Mandatory notification of newborns with different congenital abnormalities (CAs) in Hungary was ordered by the Ministry of Health in 1962. This disposition was warranted by the greatest tragedy of human teratology, the thalidomide (Contergan) disaster. A traditional mandatory notification system for infectious diseases had been working in Hungary for years, but the significance of congenital abnormalities in infant mortality was found to be fourfold higher than that of infectious diseases. Thus, Hungary established the first national-based registry of congenital abnormalities in the world.

However, it soon became clear that the evaluation of these cases needed epidemiological expertise. Therefore, based on international recommendations this registry was further developed and formed the base of the Hungarian Congenital Abnormality Registry (HCAR) at the National Institute of Public Health (NIPH) from 1970. After having detected certain CA clusters by the HCAR a population-based surveillance of cases with CA and controls, The Hungarian Case-Control Surveillance of Congenital Abnormalities (HCCSCA) was established in 1980. This registry worked in its original form until 1996. The professional network led by Prof. Andrew E. Czeizel thoroughly analyzed the data of this 17 years, publishing more than 800 scientific papers based on this database.

The collection of data was changed in 1997, slightly modifying the structure of the HCCSCA also. Cases and controls notified after 1997 formed another voluminous database that has not been analyzed yet.

Therefore, our first task was to examine and validate the records of this second part of the HCCSCA.

According to computing standards of those decades data were collected into text files between 1980 and 1996. From 1997 these data were recorded using Microsoft Excel datasheets. Slight coding differences were also present between the two segments of the database. Therefore, as the first step data of the first 17 years were converted from text files into Excel sheets. Coding differences were also resolved, and data were validated excluding erroneously recorded cases and controls. Finally, the 2 segments of the database were united into one datasheet. Thus, we managed to unify all data collected in the HCCSCA between 1980 and 2009 into a validated, single database that is now open for examination.

The following data are available for each case and control pregnancy: CA(s), gender, birth year/month/date, birth weight, gestational age, area of mother's living, maternal age, paternal age, birth order, mother's and father's qualification, employment status and type of employment, mother's marital status, outcome of previous pregnancies, maternal diseases during pregnancy (according to pregnancy months), drug intake during pregnancy (according

to pregnancy months), folic acid and/or pregnancy vitamin supplement intake (according to pregnancy months), mother's smoking habits and alcohol consumption patterns.

According to our knowledge, until now, the HCCSCA has become the largest case-control data set of CA-surveillance in the world. This is an open source of data for any researcher in the field.

The merit of the HCCSCA is not only limited to its size but it has an excellent quality of data as well. The surveillance is population-based (the population of Hungary is about 10 million), and the validity of CAs is excellent due to hospitalization for delivery. In Hungary more than 99 % of deliveries take place at obstetrical inpatients departments, therefore, birth weight and gestational age are always medically recorded. Nearly all infant deaths are autopsied, thus congenital anomalies are validated by pathologists. In addition, the diagnoses of CAs were reported by medical doctors and underwent several checking procedures by experts. Each CA case has at least two matched controls without any CA. In general, the diagnoses of maternal diseases are also scientifically accurate due to mandatory prenatal care in Hungary. Most maternal diseases were recorded prospectively by medical doctors in the prenatal logbook.

As it was mentioned previously, more than 800 papers have been published based on data in the HCCSCA collected between 1980 and 1996. The complete database now includes an additional 9,615 deliveries with CAs and 20,287 control newborns, altogether 29,902 new pregnancies. Thus, we are able to expand the previous results with much larger sample sizes to include the more recent data. The addition of these data also allows for greater power to detect associations between various maternal illnesses, drug intakes and other factors and less frequent birth defects. The immense amount and the excellent validity of data included in our database probably makes the HCCSCA a unique source of scientific information.

Besides this validation and computing work, we continuously analyzed data of some illness groups. The detailed data of our papers published during the study period based on this database are given below.

