Materials and methods

1. Data source: five National Health Surveys

Data analyzed and evaluated in this dissertation come from the five National Health Surveys that were conducted by the Institute of Social Medicine and Epidemiology of the former Federal Health Administration (Bundesgesundheitsamt, BGA) and the Robert Koch-Institute (Berlin) from 1984 to 1999 in Germany. All original data regarding drug utilization are still saved by the Pharmacoepidemiology Group, Department of Epidemiology for Non-infectious Diseases and Environmental Medicine in Robert Koch-Institute (Berlin). Overall approximately 25000 German residents under ambulant care were enrolled as representative samples of the population in the five surveys (see Table 3).

Table 3: General data of the five National Health Surveys in Germany

Surveys	Time	Area	Sample size	Men (%)	Women (%)	Age range
T0	1984-1985	West Germany	4790	2417 (50.5)	2373 (49.5)	25-69
T1	1987-1988	West Germany	5335	2649 (49.7)	2686 (50.3)	25-69
T2	1990-1991	West Germany	5311	2623 (49.4)	2688 (50.6)	25-69
T3	1991-1992	East Germany	2211	1051 (47.5)	1160 (52.5)	25-69
BGS98	1998-1999	Unified Germany	7124	3450 (48.4)	3674 (51.6)	18-79
		-West Germany	*-4705*	*-2297 (48.8)*	*-2408 (51.2)*	
		-East Germany	*-2419*	*-1153 (47.7)*	*-1266 (52.3)*	
Total	**1984-1999**	**Germany**	**24771**	**12190 (49.2)**	**12581 (50.8)**	**18-79**

In the frame of German preventive studies for cardiovascular diseases (Deutsche Herz-Kreislauf-Präventionstudie, DHP), aiming at the primary prevention of cardiovascular diseases in the general population under ambulant care with focus on cardiovascular risk factors, T0, T1 and T2 were performed continuously every three years from 1984 to 1991 with the same standardized main questionnaire [93]. Details about the design of National Health Surveys, study participants, collection and management of data are to be found in the special publications of Robert Koch-

Institute [94,95]. Soon after the reunification of East Germany and West Germany, T3 was performed in an effort to compare the differences in health status and utilization of health services between the two parts of Germany [96,97]. Based on the previous experience, the survey BGS98 was greatly improved in design such as the content of main questionnaire. Moreover, the age range of participants was extended to 18-79 years, covering more residents and thus is more representative of the German national population [98].

Each survey consisted of three parts at large, namely the main questionnaire, the questionnaire for drug use and a medical examination for participants, all of which were combined together through record-linkage for further analysis (see Fig. 2). All participants enrolled in the five surveys were medically interviewed by health professionals and a blood sample was drawn for various measurements in the laboratory such as biochemistry tests for blood, liver and kidney functions, etc.. Drug use data were collected with a standardized questionnaire (Arzneimittelfragebogen, see appendix 1) covering all drug use in the last seven days before the medical interview. The usage of some drugs such as digoxin, theophylline, diazepam, aspirin (acetylsalicylic acid), caffeine, vitamin E (tocopherols), etc. was confirmed by determining the drug concentrations in serum. This study focuses on participants who used steroid hormones for contraception and for hormone replacement therapy.

2. Identifying steroid hormone users in the five National Health Surveys

The drug use questionnaire recorded all drugs including prescribed and OTC medications taken by participants in the last 7 days. For every drug recorded, a specific code was given either according to the classification of European Pharmaceutical Market Research Association (EPHMRA-code) in the surveys T0, T1, T2 and T3 or according to the classification of Anatomical Therapeutic Chemical (ATC-code) in the survey BGS98 [99]. Steroid hormones with EPHMRA-codes and ATC-codes that were listed in the Table 4 were hormones used exclusively either for contraception or for HRT, and were identified from the medicines database of each survey. Phytoestrogens, male steroid hormones and female steroid hormones used for the treatment of various gynecological diseases such as amenorrhea, dysmenorrhea and endometriosis etc. were not included in this dissertation, but are in the RKI-database.

