available variables to prevent missing any valuable information that might currently not be known for their role in hormone-induced VTE. Second, the analysis was performed for all variables on both the absolute values at end of treatment (EOT) and on the change from baseline (CFB). This approach was chosen to accommodate differences in assay performance and avoid the potentially inflated effects, which can be caused when only percentage changes from baseline are used. Finally, the analysis was concluded with haemostasis variables only as these most likely contain the variables of interest related to VTE risk.

Here we report on an extension of the described DA approach using a data set collected in a large randomised study involving young healthy women taking one of seven oral CHCs for six cycles of 21-days-on and 7-days-off treatment [9;10]. All CHCs contained ethinyl estradiol (EE) in a dose range from 20 to 50 µg, but varied in the progestin content. The progestins were levonorgestrel (LNG), desogestrel (DSG), gestodene (GSD) or norgestimate (NGM). This

data set shows great overlap with the previously used data set although more haemostasis variables were included and multiple assays for some variables (activity and antigen) were used. We consider this large data set well-suited for DA and competent to confirm or refute our previous findings. We performed the DA using the same principles outlined above and focused on three questions.

The first question was whether DA on the large data set on the effects of seven COCs could confirm the differential effects on hepatically synthesized proteins. The second question was what factors would emerge in DA of the two of the seven COCs that differ the most in their oestrogenicity / androgenicity ratio, and whether there would be an overlap with results of the DA done on the data of all seven COCs. Finally, we questioned if our previously found algorithms (See Chapter 5 of this thesis) that identified the differential effects on FVII total and plasminogen as the most discriminative of all variables, could also be used to differentiate accurately the two COCs with the greatest difference in oestrogenicity / androgenacity ratio in this data set.

MATERIALS AND METHODS

Study design

Discriminant analysis was done on the data in which 613 young women were randomised to one of the seven oral contraceptives (Table 1) as reported in the original papers [9;10]. Carriers of factor V Leiden (n=24) and prothrombin mutation (n=14) were excluded from the analyses.

The participants were healthy, non-smoking females who had regular menstrual cycles before the trial and had not used a hormonal method of contraception for at least two menstrual cycles prior to the start of the study. None of the participants had a personal history of coagulation disorder or a family history of a thromboembolic disease among their first degree relatives <55 years of age. Measurements of blood pressure and blood sampling was done at baseline between days 18 and 21 of the menstrual cycle prior to the first cycle of study drug use and follow-up sampling was performed between days 18 and 21 of the third and the sixth (last) treatment cycle.

Table 1 Treatment groups

Treatment*	N
LNG/EE (150/30)	131
LNG/EE (125/50)	67
GSD/EE (75/20)	85
DSG/EE (150/20)	77
GSD/EE (75/30)	86
NGM/EE (250/35)	84
DSG/EE (150/30)	83

*All doses in µg; N: number of users; LNG: levonorgestrel; DSG: desogestrel; GSD: gestodene; NGM: norgestimate; EE: ethinyl estradiol

Blood sampling and measurements

The procedures for blood sampling, sample handling and the description of the performed assays have been described in detail in previous reports [9;10]. The following variables were included in the analysis: fibrinogen (FBG), factor VII antigen (FVIIag) and clotting activity (FVIIc), activated factor VII (FVIIa), factor VIII activity (FVIIIact), von Willebrand factor-ristocetin co-factor activity, antithrombin antigen (ATag) and activity (ATact), protein C antigen (PCag) and activity (PCact), Protein S clotting activity (PSact) and free protein S antigen (freePS), plasminogen (PLG), tissue-type plasminogen activator antigen (t-PAag) and activity (t-PAact), plasmin-antiplasmin complex (PAP), plasminogen activator inhibitor-1 antigen (PAI-1), thrombin-antithrombin complex (TAT), fibrinogen degradation products (FGDP), soluble fibrin, D-dimer, prothrombin fragment 1+2 (F1+2), extrinsic APC resistance (ext APCr), APTT-based APC resistance (APTT-based APCr), SHBG, and C-reactive protein (CRP).

Discriminant analysis

Values were log-transformed prior to analysis. These log-transformed values were subjected to a stepwise DA using either all available data or only the haemostasis variables. Each data set was analysed using the change from baseline (CFB) for each parameter or the absolute end of treatment (EOT) values at the sixth cycle. The thus extracted variables that together were discriminative of the treatments at a level of $p \leq .05$ were entered into the canonical discriminant analysis. The canonical coefficient was used to formulate the discriminating function Z. The algorithms are structured as: $Z = CC_1 \times V_1 + CC_2 \times V_2 + \ldots CC_n \times V_n$ in which CC indicates the canonical coefficient for a certain variable V. For each defined algorithm, mean Z and classification accuracy per treatment was calculated by using the percentages of subjects with a Z greater than the mean Z and Z smaller than the mean Z. Subjects missing measurements for any of the variables the algorithm is based on, were excluded from the population in which classification accuracy was calculated.

