Author's response to reviews

Title: Validity of information on atopic disease and other illness in young children reported by parents in a prospective birth cohort study

Authors:

Nadja H Vissing (nadja.vissing@dbac.dk)
Signe M Jensen (smj@life.ku.dk)
Hans Bisgaard (bisgaard@copsac.com)

Version: 2 Date: 8 August 2012

Author's response to reviews: see over
Dear Editor of BMC Medical Research Methodology

I hereby submit a revised version of our manuscript entitled: "Validity of information on atopic disease and other illness in young children reported by parents in a prospective birth cohort study".

We appreciate the opportunity to re-submit this. We thank the Editor and reviewers for the accolades and many positive comments and constructive input. We acknowledge all the issues raised and appreciate seeing the manuscript significantly improved after such revision.

Please find a point-to-point response to the reviewers’ comments below. Our responses are added in red. At the end of the document you will find a complete version of the revised manuscript with all corrections from the initially submitted manuscript marked in red.

We hope you will find the revision satisfactory and consider the manuscript for publication in the BMC Medical Research Methodology.

Sincerely,

Nadja Hawwa Visissing MD
COPSAC; Copenhagen Prospective Studies on Asthma in Childhood
Faculty of Health Sciences, University of Copenhagen
The Danish Pediatric Asthma Center
Copenhagen University Hospital, Gentofte
Ledreborg Alle 34; 2820 Gentofte
Phone +4539978400
www.copsac.com
nadja.visissing@dbac.dk
Response to Reviewer No 1:

General remarks
This article demonstrates the sensitivity (completeness, underreporting) of atopic and other diagnoses until age three years in the COPSAC birth cohort study, a well known Danish cohort study with its own niche and strength among the many birth cohort studies on allergy and asthma. The results are valuable for other cohorts on the topics, for the interpretation of the results of the COPSAC study and similar cohort studies, and for the design of new cohort studies.

The article however does not present result on specificity (overreporting, misdiagnosing). The authors admit this in the Discussion, nevertheless it is not quite satisfactory since specificity is indeed an issue in epidemiological studies. The authors seem to assume that GP's diagnoses are suitable as a gold standard to determine sensitivity of COPSAC diagnoses, but not suitable to determine specificity. No clear arguments are given for this assumption (see also my comment on sensitivity and specificity of GP's diagnoses in the Discussion section, below; this is especially important for the diagnosis of asthma, as argued below).

We share the view of the reviewer that it would indeed be interesting to address the issue of specificity. Unfortunately this is not possible with the data we have. GP records were received as hard copies. The identification of missed diagnoses was found by reviewing the GP record and comparing it to COPSAC data. They were not typed in independently. The main purpose of this study was to ascertain that a comprehensive birth cohort study designed like COPSAC can rely on the assumption all relevant information on their participant’s health care interactions is gathered, i.e. a data collection with high sensitivity. This is now further clarified throughout the manuscript:

Abstract: “The COPSAC study exhibited full sensitivity to the main study objectives, atopic disease, and high sensitivity to respiratory, infectious and skin related illness.”

Methods (line 154): “When the records were returned they were reviewed by a trained senior medical student and compared with COPSAC database information. When an event was captured by the GP but not by COPSAC, it was considered as a missed event and was registered in a separate data sheet by ICD10 code and dated.”

Discussion (line 236): “It is a further limitation that because of the nature of our data collection we only deal with the issue of under-reporting, i.e. sensitivity, but not over-reporting, i.e. specificity. We only captured missed diagnoses, but we cannot know the degree of over-reporting of illness, which was previously reported as a problem in interview surveys[3, 25].”

Conclusion (line 286): “Clinical interviews of parents at six-monthly intervals shows complete sensitivity to the study objectives of child’s atopic disorders and very high sensitivity to other respiratory and skin disorders”.

The limitations of our choice of gold standard are elucidated.

The order of the sections (Methods section placed at the end of the article) has some disadvantages: in the Results and in the Discussion it is not always clear how events and atopic diagnoses were defined, because the necessary information about the design and execution of the COPSAC study is only given later in the Methods section. Therefore, please consider a conventional order (Methods before Results) if possible, otherwise give some more essential details in the Introduction or Results.

The manuscript is reordered into the conventional order.
See my comments on unclarities below.

