



---

## Joint Data Analyses of European Birth Cohorts: Two Different Approaches

**Corresponding Author:**

Ms. Cynthia Hohmann,  
Research Associate, Charite Universitaetsmedizin, Berlin; Institute for Social Medicine, Epidemiology and Health Economics, 10117 - Germany

**Submitting Author:**

Ms. Cynthia Hohmann,  
Research Associate, Charite Universitaetsmedizin, Berlin; Institute for Social Medicine, Epidemiology and Health Economics, 10117 - Germany

**Article ID:** WMC003869

**Article Type:** Review articles

**Submitted on:** 12-Dec-2012, 11:32:54 AM GMT **Published on:** 12-Dec-2012, 04:48:12 PM GMT

**Article URL:** [http://www.webmedcentral.com/article\\_view/3869](http://www.webmedcentral.com/article_view/3869)

**Subject Categories:** EPIDEMIOLOGY

**Keywords:** European Birth Cohorts, Cohorts, Combined Analyses, Meta-Analyses, Networking

**How to cite the article:** Hohmann C , Govarts E, Bergström A, Cordier S, Eggesbo M , Guxens M, Heinrich J , Kasper - Sonnenberg M, Kogevinas M, Neuman A, Ranft U , Tischer C , Vandentorren S, Vardavas C , Patelarou E, Wickman M, Trnovec T, Nieuwenhuijsen M, Keil T , Bonde J. Joint Data Analyses of European Birth Cohorts: Two Different Approaches . WebmedCentral EPIDEMIOLOGY 2012;3(12):WMC003869

**Copyright:** This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC-BY\)](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Source(s) of Funding:**

**Funding:** This work was funded by **ENRIECO** (Environmental Health Risks in European Birth Cohorts), a project conducted within the European Union's 7th Framework Programme (FP7/2007-2013) under grant agreement number: 226285. Further funding was received under grant agreement OBELIX 227391, grant agreement CLEAR 226217, and with collaboration of ArcRisk agreement 226534.

**Competing Interests:**

All authors read and approved the final manuscript. The authors declare that they have no conflict of interest.

# Joint Data Analyses of European Birth Cohorts: Two Different Approaches

**Author(s):** Hohmann C , Govarts E, Bergström A, Cordier S, Eggesbo M , Guxens M, Heinrich J , Kasper - Sonnenberg M, Kogevinas M, Neuman A, Ranft U , Tischer C , Vandentorren S, Vardavas C , Patelarou E, Wickman M, Trnovec T, Nieuwenhuijsen M, Keil T , Bonde J

## Abstract

**Background:** Combined data analyses of birth cohorts can overcome the fragmentation of individual and inconclusive results obtained by analyses based upon single cohorts only. The European project Environmental Health Risks in European Birth Cohorts (ENRIECO) undertook four combined studies to evaluate the concept of added scientific value through harmonisation and exchange of birth cohort data for common analyses on environmental health risks.

**Objectives:** Two alternative analytical approaches were evaluated regarding their feasibility, benefits and limitations: (1) the *centralised approach* (pooled and non-pooled analyses which were centrally conducted by a single case study leader) and (2) the *decentralised approach* (meta-analysis of summary statistics derived by uniform methods conducted in each cohort).

**Methods:** Four main steps were identified for database building and analyses: (I) eligibility of cohorts, (II) collection of individual participant data, (III) data verification and (IV) analyses and manuscript preparation.

**Discussion:** The decentralised approach is recommended if cross-border data transfer is difficult and/or a solid basis of trust and experience still has to be established among partners. The centralised approach is recommended for combined analyses addressing variables with very heterogeneous assessments across cohorts, where a flexible handling of data is essential and trust and work experience between participating partners already exists.

**Conclusion:** Both approaches were successful, albeit laborious and time-consuming. Transparency through regular updates, presentation of results from interim analyses and the possibility for birth cohort researchers to comment and agree to each step of the analysis process builds trust and forms the basis for a sustainable collaboration.

## Background - Collaboration of

## Birth Cohorts

Over the past 30 years a large number of mother-child cohorts (birth cohorts) with various objectives have been established across Europe [1]. Although these birth cohorts were originally designed for stand-alone analyses, on a whole they build a unique data resource for investigating risk factors and determinants of pregnancy outcomes and child health. Recently, there has been a focus on collaboration between cohorts and networks have been established and maintained by projects such as GA<sup>2</sup>LEN (Global Allergy and Asthma European Network) [2-3], OBELIX (OBesogenic Endocrine disrupting chemicals: LInking prenatal eXposure to the development of obesity later in life) [4], ENRIECO (Environmental Health Risks in European Birth Cohorts) [1] or CHICOS (Developing a Child Cohort Research Strategy for Europe) (<http://www.chicosproject.eu/>). An overview of available cohort data is available under [www.birthcohorts.net](http://www.birthcohorts.net).