Firstly, we performed DA on the data set obtained for all seven COCs. We then selected from this set the two extremes in mean Z-value, and repeated the DA. Finally, we investigated how our previously found discriminating functions for an COC containing EE and LNG and a vaginal ring contraceptive containing EE and Nestorone® would perform in classifying the COCs which form the extremes in mean Z-value of the seven COCs investigated. The analyses were performed using SAS for Windows V9.1.2 (SAS Institute, Inc., Cary, NC, USA).

RESULTS

Discriminant analysis of seven COCs

Change from baseline (CFB) of all variables

CFB values of SHBG, free protein S (freePS), FVIIC, plasminogen (PLG) and protein C antigen (PCag) were found able to classify treatments with the highest possible accuracy at $p \leq .05$ (Table 2). The discriminant function based on parameters whose partial $R^2 \geq .05$ is.

z_{CFBa} = 2.4 x Δlog SHBG + 3.2 x Δlog FVIIC - 3.2 x Δlog freePS - 2.0 x Δlog PLG - 0.5 x Δlog PCag

When FVIIa and FVIIC were omitted and the analysis was repeated, FVIIC was replaced by FVIIag (p< 0.0001) in the list of the extracted variables, indicating the robustness of the FVII-related contribution.

CFB of haemostasis variables

CFB of freePS, FVIIC, PCag, PLG were found able to classify the treatments at p≤.001 (Table 2) with the highest possible accuracy. The discriminant function based on parameters whose partial R^2 ≥.05 is:

z_{CFBh}= 4.2 x Δlog FVIIC - 4.6 x Δlog freePS - 0.6 x Δlog PCag - 1.6 x Δlog PLG

Table 2 Extracted variables and their canonical coefficients resulting from discriminant analysis for the change from baseline of all measured variables and the haemostasis subset of the seven COCs

	Stepwise discriminant analysis			Canonical discriminant analysis
Variables	partial R^2	F-value	p-value	Canonical Coefficient
All variables				
SHBG	0.3407	33.4	<.0001	2.41
Free Protein S	0.2219	18.4	<.0001	-3.19
Factor VIIC	0.1186	8.7	<.0001	3.20
Plasminogen	0.0648	4.4	0.0002	-2.04
Protein C antigen	0.0541	3.7	0.0015	-0.45
Antithrombin activity	0.0448	3.0	0.0072	1.84*
FVIII activity	0.0443	2.9	0.0080	-0.22*
APTT-based APCr	0.0371	2.4	0.0251	-1.71*
Protein S activity	0.0369	2.4	0.0254	0.41*
t-PA activity	0.0326	2.1	0.0459	-0.09*
Haemostasis subset				
Free Protein S	0.3492	39.2	<.0001	-4.62
Factor VIIC	0.1862	16.7	<.0001	4.20
Protein C antigen	0.0609	4.7	0.0001	-0.60
Plasminogen	0.0473	3.6	0.0017	-1.64
Antithrombin activity	0.0403	3.0	0.0063	1.63*
Protein S activity	0.0309	2.3	0.0336	0.0005*
FVIII activity	0.0303	2.3	0.0375	-0.14*

* not included because partial R^2 ≤.05

End of treatment (EOT) values of all variables and the haemostasis subset

Comparable results were obtained when the absolute values measured at the end of the treatment were used (Table 3). Also here, free protein S, FVIIC and plasminogen appeared as the most discriminating haemostasis variables.

Table 3 Extracted variables and their canonical coefficients resulting from the discriminant analysis of the values of all measured variables and of the haemostasis subset at the end of treatment for all seven COCs

