ERJ Express. **Published on December 23, 2009 as doi: 10.1183/09031936.00154409**

Different definitions in childhood asthma: how dependable is the dependent variable?

Karina E van Wonderen[1], Lonneke B van der Mark[1], Jacob Mohrs[1], Patrick JE Bindels[2], Wim MC van Aalderen[3], Gerben ter Riet[1]

1: Department of General Practice, Academic Medical Center, Amsterdam, the Netherlands

2: Department of General Practice, Erasmus Medical Center, Rotterdam, the Netherlands

3: Pediatric Respiratory Medicine, Emma Children's Hospital – Academic Medical Center, Amsterdam, the Netherlands

Corresponding author

K.E. van Wonderen, MSc

Academic Medical Center-University of Amsterdam

Division of Clinical Methods & Public Health

Department of General Practice

P.O. Box 22700

1100 DD Amsterdam

The Netherlands

Phone: +31-205667956

Fax: +31-205669186

E-mail: k.vanwonderen@amc.uva.nl

Key words

asthma, cohort studies, illness definition, prediction models, review

Word count

Abstract 200 words

Main text 3294 words

30 references

3 tables

1 figure

Positions

Karina E van Wonderen	KvW	Researcher
Lonneke B van der Mark	LvdM	Researcher
Jacob Mohrs	JM	Scientific programmer
Patrick JE Bindels	PB	Professor
Wim MC van Aalderen	WvA	Professor
Gerben ter Riet	GtR	Associate Professor

ABSTRACT

Background

There is abundant literature on how to select and statistically deal with predictors in prediction models. Less attention has been paid to the choice of the outcome. We assessed the impact of different asthma definitions on prevalence estimates and on prediction model's performances.

Methods

We searched Pubmed and extracted data of definitions used to diagnose childhood asthma – between 6 and 18 years – in cohort studies. Next, using data from an ongoing cohort study (n=186), we constructed and compared four prediction models which all predict asthma at age six, using a fixed set of predictors and four different definitions in turn. We defined an area of clinical indecision (posterior probability between 25% and 60%) and calculated the number of children who remained inside this area.

Results

122 papers yielded 60 different definitions. Prevalence estimates varied between 15.1% and 51.1% depending on the asthma definition used. The percentage of children whose posterior asthma probability was in the area of clinical indecision varied from 14.9% to 65.3%.

Conclusions

Variation in definitions and its effect on the performance of prediction models may be another source of otherwise inexplicable variation in daily clinical decision making. More uniformity of operational asthma definitions seems needed.

INTRODUCTION

Asthma is the leading chronic disease among children. Many definitions of asthma have been proposed in guidelines and used in follow-up studies and clinical trials [1-3]. As with many other diseases, asthma prevalence estimates vary widely across time and regions. Even in one region at a single point in time, asthma prevalence estimates may differ by the use of different populations, study designs, and illness definitions. Usually, these sources are difficult to disentangle. The same applies to asthma prevalence estimates that are conditional on (one or more) risk predictors such as an atopic constitution or exposure to tobacco smoke. Such conditional prevalence estimates are usually obtained through (multivariable) prediction models [4-7]. Prevalence estimates can be useful for healthcare resource planning purposes [8,9], while prediction models are developed mostly for clinical applications [4-7].

The definition of asthma in young children is complex and varies across authoritative sources [1-3]. Even if the conceptual definition of asthma is unequivocal, the operational definitions used in empirical studies may well differ. It is unclear to which extent asthma prevalence estimates are determined by the particular operational illness definition. Therefore, we set out to provide an overview of recently used definitions to diagnose asthma in children between 6 and 18 years in published literature of research in which asthma was an endpoint. We then assessed the impact of four exemplary asthma definitions on prevalence estimates at age six. Finally, we determined the impact of definition choice on a prediction model's performance – which all predict asthma at age six – by constructing four logistic regression models with a fixed set of three known predictors of asthma using four different, but commonly used definitions in turn.

METHODS

Definitions and operationalisations of asthma

In a MEDLINE search, using PubMed, we searched for studies published between 1998 and 2008 using the MeSH-terms 'asthma', 'children', and 'cohort studies'. Studies that fulfilled the following criteria were included: (1) cohort design; (2) asthma as primary or secondary outcome; (3) participants between 0 and 18 years; (4) asthma diagnosed between 6 and 18 years; (5) ≥ 100 children included, and (6) English as language of publication.

Papers were selected by one author (KvW) based on titles and abstracts. If title and abstract were unclear, the full text papers were screened using the same criteria. A second author (LvdM) checked a randomly selected 10% of the papers that KvW had excluded for inadvertent exclusions.

From all included articles one author (KvW) extracted the following information: (1) definition of asthma; (2) operationalisation of the definition, that is, (a) source of the information used to diagnose asthma (i.e. parents or medical records, etc), and (b) the instrument used to diagnose asthma (i.e. questionnaire or list of diagnostic codes, etc). A second author (LvdM) checked the extracted information of a randomly selected 10% of all included papers.