Discretionary and Minor Essential Revisions

1. Introduction
   line 65: the next sentence is not clear to me, please reformulate: "Children that are diagnosed and treated by other physicians without our knowledge would potentially compromise the quality of our data." This sentence can easily be misunderstood when the reader does not know that children were seen by study doctors (this is only mentioned later in the Methods); e.g. where the COPSAC doctors also the treating physicians and the ones who informed the GP about the asthma diagnosis and treatment; or were other physicians (GPs or paediatricians?) the treating doctors and did the parents report on those doctor's diagnosis and treatment?
   Further clarification is added (line 63): “with a dedicated research clinic providing the families ready access to clinical evaluation and treatment with scheduled and acute clinical visits. Any atopic or respiratory illness is seen and treated by the research doctors.”

2. Results, line 82
   How are "medical events" defined? Please clarify if a repeated report of the one and the same continuing disease episode at two visits counted as one or two events? And the other way around, are repeated episodes during one period between interviews counted as separate events?
   This is clarified in the methods section (line 91-95): “The doctor assessed the reported illnesses, asking clarifying questions when needed, and classified them as ‘medical events’ by ICD10 codes. Medical events were entered online into the COPSAC database with a start and end date for each episode of illness. If a child was seen during a medical event, the end date was added at the following interview session.”

3. Point 2 is especially important for wheeze and asthma. Was asthma considered a life-time diagnosis (in other words is the no. of asthma events (48 in Table 1) reflecting the total number of children in the cohort with asthma ever), or is it reflecting distinct episodes of asthmatic symptoms (so that one child could contribute more than one asthma event)? Since no asthma diagnoses were missed by COPSAC this implies that repeated visits at the GPs (for instance for asthma medication follow-ups) were not counted as multiple events. This adds to the confusion and therefore I would suggest to explain clearly what is meant by asthma events and a diagnosis of asthma.
   The use of asthma as a life-time diagnosis is clarified in the methods section (line 134) “In the present study atopic diseases were viewed as life-time diagnoses in the follow-up period, i.e. the child was diagnosed at any time point within the first 3 years of life“, and in table 1.
   It is underlined that asthma medication follow up was carried out by COPSAC, not GP (line 130): “Parents were requested to contact the COPSAC clinic if their child developed any kind of airway or skin related symptoms. If a child developed asthma, eczema or allergic rhinitis, it attended the clinic every three months for clinical evaluation and additionally at acute disease exacerbation. All medical treatment was handled by the COPSAC clinic”.
For asthma there is another point: it is questionable whether a diagnosis of asthma can reliably be made before the age of about 6 years. My own experience is that this is even known by parents, since sometimes in our studies we encountered that parents reported in the margin of the ISAAC asthma core questions that "my doctor does not diagnose asthma before the age of 6". If this was common medical practice in Denmark too, the fact that no asthma diagnosis were missed by COPSAC may simply be due to the GPs not using this diagnostic label under age 3 at all. For that reason alone it would be informative if wheeze was added as a diagnostic category in the table. It is puzzling that wheeze (other than the repeated and persistent wheeze that was required for a diagnosis of asthma) is absent from the table; one would expect that many children would have been seen by the GP with wheeze episodes that do not comply with the stringent criteria of asthma as defined by COPSAC.

The asthma definition is clarified in the methods section and the explanation is reformulated (line 110-136): “Since the study population comprised children below 3 years of age the diagnosis of asthma was based largely on symptoms, according to international guidelines [9]. Diagnosis of asthma was based on a predefined algorithm[9, 10] focusing on persistent wheezy symptoms and subsequent response to treatment. Respiratory symptoms were recorded by the parents in daily diaries. The description of symptoms was supported by a book (written for parents, about early childhood wheeze) that was integrated with the diary cards. The COPSAC doctors reviewed the diary entries with the parents at the 6 monthly visits as well as during acute episodes of wheeze. A wheezy episode was defined on the diary card as 3 consecutive days of wheeze. Persistent wheeze was defined as five such episodes within 6 months or daily symptoms for 4 consecutive weeks leading to a 3-month course of inhaled corticosteroids (ICS). Acute severe asthmatic exacerbation also led to a 3-month course of ICS. Children responding to treatment and with a relapse when stopping treatment were diagnosed with asthma. Further treatment followed a strict algorithm previously described in details [11, 12]. “

In fact there were no wheezy episodes registered by the GP that the parents did not subsequently report at the interview at COPSAC, probably because of the high awareness about project participation among both parents and doctors.

5. For asthma it would be also be informative to know the range of severity (in the Methods criteria for a study diagnosis of asthma are given, but this seems to be a broad range): The whole section regarding asthma definition is reformulated (see above) that should clarify this to some extent, references are added that thoroughly describes our treatment algorithms (line 113+123).