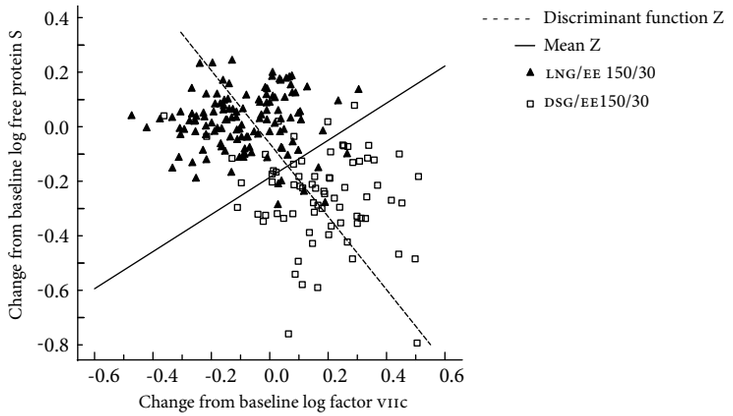
	Stepwise discriminant analysis			Canonical discriminant analysis
Variables	partial R^2	F-value	p-value	Canonical Coefficient
All variables				
SHBG	0.40	44.2	<.0001	3.76
Factor VIIC	0.11	8.5	<.0001	1.85
Free Protein S	0.11	7.9	<.0001	-2.48
Plasminogen activator inhibitor -1	0.07	4.9	<.0001	0.40*
Antithrombin activity	0.06	4.3	0.0003	1.65*
Plasminogen	0.06	4.0	0.0007	-0.78*
Protein S activity	0.05	3.4	0.0030	0.92*
Haemostasis subset				
Free Protein S	0.24	23.5	<.0001	-4.15
Factor VIIC	0.14	11.8	<.0001	2.56
Plasminogen	0.05	3.7	0.0014	-0.77
Plasminogen activator inhibitor -1	0.04	3.4	0.0028	0.38*
Antithrombin activity	0.04	2.7	0.0132	0.42*
Ext APCr	0.03	2.3	0.0314	0.58*
Protein S activity	0.03	2.3	0.0325	0.31*

* not included because partial R^2 ≤.05

Classification accuracy

These analyses identified levonorgestrel 150 µg / ethinyl estradiol 30 µg (LNG/EE 150/30) and desogestrel 150 µg / ethinyl estradiol 30 µg (DSG/EE 150/30) as the two preparations most widely separated in z-values, regardless of use of the EOT or the CFB values (Figure 1). The best classification for these 2 preparations was obtained when the algorithm for the CFB of the haemostasis parameters was used. For LNG/EE 150/30 and DSG/EE 150/30, this algorithm results in mean z-values of -1.74 and 1.20, and classification accuracy of 95 and 91%, respectively.

Figure 1 Graph of change from baseline log free protein S against that of log FVIIc when the discriminant function z (dotted line) is based on discriminant analysis of LNG/EE 150/30 and DSG/EE 150/30



DA of preparations with most widely separated z-value among the seven COCs

CFB of all variables

DA showed that with decreasing significance SHBG, PSact and FVIIC, are most discriminative of LNG/EE 150/30 and DSG/EE 150/30 (Table 4). This resulted in the following discriminant function:

z_{CFBh} = 2.25 x Δlog SHBG + 2.52 x Δlog FVIIC - 4.09 x Δlog PSact

This algorithm results in a mean z of -1.22 and 2.04 for LNG/EE 150/30 and DSG/EE 150/30 and classifies them with 92 and 97% accuracy (Figure 2).

Figure 2 Graph of change from baseline of log free protein S against that of log SHBG when discriminant function z (dotted line) is based on discriminant analysis of the seven COCs. The classification accuracy based on free protein S and SHBG is less than based on z CFBa which also includes information on other variables

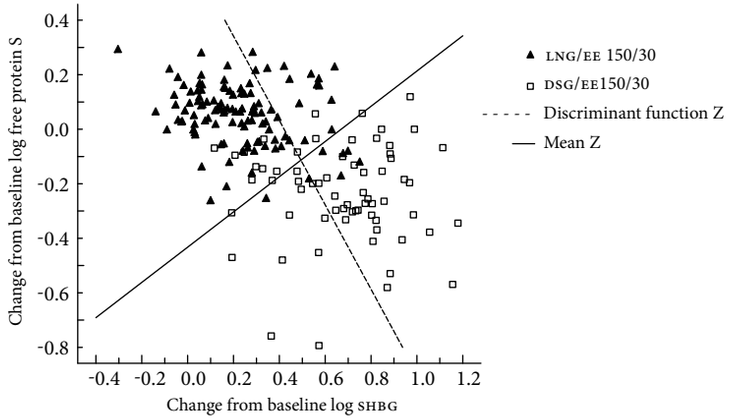


Table 4 Extracted variables and their canonical coefficients resulting from discriminant analysis of change from baseline of all measured variables and of the haemostasis subset of LNG/EE 150/30 and DSG/EE 150/30