Combined data analyses of birth cohorts can overcome the fragmentation of individual and inconclusive results obtained by analyses based upon single cohorts only. The sample size of many cohorts is often not sufficient for analysing exposures and outcomes with low prevalences. Harmonising and combining their datasets increases statistical power allowing examination of exposure-response relationships and regional differences. Consistent findings across cohorts may add to the credibility of results and robust findings are expected to have a greater impact on clinical practice and public health. Reporting bias may be reduced by promoting the reporting of results from all cohorts regardless if an effect is observed or not. Expensive laboratory analyses of xenobiotic exposures and/or advanced outcome measures are often determined in subsets of birth cohorts only and may provide added value if combined with similar data in other cohorts. Given these advantages and a general positive attitude towards collaborating, the above listed existing European birth cohort networks as well as already published and numerous ongoing combined analyses

within these networks reflect the growing interest of birth cohort study teams to participate in combined data analyses.

The ENRIECO project was a coordination project with the aim of enhancing the collaboration and networking between birth and pregnancy cohorts in Europe with a focus on exposure-response relationships between environmental exposures and health outcomes in childhood. Within the ENRIECO-framework four case studies were undertaken to evaluate the concept of added scientific value through harmonisation and exchange of birth cohort data for common analyses on environmental health risks. Two alternative analytical approaches were evaluated: (1) The first was based on the central storage of data from different birth cohorts following which pooled and non-pooled analyses were centrally conducted by a single case study leader (in the following the centralised approach) and (2) the second was based on the meta-analysis of summary statistics derived by uniform methods conducted in each cohort (in the following the decentralised approach). In this paper we report and compare feasibility, benefits and limitations of these two approaches to combine analysis.

### **Centralised Approach in three Case Studies: Dampness/Mould or Tobacco Smoke Exposure and Respiratory Health and Allergy**

Each of the three case studies aimed to examine the association between certain environmental exposures and paediatric health outcomes.

The scientific background of the first case study was suggestive evidence for an association between mould or mould spores and respiratory tract infections, bronchitis, early wheeze and physician-diagnosed allergic rhinitis [5] and between indoor dampness/mould and cough, wheeze, upper respiratory tract symptoms or physician-diagnosed asthma [6]. The first ENRIECO case study analysed data from 31,000 children from 11 birth cohorts and found early exposure to visible mould and dampness to be associated with asthma symptoms in the first two years of life and allergic rhinitis in children up to 10 years of age [7]. Exemplarily, Figure 1 shows already published results of this combined data analysis.

see Illustration 1.

The second and third case study aimed to investigate the health impact of early tobacco smoke exposure on respiratory health in children. Previous research

indicated that prenatal maternal smoking resulted in a higher risk of physician-diagnosed asthma and wheeze in preschool- and schoolchildren [8-11] and found pre- and postnatal maternal smoking to be a risk factor for asthma in young adults [12]. The second and the third centralised ENRIECO case study found that pre- and postnatal maternal smoking in the vicinity of children had independent effects on the development of wheeze and asthma in children aged 0-2 and 4-6 years. This effect could also be observed for the smoking of other persons in the vicinity of children and pregnant women. Data from 32,000 children from 8 cohorts and 49,000 children from 19 cohorts was investigated, respectively [13,14].

### **Decentralised Approach in one Case Study: Persistent Organic Pollutants and Birth Weight**

Exposure to high concentrations of persistent organochlorines may cause fetal toxicity, but the evidence at low exposure levels is limited and inconsistent [15,16]. Therefore, large studies with substantial contrast of exposure within and between populations and with reliable indicators of exposure are warranted. Tissue concentrations of some xenobiotics as persistent organochlorines are highly correlated, restricting options for inference as to which are the actual compounds causing an observed effect. If effects are observed consistently across populations with different underlying covariance patterns it increases the validity of the results. The fourth ENRIECO case study indicated an inverse relationship between the concentration of polychlorinated biphenyls (PCBs) in biological specimens and birth weight. This result was based upon 7,990 women from 12 European birth cohorts [17]. As an example, Figure 2 shows already published results of this approach.

see Illustration 2.

## **Methods**

The decentralised and centralised approaches can be divided into two main parts. The first part comprises of the preparatory steps, which is similar for both approaches, and the second main part is database building and analyses.