6. The table gives the no. of events regardless how they are distributed over children; it would also be informative to have an analysis at the child level, at least for the atopic diseases: how many children had a diagnosis of atopic dermatitis or rhinitis ever in the first three years (this can easily be added to the table without much expanding it).

As clarified above the atopic diseases are displayed as having occurred at any time point during the follow up period. Therefore the number of events for atopic diseases ("n") equals the number of children.

7. Methods line 231:
"The symptom character was judged by the doctor to be typical of asthma .": it is not clear to me whether this was a requirement for the COPSAC diagnosis of asthma, or whether persistent wheeze was sufficient for the COPSAC diagnosis of asthma. The whole section regarding asthma definition is reformulated (see above).

8. Methods line 242:
"Allergic rhinitis was diagnosed in children with sensitization to inhaled allergens clearly related to the symptomatic periods.": This sentence is unclear to me: the word sensitization suggests that this was tested by skin prick or specific IgE tests against aeroallergens, the word "inhaled allergens" suggests a challenge test by allergen inhalation, and "symptomatic periods" seems to refer to seasonal symptoms but does not describe the symptoms themselves (rhinitis without fever? rhinoconjunctivitis?); or does it just mean that allergic rhinitis was diagnosed if there was a history of seasonal rhinoconjunctivitis or rhinoconjunctivitis associated with alleged indoor allergen exposure? Please clarify.
Clarified in the methods section (line 126-129): “Allergic rhinitis was diagnosed in children with seasonal symptoms and sensitization to relevant allergens measured by a specific IgE test[14]. Symptoms were defined as persistent troublesome sneezing or blocked or runny nose severely affecting the well-being of the child in periods without common cold, fever or flu[15].”

9. Methods
What measures were taken to ensure that the coding of GPs data was done independently from the COPSAC diagnosis? Who did the coding of GP data, was it done by one observer or two (and if so, independently, and was agreement checked?).
Clarified in the methods section (line 155): “When the records were returned they were reviewed by a trained senior medical student and compared with COPSAC database information. The same person reviewed all the records.”
Were they blinded for the COPSAC data? Since for some atopic diseases the classification is difficult and arbitrary such independence is important for the credibility of the results, especially for the finding that atopic disease where 100% complete. For instance, where the 13 cases of "dermatitis" reported by GPs and missed by COPSAC really different from "atopic dermatitis"?
Since it was a comparison of data, it was not possible to blind the observer.

10.
I miss information on the compliance with study visits. This would be valuable information, since the chance of missing a diagnosis is probably bigger when one or more visits were missed; given that no diagnoses were missed it is important for future investigators to know this to be able to replicate this high sensitivity.
Information is added in the results section (line 187). “227 children (87%) attended all the study visits, 25 (10%) missed one visit, 7 (3%) missed two visits and one child (0.3%) missed three visits. “

11. Discussion line 120:
"Inter-observer differences have revealed satisfactory agreements between GPs around 80%[10, 11] and GP records therefore seems like a very sensitive, but probably not very specific source of health information." This sentence seems to justify why the authors only looked to sensitivity of COPSAC diagnoses and not specificity (false positive diagnoses cannot be identified by GP's diagnoses if
GP's diagnoses are not specific. But since the cited studies [10 and 11] only report interobserver differences no straightforward conclusion can be drawn on sensitivity and specificity of GP’s diagnoses. Please nuance this.

Is done, the last part of the sentence is deleted (line 226): “Previous reports have indicated that morbidity recorded by GP is a reliable estimate of community morbidity[19, 20], and inter-observer differences have revealed satisfactory agreements between GPs around 80%[21, 22].”

I can accept that over-reporting of symptoms was not addressed (as admitted in line 164) but over-reporting of doctor's diagnosis (specificity) seems quite possible and it would be worthwhile to have this reported too.

The issue is addressed above (first point).

12. Table 1: a confidence interval for the completeness of Atopic Diseases would be informative (given that the completeness is 100%, it should be an exact one-sided lower confidence limit, e.g. a 95% or 97.5% limit).

Confidence intervals for the atopic diseases are added to the table.

13. Abstract: the year(s) of birth of the children would be informative.

The information is added: “COpenhagen study on Asthma in Childhood (COPSAC) is a clinical birth cohort study of 411 children born of asthmatic mothers from 1999 to 2001.”
Response to Reviewer No 2:

1. It is rather peculiar, that at the end of the Discussion paragraph, additional methods are addressed. These seem to be out of order. Additionally, in the Discussion paragraph several statements are made, that should be moved to Methods (e.g. the Danish GP system). The manuscript is reordered into the conventional order. The section regarding the GP system is moved to the methods section.