	Stepwise discriminant analysis			Canonical discriminant analysis
Variables	Partial R^2	F-value	p-value	Canonical coefficient
All variables				
SHBG	0.4790	126	<.0001	-2.25
Protein S activity	0.2765	52	<.0001	4.09
Factor VIIc	0.1523	24.2	<.0001	-2.52
APTT-based APCr	0.0476	6.7	0.0107	2.97
Fibrinogen	0.0372	5.1	0.0256	-1.60*
Plasmin-antiplasmin complex	0.0299	4.1	0.0448	0.73*
Haemostasis subset				
Protein S activity	0.4905	147.3	<.0001	4.91
Factor VIIc	0.2692	49.5	<.0001	-3.84
APTT-based APCr	0.0406	4.5	0.0360	3.11*
Factor VIIa	0.0403	4.4	0.0373	0.22*
Fibrinogen	0.0340	4.4	0.0386	-1.57*
FVIIIact	0.0312	4.3	0.0404	0.90*
Plasminogen activator inhibitor -1	0.0265	3.9	0.0495	0.18*

* not included because partial $R^2 \leq .05$

EOT values of all variables

Analysis of the EOT values identified SHBG, FVIIC, PSact and F1+2 as a combination discriminative of the treatments according to the following discriminant function (Table 5):

z_{EOTa} = 3.18 x log $SHBG_{EOT}$ + 1.89 x log $FVIIC_{EOT}$ - 2.25 x log $PSact_{EOT}$

The mean z-values associated with this algorithm were -1.07 and 1.78, and the classification accuracy was 87 and 97% for LNG/EE 150/30 and DSG/EE 150/30, respectively.

CFB of haemostasis variables

CFB of PSact (p< .0001), FVIIC (p< .0001), APTT-based APCr (p=0.036), FVIIa (p=0.0373), FBG (p=0.0386), FVIIIact (p=0.0404) and PAI-1 (p=0.0495) were discriminative of the treatments according to the following algorithm (Table 6):

z_{CFBh} = -4.91 x Δlog PSact + 3.84 x Δlog FVIIC - 3.11 x Δlog APTT-based APCr – 0.22 x Δlog FVIIa + 1.57 x Δlog FBG - 0.9 Δlog FVIIIact - 0.18 x Δlog PAI-1

The associated mean z and classification accuracy for LNG/EE 150/30 and DSG/EE 150/30 was 1.00 and -1.68, and 89 and 96%, respectively (Table 6).

Table 5 Extracted variables and their canonical coefficients resulting from discriminant analysis of values of all measured variables and of the haemostasis subset at end of treatment with of LNG/EE 150/30 and DSG/EE 150/30

	Stepwise discriminant analysis			Canonical discriminant analysis
Variables	Partial R^2	F-value	p-value	Canonical coefficient
All variables				
SHBG	0.5341	159.4	<.0001	3.18
Factor VIIC	0.1435	23.1	<.0001	1.89
Protein S activity	0.1106	17	<.0001	-2.25
Fragment 1+2	0.0389	5.5	0.0203	0.39*
Haemostasis				
Protein S activity	0.3448	81.1	<.0001	3.70
Factor VIIC	0.2016	38.6	<.0001	-2.83
Antithrombin antigen	0.0340	5.4	0.0221	1.71*

* not included because partial $R^2 \leq .05$

EOT values of haemostasis variables

End of treatment values of PSact (p< .0001), FVIIC (p< .0001) and ATag (p< .0221) were discriminative of the treatments according to the following discriminant function (Table 6):

$$z_{EOTh} = -3.7 \times \log PSact_{EOT} + 2.83 \times \log FVIIC_{EOT} - 1.71 \times \log ATag_{EOT}$$

The associated mean z and classification accuracy for LNG/EE 150/30 and DSG/EE 150/30 was 0.76 and -1.27, and 77 and 91%, respectively (Table 7).

Application of previously formulated discriminant functions

In Chapter 5, we defined discriminating functions based on the change from baseline of all and haemostasis variables ($z_{CFBa/h}$) and the EOT (third cycle) values of these variables ($z_{EOTa/h}$) for two EE-containing CHCs of which, one contained LNG as progestin and the other Nestorone® [4]. To investigate the performance of these discriminant functions (listed below), they were applied to the data on the seven COCs. In the previous chapter, factor VII total (FVIIt) was the only available measurement of FVII. For the present evaluation FVIIag was used where the function requested data on FVIIt as they are closely similar.