Isolated anorectal malformation

In most patients affected by isolated anorectal malformation (IARM) the etiology is largely unknown. Thus, the aim of our project was to analyze possible risk factors for IARM. In the first step, birth outcomes of cases with IARM were analyzed on the basis of maternal socio-demographic variables. Gestational age at delivery, birth weight, preterm birth, low

birthweight and small for gestational age of cases with IARM were evaluated in the function of maternal age, birth/pregnancy order, marital and employment status of mothers in the population-based large dataset of the Hungarian Case-Control Surveillance of Congenital Abnormalities. The study samples included 231 live-born cases with IARM, 361 matched and 38,151 population controls without any defect. IARMs are more frequent in males, twins and newborn infants with low birthweight and small-for-gestational-age, the latter being the consequence of intrauterine growth restriction. In addition, mothers of cases were younger but with higher birth order, and had lower socio-economic status. These maternal variables are characteristic for the gypsy population in Hungary. The higher proportion of gipsy women among the mothers of cases with IARM was confirmed during the home visits of the study. Male sex and intrauterine growth restriction of cases, in addition low socioeconomic status and gypsy origin of mothers may have a role in the risk of IARMs.

Another aim of our studies was to analyze the possible role of maternal risk factors in the origin of IARM. Thus, cases with IARM were evaluated in the function of maternal diseases and related drug use. The findings of this case-control study suggested that cases with IARM have an obvious male excess. The mothers of cases with IARM had a lower incidence of severe nausea and vomiting in pregnancy and a higher incidence of acute infectious diseases in the urinary tract. Lack of nausea and vomiting in pregnancy and the higher incidence of urinary tract infections may have a role in the development of IARM.

Isolated infantile hypertrophic pyloric stenosis

In most patients affected by isolated infantile hypertrophic pyloric stenosis (IHPS) beyond the polygenic predisposition the other factors in the multifactorial etiology are largely unknown. The main characteristic of IHPS is the robust male predominance, thus the aim of this study was to analyze birth outcomes in males and females whether they are different or not. The study samples included 241 cases with IHPS, 357 matched and 38,151 population controls without any defect in the population-based large dataset of the Hungarian Case-Control Surveillance of Congenital Abnormalities. The findings of this case-control study confirmed the well-known strong male excess (85.5%). Mean gestational age was somewhat longer and it associated with a lower rate of preterm births. Mean birth weight did not show significant differences among the study groups but the rate of low birthweight was higher in cases with IHPS. However, these differences were found only in males. Thus, intrauterine fetal growth restriction is characteristic only for male cases with IHPS. Our study confirmed the well-known obvious male excess of cases with IHPS, bur our findings suggest some differences in

birth outcomes of male and female cases. Male cases with IHPS had intrauterine fetal growth restriction while females have not. These data may indicate some differences in the pathogenesis of IHPS in males and females.
Another aim of this study was to estimate possible maternal risk factors in the origin of IHPS. Exposures that had been medically recorded in prenatal maternity logbooks during the critical period of IHPS were evaluated separately. The findings of this case-control study suggested that - beyond the well-known robust male excess (85.5%) – maternal hyperthyroidism (OR with 95% CI: 4.17, 1.53-11.38) and oral nalidixic acid treatment (OR with 95% CI: 6.43, 3.03-14.06) is associated with a higher risk for IHPS in their children.

Isolated oesophageal atresia

In most patients affected by isolated oesophageal atresia (IOA) the etiology is largely unknown. Thus, the aim of this study was to analyze potential risk factors in mothers.
The study samples included 221 cases with IOA, 356 matched and 38,151 population controls without any defect in the population-based dataset of the Hungarian Case-Control Surveillance of Congenital Abnormalities. Only those exposures were evaluated that were medically recorded in prenatal maternity logbooks during the critical period of IOA. The findings of this case-control study suggested that the mothers of cases with IOA had a higher proportion of first delivery and lower socioeconomic status. Acute respiratory diseases (OR 95% CI: 3.8, 1.8-8.1) and essential hypertension treated with nifedipine (OR 95% CI: 3.8, 1.7-8.7) in the mothers of cases associated with a higher risk for IOA in their children.