### **Preparatory Steps for Both Approaches**

1. Definition of Key Responsibilities: For both approaches one case study leader was in charge of the coordination of case study workshops, communication with participating birth cohorts,

development of analyses and agreement plans and the coordination of the data harmonisation process. Interpretation of data was the joint responsibility of the case study leader, participating cohorts and for the centralised approach the different case study working groups. Further key responsibilities and organisational structures are displayed in Figure 3 and 4 for both approaches.

see Illustration 3.

see Illustration 4.

**II. Trust Building and Agreement:** Participating cohorts were considered to provide valuable, previously collected and longitudinal data, which made transparency of the data collection and analysis process essential.

European birth cohorts identified from former networking projects such as GA<sup>2</sup>LEN, HITEA (Health Effects of Indoor Pollutants: Integrating microbial, toxicological and epidemiological approaches), HIWATE (Health Impacts of Long-Term Exposure to Disinfection By-Products in Drinking Water) and ESCAPE (European Study of Cohorts for Air Pollution Effects) were contacted and informed about the study. Subsequently, they were invited to join a personal meeting for a general discussion of the project's content and possible analyses. This helped to promote trust building among collaborating parties, to substantiate the analyses strategy and to elaborate the analysis plan and a memorandum of understanding. These documents were circulated among the cohorts for final decision on participation.

### **Database-Building and Analyses for the Centralised Approach**

**I. Eligibility of Cohorts:** Interested cohorts had to provide the case study leader with the precise definition of their birth cohort variables, including answering categories and points of follow-up during which data was collected, e.g. for the outcome variable 'wheeze' during the first years of life, "Has your child had wheezing or whistling in the chest in the past 12 months? Yes/No". After receiving this information from all cohorts, the case study leader conducted a variable overview for each of the planned case studies. Based on this overview the case study working groups were able to decide which cohorts had eligible data (e.g. wheeze at 0-2 years) to participate in the planned analyses and which analyses were possible with the available data, respectively (see

table 1). Cohorts without the relevant outcome variable were excluded.

see Illustration 5.

None of the eligible cohorts refused participation out of other reasons. Therefore, it is unlikely that the data of excluded cohorts would have had any crucial impact on the findings.

**II. Collection of Individual Participant Data:** Individual participant data, chosen for analysis, was collected by the case study leader. Participating cohorts received the agreement form, which regulated data access, storage, management, analysis, and publication policy including authorship issues.

The initial data management and harmonisation of variables was done by the case study leader. Datasets with identical variable and value labels were prepared for the needs of different case study working groups. Original datasets and syntaxes used for variable modification were saved and made available to the cohorts and case study working groups for later replicability.

**III. Descriptive Analyses and Data Verification:** The case study leader prepared descriptive analyses of each modified dataset. The results were sent to the cohorts who had to verify the data. Discrepancies were to be reported to the case study leader and solved in discussion.

**IV. Final Analyses and Manuscript Preparation:** Each case study working group was responsible for final analyses. Interpretation of the data and manuscript preparation was a joint responsibility of the case study working groups and the case study leader.

The case study working groups were not allowed to perform any other analyses than the ones agreed upon by the cohorts, or to pass any data or results on to a third party. For future revisions and to aid transparency, it is mandatory to store the individual participant data for a certain period of time (e.g. 10 years) and to document each step of the analyses so that data can be reproduced in the future when needed.

### **Database-Building and Analyses for the Decentralised Approach**

**I. Eligibility of Cohorts:** Using the ENRIECO network, the case study leader invited all European birth

cohorts with measurements of two specified POPs (CB153 and/or p,p'-DDE) in a biological specimen (maternal blood, cord blood or breast milk) to contribute to the case study. Other eligibility criteria were data on analytical methods, birth weight, gestational age and information about several known extraneous determinants of preterm delivery and birth weight. Of 14 eligible cohorts in the ENRIECO network of European mother-child cohorts, 12 agreed to contribute. Each cohort appointed one to two representatives that joined the case study working group.

II. Development of Detailed Protocol for Data Analysis: The case study working group developed the design of the data analysis during two 2-days meetings half a year apart. This work included definition and scaling of outcome variables, exposure variables (including handling of values below the level of detection) and potential confounders and the definition of main steps in the data analysis. This included the principal analysis, possible supplementary- and sensitivity analyses and checks for violation of assumptions of the statistical analysis. The decisions on data analysis were specified in data and procedure statements in a SAS script that was translated into SPSS syntax. Those syntaxes were developed centrally. The case study working group also agreed on time schedules, authorships and publication issues.