$$z_{CFBa} = 3.2 \times \Delta\log SHBG + 15.6 \times \Delta\log FVIIag - 15.6 \times \Delta\log PLG$$

$$z_{EOTa} = 5.9 \times \log SHBG_{EOT} + 4.0 \times \log FVIIag_{EOT}$$

$$z_{CFBh} = 18.8 \times \Delta\log FVIIag - 19.3 \times \Delta\log PLG$$

$$z_{EOTh} = 8.5 \times \log FVIIag_{EOT}$$

Table 6 Mean z and classification accuracy (CA in %) per treatment of the discriminant functions formulated by discriminant analysis on the seven COCs using end of treatment (EOT) values or change from baseline (CFB) values

	End of treatment				Change from baseline			
	z_{EOTa}		z_{EOTh}		z_{CFBa}		z_{CFBa}	
Treatment	Mean	CA (%)	Mean	CA (%)	Mean	CA (%)	Mean	CA (%)
LNG/EE 150/30	1.74	93 (N=119)	1.22	93 (N=119)	1.68	95 (N=113)	1.38	90 (N=124)
LNG/EE 125/50	-0.85	73 (N=59)	0.32	73 (N=59)	0.96	83 (N=57)	0.67	71 (N=59)
GSD/EE 75/20	0.41	61 (N=81)	0.10	61 (N=81)	-0.19	64 (N=75)	0.09	43 (N=82)
DSG/EE 150/20	0.63	64 (N=70)	-0.42	64 (N=70)	-0.58	71 (N=63)	-0.48	69 (N=71)
GSD/EE 75/30	0.69	77 (N=74)	-0.50	77 (N=74)	-0.68	77 (N=74)	-0.70	76 (N=78)
NGM/EE 250/35	0.56	75 (N=79)	-0.58	75 (N=79)	-0.73	79 (N=75)	-0.64	70 (N=81)
DSG/EE 150/30	-1.74	83 (N=72)	-0.85	83 (N=72)	-1.31	91 (N=68)	-1.05	77 (N=74)

$z_{CFBa/h}$: discriminating function z based on CFB for all variables and the haemostasis subset
$z_{EOTa/h}$: discriminating function z based on EOT values of all variables and the haemostasis subset

Table 7 Mean z and classification accuracy (CA in %) per treatment resulting from applying the discriminant functions formulated in Chapter 5 of this thesis

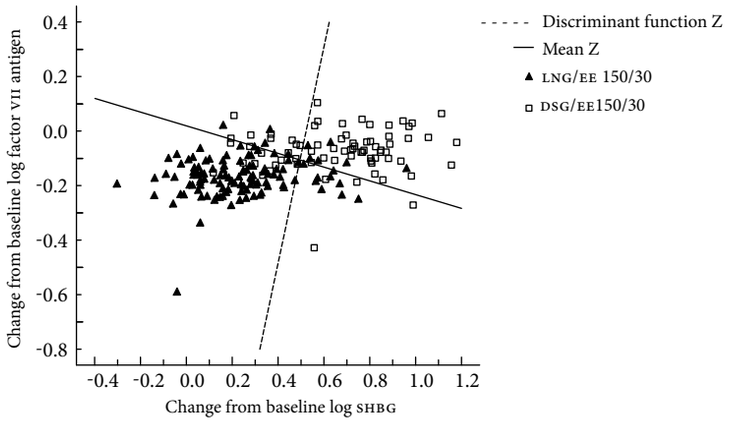
	Change from baseline				End of treatment			
	z_{CFBa}		z_{CFBh}		z_{EOTa}		z_{EOTh}	
Treatment	Mean	CA (%)	Mean	CA (%)	Mean	CA (%)	Mean	CA (%)
LNG/EE (150/30)	-1.75	91 (N=119)	-1.18	82 (N=129)	-1.76	89 (N=121)	-0.45	69 (N=129)
LNG/EE (125/50)	-1.12	79 (N=61)	-0.82	77 (N=65)	-0.78	68 (N=62)	-0.30	63 (N=65)
GSD/EE (75/20)	0.39	56 (N=77)	0.35	57 (N=83)	0.37	59 (N=78)	0.09	59 (N=83)
DSG/EE (150/20)	0.91	82 (N=66)	0.70	65 (N=74)	0.59	75 (N=68)	0.13	58 (N=74)
GSD/EE (75/30)	0.63	69 (N=80)	0.51	62 (N=84)	0.64	72 (N=81)	0.21	63 (N=84)
NGM/EE (250/35)	0.63	72 (N=76)	0.27	55 (N=83)	0.44	65 (N=77)	0.12	59 (N=83)
DSG/EE (150/30)	1.17	84 (N=76)	0.71	73 (N=82)	1.34	83 (N=78)	0.40	68 (N=82)

$z_{CFBa/h}$: discriminating function z based on CFB for all and the haemostasis subset
$z_{EOTa/h}$: discriminating function z based on EOT values of all variables and the haemostasis subset

CFB of all variables

The function z_{CFBa} = 3.2 x Δlog SHBG + 15.6 x Δlog FVIIag – 15.6 x Δlog PLG classifies LNG/EE 150/30 and DSG/EE 150/30 in 91 and 84% of the cases, respectively (Figure 3). For the other five COCs, the classification accuracy ranged between 56 and 82% (Table 7).