2. Ethical issues (paragraph ‘external standard’ in Methods section). It is not clear whether parents had provided informed consent to the procedure of the investigators asking the children’s general practitioner for a copy of their record. A statement confirming this is added in the methods section (line 152): “Written informed consent to retrieve information from other health care sources, including the GP, was obtained from both parents/legal guardians.”

3. Timing (paragraph ‘external standard’ in Methods section). ‘If the same event was registered in both files (GP record and COPSAC database) within four weeks, it was considered the same disorder’. This statement should be clarified, as COPSAC collected data only every six months. The routine of registration is clarified in the methods section,(line 93): “Medical events were entered online into the COPSAC database with a start and end date for each episode of illness. If a child was seen during a medical event, the end date was added at the following interview session” And (line 159): “Sometimes dates of the same type of event differed between COPSAC and GP records. In such cases a time span of four weeks between the registered start dates was allowed. “

4. Statistical analysis (paragraph and table 2). Not clear whether overal p-values in table 2 took ordinal nature of several variables into account (household income, mother’s education). To my knowledge GEE does not do this, but it would be preferable (like a chi-squared for trend). It is true that our model does not take a potential ordinal nature into account. The choice of the GEE model was made because of the data structure with repeated measures, where it is necessary to take the within-child variation into account. This is not possible with a chi-square trend test. The confidence limits do not suggest a strong linear effect.

5. Completeness of GP records First paragraph of Results section, line 78: 67 records turned out to be incomplete. How could they be sure that other records were complete? The sentence is reformulated (line 183): “Of the returned records 67 did not contain sufficient data for the entire study period”.

6. Limitation The above mentioned 67 incomplete records (67/327= 20%) were left out of the analysis. Consequently, results are based on complete cases. This may have resulted in an overoptimistic estimate of the primary outcome.
The issue is added as a limitation under Discussion (line 246): It is a limitation that we only received GP records with full follow up from 260 out of 350 children (74%). However, it can be argued that the individual GP’s willingness to provide COPSAC with sufficient data can be considered independent of the parent’s study compliance.
Validity of information on atopic disease and other illness in young children reported by parents in a prospective birth cohort study

Nadja Hawwa Vissing MD; Signe Marie Jensen MSc; Hans Bisgaard MD DMSci

Affiliation: Copenhagen Prospective Studies on Asthma in Childhood; The Danish Pediatric Asthma Center; Health Sciences, University of Copenhagen; Copenhagen University Hospital, Gentofte, Copenhagen, Denmark.

Correspondence
Nadja Hawwa Vissing
Copenhagen Prospective Studies on Asthma in Childhood
Danish Pediatric Asthma Center
Health Sciences, University of Copenhagen
Copenhagen University Hospital, Gentofte
Ledreborg Alle 34
2900 Hellerup, Copenhagen, Denmark
Phone +45 3977 7360
Fax +45 3977 7129
Email: nadja.vissing@dbac.dk
Abstract

Background: The longitudinal birth cohort study is the preferred design for studies of childhood health, particularly atopic disease. Still, prospective data collection depends on recollection of the medical history since the previous visit representing a potential recall-bias. We aimed to ascertain the quality of information on atopic disease and other health symptoms reported by parental interview in a closely monitored birth cohort study. Possible bias from symptom severity and socioeconomics were sought.

Methods: COOpenhagen study on Asthma in Childhood (COPSAC) is a clinical birth cohort study of 411 children born of asthmatic mothers from 1999 to 2001. Child health is monitored at six-monthly visits with particular emphasis on atopic symptoms and infections. Data from the first three study years on 260 children was compared with records from their family practitioner as an external reference.

Results: A total of 6134 medical events were reported at the COPSAC interviews. Additional 586 medical events were recorded by family practitioners but not reported at the interview. There were no missed events related to asthma, eczema or allergy. Respiratory, infectious and skin related symptoms showed completeness above 90%, other diseases showed lower completeness around 77%. There was no meaningful influence from concurrent asthma or socioeconomics.

Conclusions: The COPSAC study exhibited full sensitivity completeness to the main study objectives, atopic disease, and high sensitivity completeness to respiratory, infectious and skin related illness. Our findings support the validity of parental interviews in longitudinal cohort studies investigating atopic disease and illness in childhood.