III. Decentralised Data Analysis and Data Verification: Each birth cohort created a dataset according to the specified variable definitions and ran the SAS or SPSS program as appropriate. A number of unforeseen cohort specific problems were identified and resolved by contact among the case study leader and case study working group members. The output, including descriptive data as well as regression analysis, were collected at one centre and aggregated for discussion and interpretation at a subsequent working group meeting. This resulted in updates of the SAS/SPSS program and iterations of all data-analyses in each centre.

IV. Meta-analysis and Manuscript Preparation: Following verification of all updated output from each birth cohort centre, a meta-analysis of risk across cohorts was performed in one centre based upon summary statistics created from each cohort by the uniform program. Interpretation of the data and manuscript preparation was a joint responsibility of the case study leader, the case study working group and other representatives of the participating birth cohorts (1-3 additional persons from each cohort).

## Data Harmonisation

To harmonise different datasets, three aspects of variables characteristics determine the comparability of questionnaires of different cohorts:

- I. similar wording of variables,
- II. answering categories and
- III. the timing of the follow-ups when data was assessed. Differences in one or more of the three aspects and/or missing data can complicate or even make the harmonisation process impossible.

The variables of interest were collected to check comparability across cohorts. For this purpose a definition of each variable used in the combined analyses had to be agreed on, e.g. when the outcome of interest is early wheeze: "Has your child had wheezing or whistling in the chest in the past 12 months (Yes/No)?" asked at any time during the age period 0 to 2 years.

If cohorts had obtained data on the variable at several follow-ups, e.g. at 3, 12 and 24 months, the information was integrated in one variable e.g. "wheeze\_0-2years" to use as much information as possible per each cohort. Important is a consensus of how to combine several follow-ups with a special focus on the handling of missings. A conservative approach might be to handle missings as described in the illustration 6.

## Discussion

We found that a decentralized approach should be recommended for combined birth cohort data analysis if barriers for cross-border data transfer exist and/or a solid basis of trust still has to be established among newly collaborating partners. The centralized approach seemed more suitable for combined analyses addressing variables with very heterogeneous assessments across cohorts, where a flexible handling of data is essential and an established basis of trust among participating partners already exists.

### Strengths and Limitations

In the centralised approach, the central storage of data allowed a flexible handling of data and with a single data collection and harmonisation process it was possible to conduct not only one but three combined analyses with different foci. Moreover, we had the possibility to test different analytic approaches (i.e. pooling data, or conducting a two-stage meta-analysis) for each specific question. For the success, several

personal meetings with the working groups were necessary to discuss and agree on the final analysis plans for the centralised approach. Each cohort needed only few personnel resources, but as a consequence was only minimally involved in the harmonisation and analyses process. The decentralised approach has the advantage of enabling cohort leaders to build trust over time and creates an efficient framework for identifying and resolving issues related to variable definitions. The centralised method may become more cost-efficient considering that cohorts agree to address further research questions with the established centralised dataset.

Meta-analyses have the advantage of showing effect estimates for each individual cohort and allow e.g. slightly different confounder adjustments, while pooled analyses are eligible for more homogenous data where the effect of individual cohorts is not of main concern.

Analysis of pooled data including all information at the individual level has some advantages compared to meta-analysis irrespective whether this is undertaken by the centralised or the decentralised approach. First, pooled analysis based upon full individual records bypass the risk of ecologic bias inherent in all meta-analyses. In addition, analysis of pooled data allows modelling of exposure-response relationships in more detail than possible by meta-regression analysis. More power to identify of lowest adverse effect thresholds is in particular important in studies addressing potential low dose toxic exposures.

In the decentralised approach the cohorts were maximally involved in decisions on analyses and the need to write the detailed software script at an early phase ensured adherence to a prior hypothesis. At a later stage, the approach was less flexible. Reanalyses were time-consuming as they imply the involvement of the case study leader and all participating cohorts. Since only the summary estimates were used in the meta-analyses, further analysis such as the examination of the shape of the exposure-response curve, of effect-modification across cohorts or other analyses based upon a pooled dataset were limited. The decentralised approach does not necessitate transfer of complete datasets across borders and may thus bypass legal or ethical constraints linked to the use of the datasets.

see Illustration 7.

## Conclusions

In light of their strengths and weaknesses both the centralised and decentralised approach are recommendable for combined data analyses. In general, combined approaches are valuable to obtain sufficient sample sizes with enough power to detect association between rare events.

It is strongly recommended that resources for personnel and meetings are carefully planned. For both approaches, about four face-to-face meetings should be held to establish personal contact, discuss data harmonisation, analyses and interpretation of findings.

In a centralised approach with data collection and harmonisation of up to 19 cohort datasets, resources of a post doc would be needed for nine months and an additional two months should be planned per each analysis and manuscript preparation of the case study working groups.