Figure 3 Graph of change from baseline of log factor VII antigen against change from baseline of log SHBG with discriminant function z_{CFBa} = 3.2 x Δlog SHBG + 15.6 x Δlog FVIIag – 15.6 x Δlog PLG (dotted line) formulated in Chapter 5. Note that the classification accuracy based on FVIIag and SHBG is less than based on z_{CFBa} which also includes the change from baseline of plasminogen



EOT values of all variables

The classification accuracy for the function z_{EOTa} = 5.4 x log $SHBG_{EOTa}$ 4.0 x log $FVIIag_{EOT}$ was 89 and 83% for LNG/EE 150/30 and DSG/EE 150/30, respectively, and varied between 59 and 75% for the preparations other treatments (Table 7).

CFB of haemostasis variables

The function z_{CFBh} = 18.8 x Δlog FVIIag – 19.3 x Δlog PLG classifies LNG/EE 150/30 and DSG/EE 150/30 correctly in 82 and 73% of the cases, respectively, while it ranged between 55 and 77% for the other COCs (Table 7).

EOT values of haemostasis variables

The function z_{EOTh} = 8.5 x log $FVIIag_{EOT}$ is able to classify LNG/EE 150/30 and DSG/EE 150/30 with 69 and 68% accuracy, respectively. The classification accuracy ranged between 58 and 63% for the other COCs (Table 7).

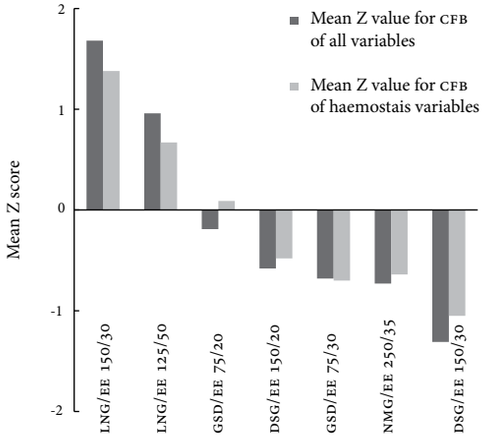
DISCUSSION

In this study we investigated the utilization of discriminant analysis to classify several combined oral contraceptives. The discriminant analysis was carried out using large series of haemostatic variables including all those recommended by the European Medicines Agency (EMEA) to evaluate for novel hormonal contraceptives, except for factor II [11].

All preparations contained EE as oestrogen, but contained four different progestins; levonorgestrel which is an androgenic progestin, gestodene and desogestrel that are strong progestins with low androgenic activity, and norgestimate which is a pro-drug partially metabolised to levonorgestrel. Consequently the treatments differed in oestrogenicity / androgenicity ratio.

The current study shows that the effects of seven COCs with different extents of oestrogenicity / androgenicity ratio are best discriminated by a combination of proteins which invariably included SHBG and measures of coagulation: factor VII and free protein S. There appears no major difference when either the change from baseline or the absolute end of treatment values are included in the discriminant analysis. Both approaches classify the COCs similarly, although for the haemostasis variables using the change of baseline performs slightly better. Whether the full set including SHBG was used or only the haemostasis subset did not matter for the ranking (Figure 4). When the analysis was limited to haemostasis variables only, the discriminative contribution of FVII and free protein S remained, but also a contribution of protein C, and plasminogen appeared among the most important variables, albeit with a smaller impact.

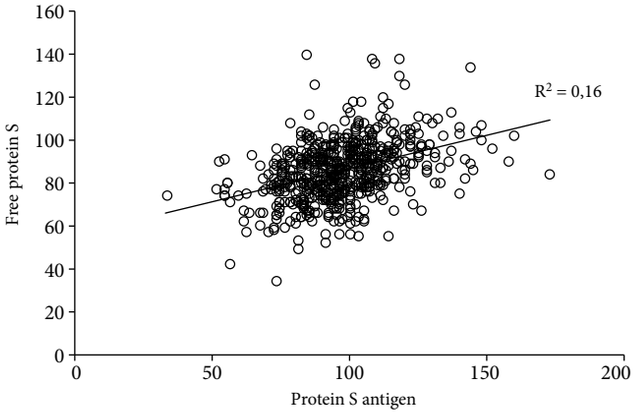
Figure 4 Mean Z scores for CFB of all variables and of haemostasis variables for all seven COCs ranked from high to low