Key words: Validation studies, Cohort Studies, Child, Asthma, Atopic dermatitis, Diagnosis, Interviews, Recall bias, Infections.
Background

The longitudinal birth cohort study is the preferred design for studies of the origins of chronic diseases in childhood such as asthma[1, 2], partly because recall bias is minimized. The prospective clinical cohort study with doctor’s interview and examination at regular visits to the clinic is expected to have the highest completeness capturing medical and exposure history. Recall bias is reduced with a high frequency of visits to the clinic, but information still depends on recollection of the medical history since the previous visit. Particularly short-term symptoms such as common childhood infections may be influenced by recall bias[3, 4]. Diary data is considered a valid source and provides strong support to the history taking at the clinic. But such data collection also depends on a high level of compliance and the risk of data being influenced by socioeconomics of the subjects[4, 5].

The validity of information obtained in prospective clinical cohort studies have been sparsely investigated, in part because this requires a comparison with an external standard, which is rarely available[6, 7].

The Copenhagen Prospective Study on Asthma in Childhood (COPSAC) is an ongoing clinical, prospective, longitudinal birth-cohort study of 411 infants born to mothers with asthma designed to intensively investigate the development of atopic diseases in high-risk children, with a dedicated research clinic providing the families ready access to clinical evaluation and treatment with scheduled and acute clinical visits. Any atopic or respiratory illness is seen and treated by the research doctors.

Precise case definition of atopic disease is the hallmark of the COPSAC cohort study. Nevertheless, as dedicated as the research unit can be, it still requires a high level of compliance from study participants. Children that are diagnosed and treated by other physicians without our knowledge would potentially compromise the quality of our data.
We aimed to ascertain the completeness of our data on atopic diseases during the first three years of life in an ongoing prospective clinical birth cohort study. Secondly we aimed to investigate potential underreporting of the history of symptoms, diagnoses and other health care resource utilization reported by interview at six-monthly visits at our clinical research unit. Information collected by the family’s General Practitioner was used as external reference for such missing data. As a part of this analysis we tried to look at bias in parent-reporting related to the child’s asthma status and socioeconomics of the family.

Methods

Study cohort:
COPSAC is a longitudinal birth cohort study designed to examine the relation between the genetics, environmental and lifestyle factors and development of asthma, eczema and allergic symptoms in early life. The cohort study design and characteristics have previously been described in detail [8].

Subjects participating in COPSAC were recruited between August 1998 and December 2001 among pregnant mothers with a history of asthma diagnosed by a doctor and requiring medication. 411 children of asthmatic mothers were included in a comprehensive program of clinical and objective assessments. At three years of age 350 children (85%) were still active in the cohort. Children attended the COPSAC clinic for planned visits at one month of age and every six months. Additionally they were seen for any acute airway and/or skin symptoms. **If a child developed asthma it attended the clinic every three months for clinical evaluation.**
At each six-monthly visit, the medical doctors at COPSAC interviewed the parents about any illnesses, symptoms and use of medication during the previous six months. The doctor assessed the reported illnesses, asking clarifying questions when needed, and classified them as ‘medical events’ by according to ICD10 codes. Medical events Diagnoses were entered online into the COPSAC database with a start and end date for each episode of illness. If a child was seen during a medical event, the end date was added at the following interview session.

All data were collected according to Good Clinical Practice data management and quality control procedures including external monitoring. The study is conducted in accordance with the Declaration of Helsinki and approved by the Copenhagen Ethics Committee (KF01-289/96) and the Danish Data Protection Agency (2008-41-1754). Written informed consent was obtained from both parents/legal guardians.

Disease categories

In the present analysis we classified all diagnoses into the following five groups: 1. Atopic diseases (asthma, atopic dermatitis and allergic rhinitis), 2. Airway related illness (upper and lower airway infections), 3. Skin related illness (all diagnoses related to skin other than atopic dermatitis), 4. Childhood infections (infections with fever, without airway symptoms) and 5. Other illness. For group details see TABLE 1.

Atopic disease

Asthma:

Since the study population comprised children below 3 years of age the diagnosis of asthma was based largely on symptoms, according to international guidelines [9]. Diagnosis of asthma was based on a predefined algorithm[9, 10] focusing on persistent wheezy symptoms and subsequent response to treatment. Respiratory symptoms were recorded by the parents in daily diaries. The description of symptoms was supported by a book (written for parents, about early childhood wheeze) that was
integrated with the diary cards. The COPSAC doctors reviewed the diary entries with the parents at the 6 monthly visits as well as during acute episodes of wheeze. A wheezy episode was defined on the diary card as 3 consecutive days of wheeze. Persistent wheeze was defined as five such episodes within 6 months or daily symptoms for 4 consecutive weeks leading to a 3-month course of inhaled corticosteroids (ICS). Acute severe asthmatic exacerbation also led to a 3-month course of ICS. Children responding to treatment and with a relapse when stopping treatment were diagnosed with asthma. Further treatment followed a strict algorithm previously described in details [11, 12].