In the decentralised approach each participating cohort would need two to three months to define the subset of data for the meta-analysis, to run programs, to format output and to participate in meetings. In addition to time for preparing and organising the study, a post doc for six months would be needed to analyse data and draft papers.

To increase the willingness of birth cohorts to participate in collaborative projects on combined data analyses, financial reimbursement for time and effort to provide previously collected datasets should be considered for both approaches.

The ENRIECO case studies showed that both the centralised and the decentralised approach of combined analyses with individual participant data from European birth cohorts were successful [7, 13, 14, 17]. Data collection and management, particularly the harmonisation process of often heterogeneous variables, are laborious and time-consuming efforts. Transparency through regular updates, presentation of results from interim analyses and the possibility for birth cohort researchers to comment and agree to each step of the analysis process builds trust and forms the basis for a sustainable collaboration with the aim to perform further combined birth cohort data analyses.

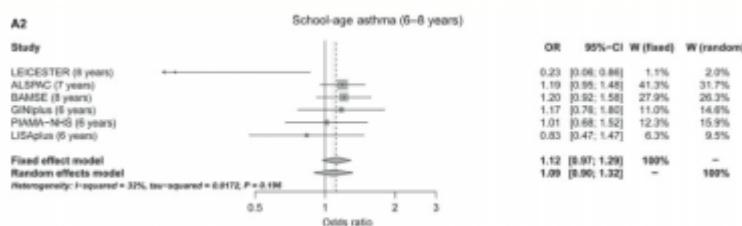
## References

1. Vrijheid M, Casas M, Bergstrom A, Carmichael A, Cordier S, Eggesbo M, Eller E, Fantini MP, Fernández MF, Fernández-Somonano A, Gehring U, Grazuleviciene R, Hohmann C, Karvonen AM, Keil T, Kogevinas M, Koppen G, Krämer U, Kuehni C, Magnus P, Majewska R, Andersen AM, Patelarou E, Petersen MS, Pierik FH, Polanska K, Porta D, Richiardi L, Santos AC, Slama R, Sram RJ, Thijs C, Tischer C, Toft G, Trnovec T, Vandentorren S, Vrijkotte TG, Wilhelm M, Wright J, Nieuwenhuijsen M: European Birth Cohorts for Environmental Health Research. *Environ Health Perspect* 2012; 120(1):29-37.
2. Frew AJ: GA2LEN-The Global Allergy and Asthma European Network. *Clin Exp Allergy* 2005;35(2):122-5.
3. Munoz-Lopez F: Ga2Len: coordinated research. *Allergol Immunopathol (Madr)* 2005;33(5):243-4.
4. Legler J, Hamers T, van Eck van der Sluijs-van de Bor, Schoeters G, van der Ven L, Eggesbo M, Koppe J, Feinberg M, Trnovec T: The OBELIX project: early life exposure to endocrine disruptors and obesity. *Am J Clin Nutr* 2011; 94(6):1933-1938.
5. Tischer C, Chen CM, Heinrich J: Association between Domestic Mould and Mould Components, and Asthma and Allergy in Children: A Systematic Review. *Eur Respir J*. 2011;38(4):812-24. doi: 10.1183/09031936.00184010 [online 3rd May 2011]
6. Fisk WJ, Lei-Gomez Q, Mendell MJ: Meta-analyses of the associations of respiratory health effects with dampness and mold in homes. *Indoor Air* 2007;17(4):284-96.