This finding seems at odds with our previous findings in which protein S did not appear in the DA. This is, however, possibly explained as in the previous study protein S activity was included as measure. Although the variables protein S activity, protein S antigen and free protein S are obviously related, they convey different information and cannot be used interchangeably. We do not have direct comparisons of protein S activity and protein S antigen or free protein S antigen available, but can illustrate the situation with comparing protein S antigen and free protein S antigen (Figure 5).Total protein S is, in contrast to free protein S, insensitive to changes in the protein binding protein C4b. For example, the

influence of other clotting factors on results of the clotting assay for protein S activity is conceivable. Interestingly, comparable findings were reported by Lefkowitz et al on changes in multiple clotting factors occurring during pregnancy. They found differential effects on protein S activity compared with free protein S [12].

Figure 5 Correlation between free protein S and protein S antigen in 607 subjects of the study in the pre-treatment sample. Linear curve fit: $R^2 = 0.16$



The subsequent DA performed on the two COCs most widely separated in Z-value showed the importance of variables such as SHBG, PSact, and FVIIc, compared to those with lesser contributions. This suggests consistency of the DA and our approach.

Despite some differences between our two data sets, we consider it remarkable that we found a rather good classification accuracy of the seven COCs using a discriminant function that was formulated based upon two CHCs in Chapter 5 of this thesis [4]. We recognize that in the current data set more variables were necessary to accurately discriminate the seven COCs. We argue that this makes sense as treatments that are more similar in oestrogenicity / androgenicity ratio can only be classified with more information. In fact, the high partial R-squares and F-values suggest that SHBG, protein S, FVII and plasminogen are the most important discriminating factors in all algorithms, while other contributing variables probably merely optimize the classification accuracy.

Based on our DA analysis it appears that COCs show gradual differences on a continuous scale. This is illustrated in Figure 4 showing the COCs lined up in decreasing Z-value. It appears that classification of COCs as second- or third-generation will hold true only when the compared COCs lie at different ends of the spectrum. Consequently, it is even more unlikely that differences in risk on clinical endpoints such as VTE associated with CHCs can be explained by a rather arbitrary dichotomous classification of CHCs as is commonly done. Our data show that there is a gradual difference between COCs with regard to their effect on haemostasis variables and SHBG. This would suggest also a gradual difference in contribution to risk.

Another difference between the findings in current study as compared to those in Chapter 5 is identification of VIIag and FVIIc as important discriminators in the data set of the seven COCs, versus FVIIt in the previous report. When we repeated the DA of the seven COCs without FVIIc and FVIIact, FVIIag replaced FVIIc in the algorithm. This suggests that the change in the coagulant activity of FVII (as measured with the *in vitro* FVIIc test) is the best discriminatory test, but that the increase in FVII antigen is a major determinant. It might be surprising that FVII is highly discriminative of the treatments, as epidemiological reports normally show no significant contribution of this factor in VTE risk. However, FVII levels are mostly within its normal range among general population, while strongly increased and exceeding the normal range in our data sets. Although we are unaware of mechanistic studies relating relatively high or strongly elevated factor FVII levels above the normal range and VTE, our data suggest that they may be related. This suggestion warrants further investigations.

Similar to our previous report in Chapter 5, ext APCr did not emerge as a discriminating variable in the current analysis. The most likely explanation is the large overlap in ext APCr among the treatments [4]. This may suggest that the effect of different COCs on ext APCr is not materially relevant for the difference in the VTE risk among the studied set of COCs. Changes in ext APCr may contribute to an increase in risk particularly for carriers of the Factor V Leiden mutation, but not to a difference in risk for users of various COCs. The consistency in this and our previous report suggests that ext APCr may play a role in VTE risk, but to a lesser extent than previously suggested.

It is important to mention that DA has potential weaknesses. This type of analysis identifies variables that are most differently affected by different treatments. A low degree of inter-individual variability, beside a large difference in treatment effect, will increase the discriminative power of a variable. However, such discriminative power defines a large difference in biological status, but not necessarily pathophysiological consequences. That the identified variables are not contributors to VTE risk per se, is obvious as SHBG emerges in DA as highly discriminative of the treatments without being known for its involvement in VTE.