Asthma is a primary outcome in the COPSAC cohort. It was diagnosed according to international guidelines as previously detailed[23] based on persistent wheezing, defined as five episodes (each lasting at least three consecutive days) within six months or acute severe asthma symptoms (resulting in hospitalization or the need for systemic corticosteroid treatment, as judged by the COPSAC doctors). The symptom character was judged by the doctor to be typical of asthma with symptoms between episodes, such as exercise induced symptoms; prolonged nocturnal cough; persistent cough outside common cold; symptoms causing wakening at night; in need of intermittent rescue use of inhaled \( \beta_2 \)-agonist; and responding to a 3-month course of inhaled corticosteroids and relapsing when stopping treatment.

Atopic dermatitis was defined by the criteria of Hanifin and Rajka based on the presence of 3 of 4 major criteria and at least 3 of 23 minor signs as previously detailed [13]. Allergic rhinitis was diagnosed in children with seasonal symptoms and sensitization to relevant allergens measured by a specific IgE test[14]. Symptoms were defined as persistent troublesome sneezing or blocked or runny nose severely affecting the well-being of the child in periods without common cold, fever or flu[15]. Rhinitis was defined by persistent troublesome sneezing or blocked or runny nose severely affecting the well-being of the child in periods without common cold or flu[25]. Allergic rhinitis was diagnosed...
in children with sensitization to inhaled allergens clearly related to the symptomatic periods. Non-
allergic rhinitis was diagnosed in children without sensitization or without symptoms during periods of
exposure to such allergens[26].

Parents were requested to contact the COPSAC clinic if their child developed any kind of airway or
skin related symptoms. If a child developed asthma, eczema or allergic rhinitis, it attended the clinic
every three months for clinical evaluation and additionally at acute disease exacerbation. All medical
treatment was handled by the COPSAC clinic.

In the present study atopic diseases were viewed as life-time diagnoses in the follow-up period, i.e. the
child was diagnosed at any time point within the first 3 years of life.

*External standard*

*(Section moved from discussion):* The GP occupies a central position in the Danish health care service
and the GP records are considered a reliable external data source. Citizens are signed up with a GP of
his own choice and are allowed to change only once a year. The GP is the patients’ primary contact
with the health service and act as "gate keeper" to secondary care specialists. The GP keeps a record on
every patient. A personal identification number is assigned to every Danish citizen linking all health
care utilizations to the patient. If a patient seeks other medical health care, such as outpatient clinic,
emergency room and hospital, a discharge summary is sent to the GP. Consequently no patients are
treated for an illness without automatic notification of the GP, and any illness severe enough to cause
medical attention, even by phone, is registered by the GP. Prescription rules in Denmark are strict and
only very few drugs can be bought over the counter without doctor’s prescription, which excludes
patients from use of any antibiotics or anti-asthmatics without prescription automatically recorded by
the GP.
The family practitioner (GP) of every COPSAC child was identified in the national health registry. Written informed consent to retrieve information from other health care sources, including the GP, was obtained from both parents/legal guardians. The GP of children who completed three study years was requested by mail to send a copy of his record from the child’s first three years of life. When the records were returned they were reviewed by a trained senior medical student and compared with COPSAC database information. The same person reviewed all the records. When an event was captured by the GP but not by COPSAC, it was considered as a missed event and was registered in a separate data sheet by ICD10 code and dated. Sometimes dates of the same type of event differed between COPSAC and GP records. In such cases a time span of four weeks between the registered start dates was allowed. If the same event was registered in both files within four weeks, it was considered the same disorder.

Socioeconomics.

As a proxy for socioeconomic status information on household income, educational level and occupational status of the mother was obtained at the one-year-visit. Household income was classified into three groups: Income below average (<400.000 DKR), around average (400.000 – 600.000 DKR) and above average (>600.000 DKR)[8]. The highest level of completed education was divided into four categories: elementary school, college, medium and university education. Occupational status was described in four groups; professionals, non-professionals, unemployed and student; based on DISCO classification, a national statistic used by the Danish National Statistic Agency (www.dst.dk).
Statistical analysis

The total number of medical events for a child (N) was considered as the sum of events registered by COPSAC (n) plus the additional missed events found in the GP record. The completeness of COPSAC information was estimated as the number of medical events recorded by COPSAC (n) as a percentage of the total number of medical events (N).