Paroxysmal supraventricular tachycardia

The aim of the study was to estimate the possible association of pregnant women with paroxysmal supraventricular tachycardia (PSVT) with the possible risk for adverse birth outcomes, particularly different congenital abnormalities (CAs) in their children.
Prospectively and medically recorded PSVT was evaluated in 103 pregnant women who later had offspring with CA (case group) and 149 pregnant women who later delivered newborn infants without CA (control group) and matched to cases in the population-based data set of the Hungarian Case-Control Surveillance System of Congenital Abnormalities. Of 252 pregnant women with PSVT, 115 (45.6%) had the onset of this condition before the study pregnancy, that is, their PSVT was a chronic condition, while the rest (N = 137) of PSVT was considered as new onset in the study pregnancy. The comparison of occurrence of PSVT in pregnant women who had offspring with different CA groups and in control mothers showed

a higher risk for cardiovascular CAs (adjusted OR with 95% CI: 2.1, 1.1-3.8) explained mainly by secundum atrial septal defect. This association was confirmed in pregnant women with PSVT in the second and/or third gestational month, that is, critical period of cardiovascular CAs. In conclusion PSVT in pregnant women associates with a higher risk of secundum atrial septal defect in their children.

The rate of undescended testis in Hungary

Undescended testis (cryptorchidism) is a common congenital abnormality of the male genital organs and increasing trend in its birth prevalence was reported in some countries. The aim of this study was to analyze the recorded annual birth prevalence of isolated undescended testis (IUT) in the population-based large dataset of the Hungarian Congenital Abnormality Registry for the period between 1962 and 2011, i.e. during the last 50 years. Cases with IUT reported after births were evaluated, and their annual rate per 1000 live-births was calculated. The rates of cases with IUT were compared with the so-called true rate of IUT measured in a previous clinical-epidemiological study based on the personal examination of 10,203 newborn infants. The birth prevalence of cases with recorded IUT in Hungary was lower than expected based on the true rate of IUT. Thus, the two waves in the rate of IUT were connected with the different completeness of reporting. In conclusion the birth prevalence of cases with IUT in Hungary did not indicate a real increasing trend during the last 50 years.

Preventive effect of folic acid for isolated hypospadias

Hypospadias is a common structural birth defect of the male genital organ. The objective of this study was to test the hypothesis regarding the possible preventive effect of folic acid for isolated hypospadias (IH). Folic acid use was compared in 3,038 cases with IH and 24,814 male controls without any defects in the national population-based Hungarian Case-Control Surveillance of Congenital Abnormalities. In Hungary only one kind of folic acid tablet, containing 3 mg/tablet was available during the study period. Hungarian obstetricians recommended daily use of 1-3 tablets, that is, 3-9 mg of folic acid during pregnancy; the estimated daily dose was 5.7 or 5.6 mg in the mothers of cases with IH and controls, respectively. Of 3,038 mothers of cases 1,474 (48.5%) were supplemented with high doses of folic acid during pregnancy and 13,509/24,814 mothers of controls (54.9%) as reference were supplemented at this level as well (OR 95% CI: 0.79, 0.73-0.85). If only medically recorded folic acid use in the critical period of IH was evaluated, the preventive effect was more apparent (OR 95% CI: 0.36, 0.32-0.41). The intake of folic acid among mothers of infants

with severe IH was lower than among mothers of infants with mild IH, suggesting a dose-response relationship. In conclusion, this study suggests that high doses of folic acid are associated with a reduced risk of IH. However, this analysis was based on observational data; therefore, confirmation in a well-controlled study is needed.

As a conclusion our group has successfully fulfilled the study plan during the period between 2014 and 2017. The greatest challenge has been the immense volume of data to be analyzed in the unified database. Our computers failed to handle these voluminous datasheets, therefore we had to buy new hardware equipped with extraordinary memory capacities and high-performance processors. At the end, however we managed to create the unified database of the HCCSCA, that is a really unique source of data in the world. This database is now open for analyses, giving us study perspectives for the oncoming years.