Within the haemostasis, variables that are presumed to fulfil potential roles in the pathophysiology of the disease are often selected to study with regard to the steroid-induced VTE risk. We realize that we are unable to guarantee the completeness of our set of variables subjected to DA. There may be other important contributors to VTE risk that were not included in our data set and may therefore have been missed.

Nevertheless, this report shows a set of haemostasis variables that may plausibly be related to the mechanism of steroid-induced VTE. Given the consistency of the results on factor VII, free protein S and plasminogen, we suggest that these variables are of prime importance. These variables should always be measured with appropriate, and preferably standardized, assays. DA offers the advantage of identifying combinations of variables that may play substantial roles in risk alteration. This is, at least intuitively, important when risk is a function of multiple variables, as is the case in steroid-induced VTE.

It is not possible to validate our findings with clinical data as no study combines measuring haemostasis variables with long-term follow-up to collect data on the relevant clinical endpoint: VTE. At present the interpretation of

clinical studies is different in different groups of investigators and regulatory agencies. One group and the EMEA follow the data of large retrospective studies on cohorts and acknowledges a difference between COCs with different oestrogenicity / androgenicity ratio's [5;6;13]. If this conclusion is justified, the results of our DA may indeed indicate that changes in FVIIc, PSact and plasminogen are highly significant in the pathophysiology of steroid-induced VTE. Another group and the FDA follow the data of active surveillance prospective studies concluding there is no proven difference in VTE risk associated with various COCs [7;8;14]. If true, this would imply that the differences we noted with the DA are not relevant. The markers we discussed are in essence markers of oestrogenicity / androgenicity ratio, so the question is whether oestrogenicity / androgenicity ratio is a determinant of VTE.

In conclusion, DA identified combinations of proteins synthesized in the liver and the endothelium: protein S, FVII and plasminogen as most discriminative of treatment effect of seven COCs that differ in oestrogenicity / androgenicity ratio. It is biologically plausible that these variables would be related to the pathogenesis of VTE. This information cannot be captured by any single (coagulation) variable. Clinical validation is required to judge the relevance of these findings.

REFERENCES

1 Hand DJ. Discrimination and classification. New York: John Wiley and Sons, Inc.; 1981.

2 Lachenbruch PA, Clarke WR. Discriminant analysis and its applications in epidemiology. Methods Inf Med 1980 October;19(4):220-6.

3 Dillion W, Goldstein M.Multivariate analysis: Methods and Applications. New York: John Wiley and Sons Inc.; 1984.

4 Rad M, Burggraaf J, de Kam ML, Cohen AF, Kluft C. Discriminant analysis of the metabolic effects of a new combined contraceptive vaginal ring containing Nestorone/EE vs. a second-generation oral contraceptive containing levonorgestrel/EE. Contraception 2012 September;86(3):231-7.

5 Lidegaard O, Milsom I, Geirsson RT, Skjeldestad FE. Hormonal contraception and venous thromboembolism. Acta Obstet Gynecol Scand 2012 July;91(7):769-78.

6 Lidegaard O, Nielsen LH, Skovlund CW, Skjeldestad FE, Lokkegaard E. Risk of venous thromboembolism from use of oral contraceptives containing different progestogens and oestrogen doses: Danish cohort study, 2001-9. BMJ 2011;343:d6423.

7 Dinger JC, Heinemann LA, Kuhl-Habich D. 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 2007 May;75(5):344-54.

8 Shapiro S, Dinger J. Risk of venous thromboembolism among users of oral contraceptives: a review of two recently published studies. J Fam Plann Reprod Health Care 2010 January;36(1):33-8.

9 Hemostasis study Group. An open label, randomized study to evaluate the effects of seven monophasic oral contraceptive regimens on hemostatic variables. Outline of the protocol. Oral Contraceptive and Hemostasis Study Group. Contraception 1999 June;59(6):345-55.

10 Hemostasis study Group. The effects of seven monophasic oral contraceptive regimens on hemostatic variables: conclusions from a large randomized multicenter study. Contraception 2003 March;67(3):173-85.

11 European Medicines Agency – Committee for MedicinalProducts for Human Use. Guideline on clinical investigation of steroid contraceptives in women. 2005; EMEA/CPMP/EWP/519/98 Rev 1.

12 Lefkowitz JB, Clarke SH, Barbour LA. Comparison of protein S functional and antigenic assays in normal pregnancy. Am J Obstet Gynecol 1996 September;175(3 Pt 1):657-60.

13 EMEA. Benefits of combined hormonal contraceptives (CHCs) continue to outweigh risks – CHMP endorses PRAC recommendation 22 November 2013, EMEA/709120/2013, 2013.

14 Woodcock J, Response to Docket FDA-2007-8-0190.