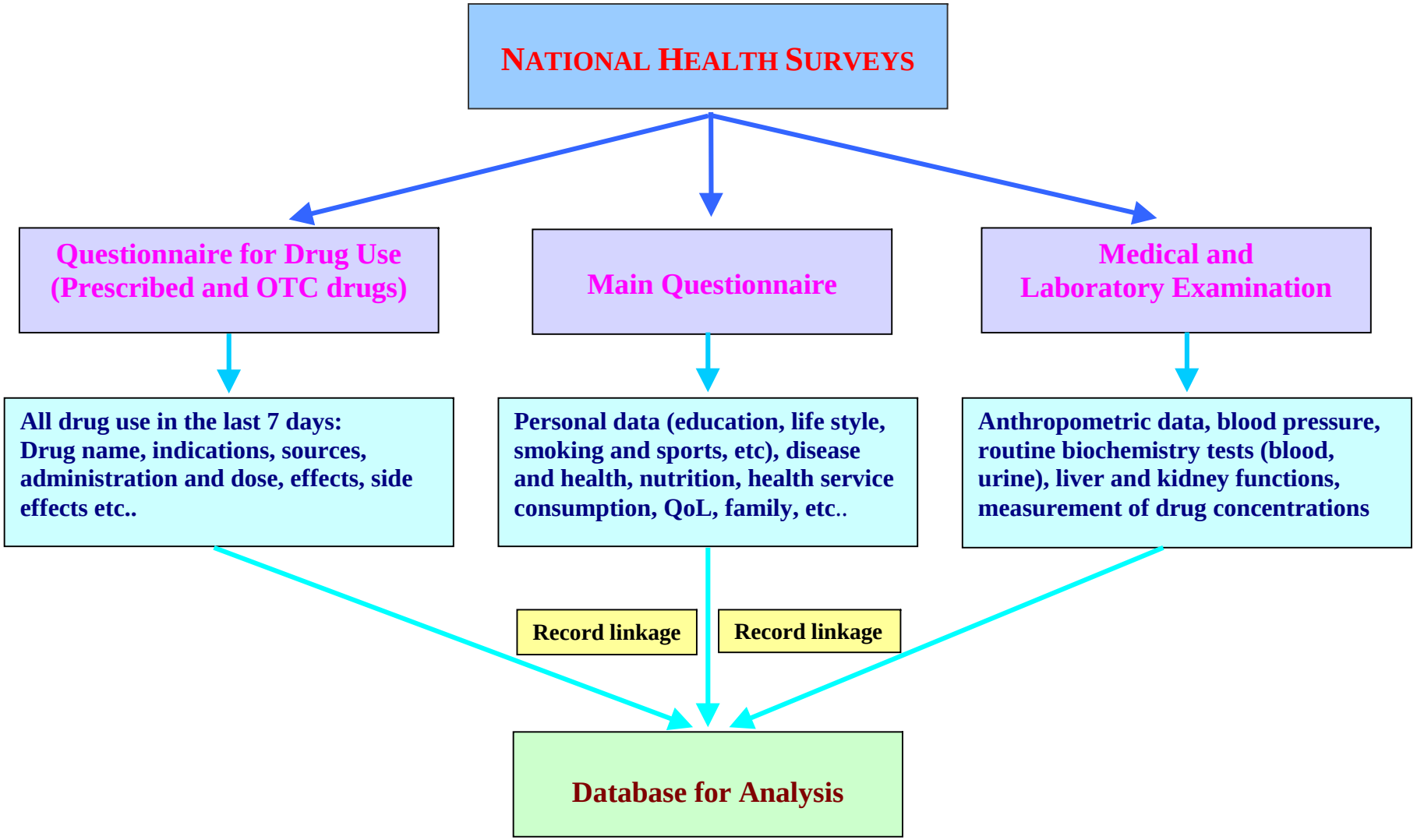

Fig. 2: Contents of National Health Surveys in Germany from 1984 to 1999

Table 4. Steroid hormones used for contraception and for hormone replacement therapy: EPHMRA- and ATC-code and their implications