After that we solved the problem of differences between the structures of the two chapters. The details of this unified database are given in this paper.

Types of data recorded

Unifying the two segments of the database

Statistical Analysis

The software R was used for the analysis of variables. At the evaluation of quantitative variables such as mean maternal age and birth order Student t test was used. We used chi square test at the evaluation of distribution of maternal age, birth order and employment status. At the evaluation of categorical variables: pregnancy complications, maternal diseases, drugs and pregnancy supplements adjusted odds ratios (OR) with 95% confidence intervals (CI) were estimated. At the comparison of mothers of cases and matched controls multivariable conditional logistic regression model was used. The exposures in the mothers of cases and population controls were also compared and adjusted OR with 95% CI were calculated in multivariable unconditional logistic regression model. Among confounding factors maternal age (continuous variable), birth order (parity) (2 vs. 1, 3+ vs. 1) and employment status (skilled/semiskilled worker vs. professional/managerial, unskilled worker/other vs. professional/managerial) as indicator of socio-economic status were considered [4].

Results

The mean time elapsed between the end of pregnancy and return of the „information package" (including prenatal logbook, questionnaire, signed informed consent, etc.) was 3.5 ± 1.2, 5.2 ± 2.9, and 3.3 ±1.8 months in the case, control, and malformed control groups, respectively.

The numbers of cases and controls collected in the HCCSCA are presented in Table 1. Altogether we have 32,345 cases and 57,231 controls recorded in our database. During the study period there were 3,009,303 live-births in Hungary, thus our control sample represented 2.98% of total births. The overall incidence of CAs recorded in the HCCSCA during the study period

was 10.7/1,000 live-births (as mentioned earlier, some mild CAs of high prevalence, but of limited clinical relevance were excluded from the database).

The detailed numbers of specific CAs and CA groups in the HCCSCA are given in Table 2. The most frequent anomalies detected were ventricular septal defect (2,864), atrial septal defect (1,895), polydactyly (1,499), hypospadias (1,083) and unilateral cleft lip ± palate (961).

The age of mothers was 26.1±5.5 years and 27.5±8.4 years in the case and control groups, respectively.

We analyzed the prevalence of CAs as a function of maternal age. These data are given in Fig. 1. The lowest prevalence rates have been detected between the ages of 23 and 33 years. In general, younger and older maternal age are both associated with higher rates of birth defects.

According to ICD-10 altogether we found 701 diseases that affected case mothers during pregnancy. The 25 most frequently detected diseases during pregnancy among the case and control mothers are given in Table 3. The top 5 among case mothers were: common cold: 4,498, acute upper respiratory infections: 3,026, acute urinary tract infections: 2,265, influenza: 1,540 and hypertension: 1,365.

We identified 816 drugs that were taken by mothers during pregnancy. Apart from vitamin intake the most frequent ones are shown in Table 4.

We also examined the potential protective effect of folic acid or complex pregnancy supplement intake against the development of CAs. These data are given in Table 5. Although folic acid intake significantly decreased the prevalence of CAs, the protective effect of complex pregnancy supplements was found to be much more pronounced.

Discussion

In this paper we aimed to introduce some preliminary data from the new, completed HCCSCA to show the scientific opportunities provided by this database.

Analyzing maternal age as a risk factor for CAs is just an interesting example for the opportunities of our database. Based on the literature data on this question are inconclusive. Analyzing data from the first part of the HCCSCA Csermely et al. have described the same trends, younger and older maternal age being a risk factor for birth defects [5]. On the other hand, Goetzinger at al. have found that advanced maternal age decreases the risk of CAs [6]. One possible explanation for this obvious difference may be that this study group compared data of

This system is appropriate for postmarketing surveillance of drug teratogenicity, for the improvement of congenital abnormality diagnosis, to get informed consent, to have a communication with parents and to provide material for research.

We are absolutely open for any scientific cooperation based on these data.