7. Tischer CG, Hohmann C, Thiering E, Herbarth O, Muller A, Henderson J, Granell R, Fantini MP, Luciano L, Bergström A, Kull I, Link E, von Berg A, Kuehni CE, Strippoli MP, Gehring U, Wijga A, Eller E, Bindslev-Jensen C, Keil T, Heinrich J, ENRIECO consortium: Meta-analysis of mould and dampness exposure on asthma and allergy in eight European birth cohorts: an ENRIECO initiative. *Allergy* 2011; 66(12):1570-9. doi: 10.1111/j.1398-9995.2011.02712.x.
8. Gilliland FD, Li YF, Peters JM: Effects of maternal smoking during pregnancy and environmental tobacco smoke on asthma and wheezing in children. *Am J Respir Crit Care Med* 2001;163(2):429-36.
9. Lannero E, Wickman M, Pershagen G, Nordvall L: Maternal smoking during pregnancy increases the risk of recurrent wheezing during the first years of life (BAMSE). *Respir Res* 2006;7:3; doi:10.1186/1465-9921-7-3.
10. Pattenden S, Antova T, Neuberger M, Nikiforov B, De Sario M, Grize L, Heinrich J, Hrubá F, Janssen N, Luttmann-Gibson H, Privalova L, Rudnai P, Splichalova A, Zlotkowska R, Fletcher T: Parental smoking and children's respiratory health: independent effects of prenatal and postnatal exposure. *Tob Control* 2006;15(4):294-301.
11. Neuman A, Hohmann C, Orsini N, Pershagen G, Eller E, Fomsgaard Kjaer H, Gehring U, Granell R, Henderson J, Heinrich J, Lau S, Nieuwenhuijsen M, Sunyer J, Tischer C, Torrent M, Wahn U, Wijga AH, Wickman M, Keil T, Bergström A, as part of the ENRIECO consortium: Maternal smoking in pregnancy and asthma in preschool children: a pooled analysis of 8 birth cohorts. *AJRCCM* 2012; doi:10.1164/rccm.201203-0501OC.
12. Longnecker MP, Korrick SA, Moysich KB: Human Health Effects of Polychlorinated Biphenyls. In *Dioxins and Health*. 2nd edition. Edited by Schecter A, Gasiewicz TA. John Wiley & Sons; 2003:679-728.
13. Wang C, Salam MT, Islam T, Wenten M, Gauderman WJ, Gilliland FD. Effects of in utero and childhood tobacco smoke exposure and beta2-adrenergic receptor genotype on childhood asthma and wheezing. *Pediatrics* 2008;122(1):e107-14.
14. Goksoy E, Amark M, Alm B, Gustafsson PM, Wennergren G. The impact of pre- and post-natal smoke exposure on future asthma and bronchial hyper-responsiveness. *Acta Paediatr* 2007;96(7):1030-5.
15. Vardavas CI, Hohmann C, Patelarou E, Martinez D, Keil T, Nieuwenhuijsen M, Kogevinas M. et al: The independent role of prenatal and postnatal active and passive smoking on early wheeze in 20 European birth cohorts. 2012 in press .
16. Toft G, Hagmar L, Giwercman A, Bonde JP: Epidemiological evidence on reproductive effects of persistent organochlorines in humans. *Reprod Toxicol* 2004;19(1):5-26.
17. Govarts E, Nieuwenhuijsen M, Schoeters G, Ballester F, Bloemen K, de Boer M, Chevrier C, Eggesbo M, Guxens M, Krämer U, Legler J, Martínez D, Palkovicova L, Patelarou E, Ranft U, Rautio A, Petersen MS, Slama R, Stigum H, Toft G, Trnovec T, Vandentorren S, Weihe P, Kuperus NW, Wilhelm M, Wittsiepe J, Bonde JP, OBELIX, ENRIECO: Prenatal Exposure to Polychlorinated Biphenyls (PCB) and Dichlorodiphenyldichloroethylene (DDE) and Birth Weight: A Meta-analysis within 12 European Birth Cohorts. *Environ Health Perspect* 2012;120(2):162-70.