Completeness was estimated using a GEE model with the logit link function taking into account the child variation. P-values corresponded to score tests. Confidence intervals of sensitivities were estimated on a logit scale, back-transformed and presented in brackets. Analyses were done using PROC GENMOD in SAS 9.1

Results

The first three years of longitudinal data collection was completed by 350 children. 327 (93%) GPs responded to the request of a copy of the child’s record. Of the returned records 67 did not contain sufficient data for the entire study period: were incomplete for the three year period. In 65 cases the families had moved from one city to another within the first 3 years of life, and the new GP did not return information from the former. In two cases the records were handwritten and unreadable. The final population for this study was therefore 260 children. 227 children (87%) attended all the study visits, 25 (10%) missed one visit, 7 (3%) missed two visits and one child (0.3%) missed three visits.

For these 260 children a total of 6134 medical events were recorded in the COPSAC database. In 586 cases a medical event was registered by the GP but not by COPSAC.
TABLE 1 displays the completeness of the groups. The overall completeness including all the medical events was 0.91 [0.90;0.92]). Completeness for the atopic diseases alone was 100%. The completeness differed significantly (p < 0.0001) between the four remaining disease groups. 

TABLE 2 displays the relationship between completeness and child’s asthma and socioeconomic status. The overall completeness was not associated with either child’s asthma (p=0.27), mother’s education (p=0.19) or mother’s occupation (p=0.28). The association with household income was borderline significant (p=0.06) with the completeness highest for the high income group (0.93 [0.91;0.94]) and lowest for the low income group (0.89 [0.86;0.91]) compared to the average income group (0.91 [0.90;0.93]). We also looked at the relationship between completeness within the five disease groups and asthma and socioeconomic status and found no significant associations. 

Discussion

Main findings

Doctor’s interviews at six-monthly clinic visits on child’s illness and symptoms supported by daily diary cards on lung symptoms provided 100% completeness to the main objectives of the study, the atopic diseases. Furthermore the current study revealed completeness above 90% to other respiratory and skin related illness and childhood infections. Other illnesses were only captured with a lower completeness around 77%. We saw an insignificant trend of better completeness in families from higher social status, but no influence from child’s asthma status.
Strengths and limitations (Strength and limitations rearranged into one section)

(Section moved): The strength of our data also relates to the close longitudinal surveillance at the
COPSAC clinical research unit. The study is a single-centre study with six-monthly assessments by
experienced study-doctors examining and taking clinical history based on standard operating
procedures supported by diary cards on atopy-related symptoms. This assures consistency in
procedures, definitions of conditions and data capture methods and reduce risk of misclassification.
Atopic disorders such as recurrent wheeze, asthma and atopic dermatitis in young children display
significant between-observer variation [16–18]. The risk of misclassification is important particularly in
young children with respiratory and skin disorders because there is gross inconsistency among doctors
in their diagnostic and treatment practice, reflecting little consensus on definition and best practices.

It is The a strength of our study is that we have provided an external reference to validate symptom
history, but this choice can be debated. (The descriptions of the GP moved to the methods section): The
GP occupies a central position in the Danish health care service and the GP records are considered a
reliable external data source. Citizens are signed up with a GP of his own choice and are allowed to
change only once a year. The GP is the patients’ primary contact with the health service and act as
"gate keeper" to secondary-care specialists. The GP keeps a record on every patient. A personal
identification number is assigned to every Danish citizen linking all health care utilizations to the
patient. If a patient seeks other medical health care, such as outpatient clinic, emergency room and
hospital, a discharge summary is sent to the GP. Consequently no patients are treated for an illness
without automatic notification of the GP, and any illness severe enough to cause medical attention,
even by phone, is registered by the GP. Previous reports have indicated that morbidity recorded by GP
is a reliable estimate of community morbidity [19, 20], and inter-observer differences have revealed
satisfactory agreements between GPs around 80%[21, 22] and GP records therefore seems like a very sensitive, but probably not very specific source of health information. Unfortunately GP records are only kept by the GP personally. So far there is no central registration of GP diagnose coding or link to a national database. The gate keeper function of the GP and the automatic notification of all health care contacts make the GP record a very reliable source of information on health care interactions.

(moved to methods:) Prescription rules in Denmark are strict and only very few drugs can be bought over the counter without physician’s prescription. This excludes patients from use of any antibiotics or anti-asthmatics without prescription automatically recorded by the GP.

The fact It is a strong point of our study that health care is free to all Danish citizens minimizes socioeconomic status as a significant confounder of contacts to the health care system.