EPHMRA-code		ATC-code	
Codes	***implications***	***Codes***	***Implications***
G03A	Contraceptives	**G03A**	Hormone contraceptives for systemic use
G03A1	Monophasic contraceptives	**G03AA**	Progestogens and estrogens, fixed combination
G03A2	Biphasic contraceptives	**G03AB**	Progestogens and estrogens, sequential preparations
G03A3	Triphasic contraceptives	**G03AC**	Progestogens
G03A4	Progestogen-only contraceptives	**G03C**	Estrogens
G03C	Estrogens and combinations	**G03CA**	Natural and semisynthetic estrogens, plain
G03C1	Estrogens only	**G03CB**	Synthetic estrogens, plain
G03C2	Combined estrogens	**G03CC**	Estrogens, combinations with other drugs
G03D	Combined progestogene or progestogene only,	**G03D**	Progestogens
G03E	Female sex hormone with androgene,	**G03DA**	Pregnene derivatives
G03F	Combined estrogens and progestogene	**G03DB**	Pregnadiene derivatives
G03H	Other sex hormone products	**G03DC**	Estren derivatives
		G03E	Androgens and female sex hormones in combination
		G03EA	Androgens and estrogens
		G03EB	Androgens, progestogens and estrogens in combination
		G03EK	Androgens and female sex hormones in combination with other drugs
		G03F	Progestogens and estrogens in combination
		G03FA	Progestogens and estrogens, combinations
		G03FB	Progestogens and estrogens, sequential preparations
		G03HB	Antiandrogens and estrogens

Totally 1365, 368 and 1090 uses of steroid hormones with one of the above EPHMRA-codes or ATC-codes in Table 4 were identified from surveys T0-T2, survey T3 and BGS98, respectively. 15 and 21 women in surveys T0-T2 and BGS98, respectively, had taken two different hormone agents, for which the main hormone agent was adopted. Therefore, 2787 users of female sex hormones were identified

from the five German National Health Surveys, of which 1862 and 827 women used steroid hormones for contraception and for HRT, respectively (Table 5).

Table 5: Gender of users and indications of steroid hormones in the five National Health Surveys from 1984 to 1999 in Germany

Survey	Gender of users		Indications			Total(%)
	Men (%)	Women (%)	OC(%)	HRT(%)	Others(%)	
T0 1984-1985	1(0.3)	321(99.7)	258(80.4)	46(14.3)	18(5.3)	322(100)
T1 1987-1988	1(0.2)	458(99.8)	353(76.9)	81(17.6)	25(5.4)	459(100)
T2 1990-1991	1(0.2)	568(99.8)	328(57.6)	210(36.9)	31(5.4)	569(100)
T3 1991-1992	0(0)	368(100)	303(82.3)	63(17.1)	2(0.6)	368(100)
BGS98 1998-1999	5(0.5)	1064(99.5)	620(58)	427(39.9)	22(2.1)	1069(100)
Total 1984-1999	8(0.3)	2779(99.7)	1862(66.8)	827(29.7)	98(3.5)	2787(100)

3. Randomly choosing controls for the users of oral contraceptives and HRT

Considering that hormone use is associated with sociodemographic factors [100] and the marked age difference between hormone users and nonusers in the general population, for example, OC users tended to be younger women, whereas HRT users tended to be those from higher-age groups, it is necessary to choose controls randomly among nonusers for the further comparison. In order to yield an age distribution similar to that of hormone users, controls were chosen according to following procedures.

In a specific age group (5 years), with the help of the random-program of SPSS, controls of hormone users were obtained by choosing, according to the certain proportion p, from hormone nonusers that were calculated by the total women population subtracting hormone users in the same age group (Table 6). The value of p was determined by the relative proportion between the numbers of hormone users (a_i) and of nonusers (A_i-a_i), ranging from 1 to 5. If A_i-a_i < a_i, nonusers may be supplemented from the nearest age group of a_i as the controls of a_i.