## Illustrations

### Illustration 1

"Adjusted odds ratios and 95%confidence intervals (CI) of asthma in relation to early exposure to mould and/or dampness (0-2years), from random effect meta-analyses (combined effect) and separately by each cohort. For each study, the size of the box represents the variance, the horizontal linke the CI of each cohort. W (fixed) and W (random) indicate the percentage weight of each cohort contributing to the combined summary estimate" [7].



### Illustration 2

"Adjusted regression coefficients (95%CI) of cord serum PCB.153 with birth weight (g). Covariates included in the regression model: child's gestational age and sex, mother's region, maternal BMI, height, smoking status during pregnancy, socioeconomic status, mother's age, parity, ethnicity and time of sampling" [17].

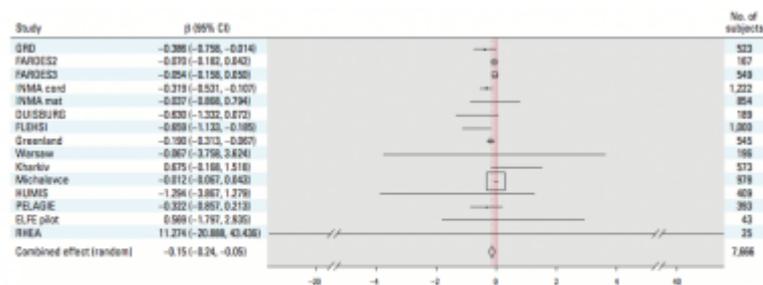
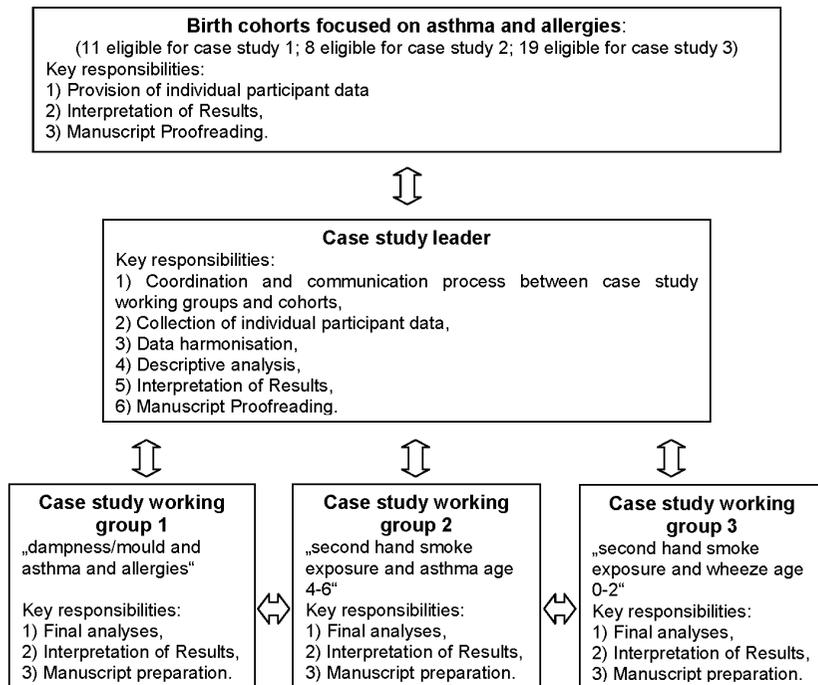


Figure 1. Adjusted regression coefficients (95% CI) of cord serum PCB-153 (ng/L) with birth weight (g). The squares are proportional to the inverse variance of the effect estimation of each cohort. Covariates included in the regression model: child's gestational age and sex, mother's region, maternal BMI, height, smoking status during pregnancy, socioeconomic status, mother's age, parity, and ethnicity, and time of sampling. Greenland, Warsaw, and Kharkiv are part of the INUENDO cohort.

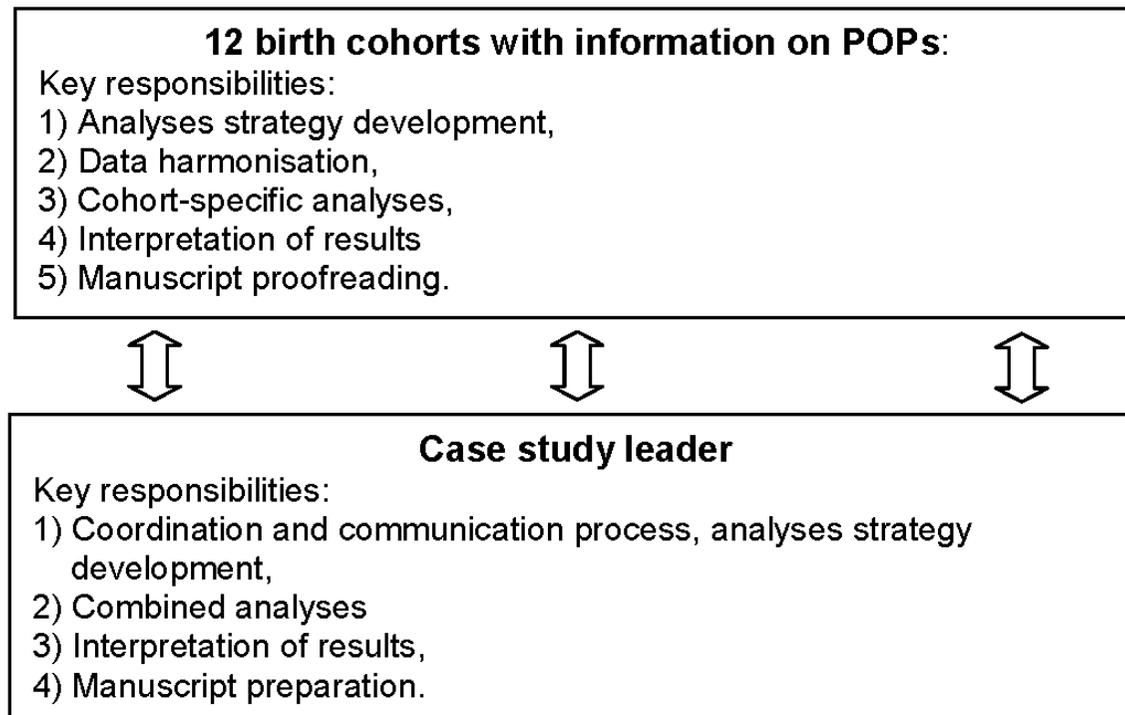
### Illustration 3

Key responsibilities and working flow of the centralised approach.