Nevertheless, the study is limited by the fact that the GP record can only give information on illness where parents considered medical attention necessary. It is an important limitation to our study that we only look at doctor diagnosed illnesses. By nature we could not capture illness where parents did not consider medical attention necessary. In fact, it is known that a large proportion of childhood illness resolves spontaneously and is not reported to the professional health care[23, 24].

It is an important further limitation that because of the nature of our data collection the study only addresses we only deal with the issue of under-reporting, i.e. sensitivity, but not over-reporting, i.e. specificity. We only captured missed diagnoses, but we cannot know the degree of over-reporting of illness symptoms, which was previously reported as a problem in interview surveys[3, 25].
It is a limitation that our results relate only to interviews done six-monthly and do not permit us to draw conclusions on interview studies with shorter or longer intervals. We allowed a time span of 4 weeks between family practitioner and COPSAC registration of events. This choice is debatable since several events in either source could occur within this period and potentially lead to an underestimation of unreported events.

It is a limitation that we only received GP records with full follow up from 260 out of 350 children (74%). However, it can be argued that the individual GP’s willingness to provide COPSAC with sufficient data can be considered independent of the parent’s study compliance.

Socioeconomic status was determined at one year of age, but household income, educational and occupational status are dynamic and can change over the course of a study. However this study is carried out in a relatively short period of time, so we consider status at one year of age as a reasonable estimate for the entire study period. Another limitation in measuring socioeconomic status is that only education and occupation of the mother, but not the father was included. We used univariate analysis; this choice can also be questioned, since education and income are likely to be correlated and interactive effects can be missed by this procedure, although consensus on this matter is not clear[26].

The generalizability of our conclusions to a general population can be questioned because of the high-risk nature of the population. Subgroup analysis showed that the completeness of symptom history was independent of the child’s asthma symptoms, but the fact that all mothers had a history of asthma is likely to increase their awareness of symptoms associated with lung symptoms.
Interpretation

This study primarily addresses the validity of COPSAC data. It also addresses the reliability of information on child health collected by parental interviews. The results can be beneficial to researchers working with cohort studies investigating pediatric public health, particularly atopic diseases. Data on illness are not perfect, but acceptable. The majority of studies on general illness in early childhood in Denmark are either retrospective[27], cross-sectional[28] or follow children for a shorter period of time[23, 29], and little is known about the validity of such data. We would expect them to be less accurate without the close follow-up that the clinical cohort study provides, but further studies on this subject would be of great value.

Parents of children with many symptoms may be suspected to be more compliant with the project than the parents of children without symptoms, which could introduce confounding to further analysis. However, subgroup analysis showed sensitivity was independent on the child’s asthma.

Data on non-infectious events are not appropriate for further analysis due to the low completeness. Comparison of medical events with low sensitivities to those with higher suggests that understanding of the concept of illness influences reporting (TABLE 2). When asked about illness in childhood the parents preferentially report infectious diseases or diseases with similar symptoms, even though the interviews specifically invite any kind of illness. Participating in a study focused on asthma, eczema and allergy, parents may consider common events such as conjunctivitis and thrush as everyday complaints not severe enough to be considered as illness, and non-infectious events such as fractures, burns and wounds are not considered relevant for reporting.
We saw an insignificant trend of better completeness in families from higher social status. Although insignificant in this study, it may be prudent to consider this as a potential confounder.

Conclusion

Clinical interviews of parents at six-monthly intervals show complete sensitivity full completeness to the study objectives of child’s atopic disorders and very high sensitivity completeness to other respiratory and skin disorders with no important influence from socioeconomic status or concurrent asthma. Our findings support the completeness of prospective doctor’s interviews at the clinic supported by diaries when assessing childhood health and symptoms related to atopic disease.

List of abbreviations

COPSAC: Copenhagen Prospective Studies on Asthma in Childhood
GEE: generalized estimating equation
GP: General Practitioner
ICD10: International Classification of Diseases Version 10, WHO
ICS: Inhaled Corticosteroids
SAS: Statistical Analysis System

Competing interests

The authors declare that they have no competing interests.
Authors’ contributions

NHV is responsible for the acquisition of data, obtaining and reviewing records from GPs, data analysis and interpretation, important intellectual input and writing of the manuscript. SMJ performed the statistical analyses and contributed with important intellectual input. HB is responsible for the integrity of the study as a whole, from conception and design to acquisition of data, analysis and interpretation of data and writing of the manuscript. All authors red and approved the final manuscript.

Acknowledgements

The authors wish to thank the children and parents participating in the COPSAC cohort as well as the COPSAC study team.