Table 6: Choosing of controls for steroid hormone users in a specific National Health Survey

Age group (years)	women Population	No. of hormone users	No. of hormone nonusers	Randomly chosen from hormone nonusers (A_i-a_i) (1:p)*	No. of controls for hormone users
17-19	A_1	a_1	A_1-a_1		a_1p
20-24	A_2	a_2	A_2-a_2		a_2p
25-29	A_3	a_3	A_3-a_3		a_3p
30-34	A_4	a_4	A_4-a_4		a_4p
35-39	A_5	a_5	A_5-a_5		a_5p
40-44	A_6	a_6	A_6-a_6		a_6p
45-49	A_7	a_7	A_7-a_7	→	a_7p
50-54	A_8	a_8	A_8-a_8		a_8p
55-59	A_9	a_9	A_9-a_9		a_9p
60-64	A_{10}	a_{10}	A_{10}-a_{10}		$a_{10}p$
65-69	A_{11}	a_{11}	A_{11}-a_{11}		$a_{11}p$
70-74	A_{12}	a_{12}	A_{12}-a_{12}		$a_{12}p$
75-79	A_{13}	a_{13}	A_{13}-a_{13}		$a_{13}p$
Total	**∑A**	**∑a**	**∑(A-a)**		**p∑a**

***: p=1-5, depending on the relative proportion of hormone users and nonusers in the same age group; by using random-program of SPSS software;**
∑a: total hormone users in a specific survey; p∑a: total controls of hormone users in this survey.

Therefore, we could obtain the study subjects in this dissertation, which accounted for 44.25% (1050/2373), 50.97% (1369/2686), 67.86% (1824/2688), 68.53% (795/1160) and 68.62% (2521/3674) all women population in survey T0, T1, T2, T3 and BGS98, respectively (Table 7). In survey BGS98, contraceptive users, though 1 year younger than controls on average overall, who aged over 25 years had no significant difference in a 5-year age distribution in comparison to their controls, the same as in the other four surveys. Further, no difference could be found for the unweighted data. So, one-year difference in average age between contraceptive users and controls in BGS98 presented more statistical meaning. The actual influence on later analysis regards to disease and health should be very small.

Table 7: Steroid hormone users and their controls in the five German National Health Surveys

Surveys	Women population	Oral contraceptives		HRT		Total
		Users	Controls	Users	Controls	
T0	**2373**	258 (33.26±6.48)	516 (33.48±6.70)	46 (56.50±6.54)	230 (56.43±6.70)	1050
T1	**2686**	353 (32.34±6.62)	530 (32.68±6.41)	81 (51.84±7.89)	405 (52.13±7.85)	1369
T2	**2688**	328 (32.95±6.63)	656 (33.11±6.60)	210 (52.86±6.92)	630 (53.29±6.37)	1824
T3	**1160**	303 (33.85±6.10)	303 (35.34±6.15)	63 (51.77±6.47)	252 (51.45±6.49)	921
BGS98	**3674**	620 (29.75±7.92)*	620 (30.76±7.77)	427 (56.5±8.19)	854 (56.10±8.10)	2521
Total	**12581**	**1862**	**2625**	**827**	**2371**	**7685**

( ): the average age of users or controls, mean±SD; all means were weighted according to national population to represent population values.

***: p=0.027 comparison with controls in the same survey (comparing in unweighted data: 30.10±8.27 vs. 30.69±7.93, p=0.199)**

4. Statistical methods

In the first part of 'Results', 'DESCRIPTIVE EPIDEMIOLOGICAL STUDIES', primary analyses were focused mainly on the cross-sectional difference between steroid hormone users and controls who were matched on age ± 5 years. Student-t test and χ^2 or Fisher's exact tests served to compare steroid hormone users and controls within the same national health survey for continuous variables and for categorical variables, respectively. In the second part of 'Results', 'ANALYTICAL EPIDEMIOLOGICAL STUDIES', which embraces 4 independent case-control studies, ORs and 95% confidence intervals (95% CI) were calculated to represent relative risks (RR). Given the fact that our controls were very 'crude', matched only on age ±5 years, unconditional logistic regression models were used to assess the independent contribution of covariables to the use of steroid hormones or the occurrence of diseases. ORs and 95% confidence intervals (95% CI) were obtained

from the unconditional logistic regression models. The significance of regression coefficients was evaluated using the Wald test.

Statistical analyses were conducted using SPSS version 11.5.2.1, incorporating the examination sampling weights of five National Health Surveys when necessary. Statistical tests with $p<0.05$ were considered statistically significant.