## Illustration 4

Key responsibilities and working flow of the decentralised approach.



## Illustration 5

Example for variable overview, identifying the availability of data: variable wording, answering categories and follow-ups.

<b>Cohort</b>	<b>Wording of Variable</b>	<b>Answering Categories</b>	<b>Asked at which follow-up</b>
AMICS-M enorca	Did your child ever wheeze in the last 12 months?	Never/ occasionally (1-6 episodes)/ most of the time (7+ episodes)	1y,2y
BAMSE	Has your child had problems involving wheezy breathing after x years of age?	yes/no	1y, 2y
CO.N.ER	In the last 12 months has your child had breath difficulty with wheezing symptoms?	yes/no	3y
DARC	Has or has had the child occasionally suffered from wheezing/restricted breathing?	yes/no	3m, 6m, 9m, 12m, 18m, 3y,
GINI	Did your child suffer from wheezing in the chest while breathing in the last 12 months?	yes/no	0,1y,2y,3y
LISA	Did your child have chest wheezing in the last 6/12months?	yes/no	6m, 1y,1y, 2y
MAS	Did your child suffer from wheezing in the last 3/6/12 months?	yes/no	3m,6m,1y,2y,3y

m = months, y = years.

## Illustration 6

Example for integrating information and handling of missings of several follow-up assessments for combined analyses.

variable answering categories	<u>original cohort</u> <u>variable:</u> wheeze at 3 months	<u>original cohort</u> <u>variable:</u> wheeze at 12 months	<u>original cohort</u> <u>variable:</u> wheeze at 24 months	<u>combined</u> <u>variable:</u> wheeze at 0 to 24 months
<b>two clearly defined cases:</b>				
a n s w e r i n g categories: yes/no	yes	yes	yes	yes
	no	no	no	no
<b><u>“yes” at any follow-up results in a “yes”</u></b> <b>for the harmonized variable, e.g.:</b>				
a n s w e r i n g categories: yes/no	no	yes	no	yes
	yes	no	missing	yes
<b><u>combinations of “no” and missings at the follow-ups</u></b> <b>result in missings for the harmonized variable, e.g.:</b>				
a n s w e r i n g categories: yes/no	no	no	missing	missing
	missing	no	missing	missing

## Illustration 7

Limitations and strengths of a centralised vs. decentralised approach of combined data analysis.

Centralised Approach	Decentralised Approach
<b>Strengths</b>	
Combined analyses avoid fragmentation of published results	
flexible handling of data enables different analyses with one data collection (different analytic strategies e.g. meta-regression and two-stage meta-analysis can be performed)	maximal involvement of cohorts in decisions on analyses
handling of heterogeneous datasets possible	no transfer of datasets across national borders (bypassing legal or ethical constraints linked to use of datasets)
<b>Limitations</b>	
only possible if strong basis of trust is established beforehand	limited to originally planned analyses-reanalyses highly time-consuming (involvement of case study leader + all participating cohorts)
cohorts only little involved in analyses process	pooled data analyses not possible
high resources of personnel for work package leader for central data collection and harmonisation	high resources of personnel in each cohort for data harmonisation and analyses
possibility of misinterpretation of variables (assessment and analyses do not remain in one institution)	
since it may not be possible to retrieve all data, it is important to have extended knowledge on the exposures and outcomes of interest when the data is centralised.	

## Disclaimer

This article has been downloaded from WebmedCentral. With our unique author driven post publication peer review, contents posted on this web portal do not undergo any prepublication peer or editorial review. It is completely the responsibility of the authors to ensure not only scientific and ethical standards of the manuscript but also its grammatical accuracy. Authors must ensure that they obtain all the necessary permissions before submitting any information that requires obtaining a consent or approval from a third party. Authors should also ensure not to submit any information which they do not have the copyright of or of which they have transferred the copyrights to a third party.

Contents on WebmedCentral are purely for biomedical researchers and scientists. They are not meant to cater to the needs of an individual patient. The web portal or any content(s) therein is neither designed to support, nor replace, the relationship that exists between a patient/site visitor and his/her physician. Your use of the WebmedCentral site and its contents is entirely at your own risk. We do not take any responsibility for any harm that you may suffer or inflict on a third person by following the contents of this website.