Prevalence estimates and prediction models' performances using different definitions

To assess the variation in prevalence and prediction model performance we used data from the ARCADE study, an ongoing prospective cohort study [10]. One of the aims of ARCADE is to construct a primary care based asthma prediction model for preschool children at risk of developing asthma. Briefly, between 2004 and 2006, one to five year old children at risk for developing asthma were selected from general practices in the Netherlands. 'At risk' in this study was defined as 'visited the general practitioner with recurrent coughing (≥ 2 visits), wheezing (≥ 1) or shortness of breath (≥ 1) in the 12 months previous to enrolment'. All children are being followed up to the age of six. At age six, a

definitive diagnosis of asthma is made according to the operational definition used in the ARCADE study (see below).

For this contribution, we used data of 186 children, between 2 and 4-years old at enrolment, whose follow-up was completed and diagnosis of asthma was made at age 6. Five year old children were excluded because asthma definitions cover time periods of at least 12 months back. This precludes *pre*diction in a strict sense.

ARCADE was approved by the Central Committee on Research Involving Human Subjects (CCMO/P04.0098C). Written informed consent was obtained from the parents prior to all measurements.

Development of different prediction models

We constructed four logistic regression models using a fixed set of three binary known predictors of asthma [11-14]. Predictors were (1) wheezing (during the previous year, but apart from colds), (2) eczema (during the previous year), and (3) specific immunoglobulin E (IgE) directed against house dust mite, cat, or dog dander [12,15]. Information on the predictors was collected at time of enrolment (in the ARCADE study) at age 2 and 4- years old [10].

The first three prediction models were constructed using three operational definitions taken from the literature search (see Results for details, table 2). The definitions were selected based on the following criteria: (1) definition could be constructed using the ARCADE data and (2) definitions differed by at least one key clinical component to prevent comparing definitions that are almost similar (see table 1). A fourth prediction model was added using the operational definition used in the ARCADE study; that is, a combination of current symptoms (complaints of wheezing and/or shortness of breath and/or recurrent coughing) and/or use of β_2 agonists and/or inhaled corticosteroids both for any length of time during the previous 12 months in combination with airway hyperresponsiveness to methacholine. Hyperresponsiveness is defined as a provocation concentration of methacholine inducing a 20% fall in FEV_1 (PC_{20}) $\leq$ 8.0 mg/ml [12,16,17].

Thus, four asthma prediction models were constructed with a fixed set of three known predictors determined at enrolment (children aged between 2 and 4- years old) using four different definitions which all predict asthma at the age of six.

Statistical methods

By multiple imputation 44 missing IgE values were estimated using several baseline variables collected in ARCADE like breastfeeding, history of asthma of the parents, and whether the child awoke as a consequence of shortness of breath [18,19]. Five imputed datasets were created and 5×4 regression analyses were run, one for each dataset-definition combination. All further analyses used the mean of the five datasets per definition. Conservatively, per definition, 95% confidence intervals (CI) were determined taking the lowest lower bound and highest upper bound of all imputed datasets.

First, the prevalences for the four definitions were compared. Next, the posterior probabilities for the four prediction models were summarized using the 10th, 50th, and 90th centiles of their distributions. To illustrate the potential clinical consequences of these differences between the posterior probability distributions, two decision thresholds were selected. The first threshold we set at 25%, assuming that below that threshold a clinician may well choose a ‘wait and see’ policy as the chance that the child has asthma at age six is relatively small. The second threshold was set at 60% assuming that a clinician may pursue a more active management strategy, perhaps including a prescription of anti-inflammatory drugs. Thus, an area of clinical indecision was defined. To be able to focus on a single outcome, the performances of the prediction models were compared, using the proportion of patients who remained in the area of clinical indecision, that is, whose posterior asthma probabilities were between 25% and 60%.

Finally, the areas under the receiver operating curves (AUC) between the models as a commonly used measure of overall predictive performance were compared. All differences and their 95% CI were calculated using bootstrapping procedures (1000 times). All calculations were performed using Stata version 10 (Stata corp, College Station TX, USA).

RESULTS

Literature search for definitions and operationalisations

The overall search yielded 1238 papers, of which 122 were included. There were no discordances between the two authors with respect to inclusion or extracting information on definitions and operationalisations.

In total, the 122 included papers yielded 60 different definitions (table 1). The most common definitions were: (1) a doctor's diagnosis of asthma ever (10%), (2) a doctor's diagnosis of asthma (time unspecified) (8%), (3) asthma ever (6%), (4) a doctor's diagnosis of asthma ever in combination with asthma symptoms in the previous 12 months (5%), (5) a doctor's diagnosis of asthma ever in combination with symptoms of asthma in the previous 12 months or the use of asthma medication (5%). In total, 34% of the papers used one of these definitions.

The 60 definitions may be categorised in various groups. Sixty-two papers (51%) used a definition which was based on a doctor's diagnosis of asthma with or without other symptoms, medication use or any time constraint. Bronchial hyperresponsiveness or spirometry was a component of the definition in 13 (11%) of the papers. Definitions based on symptoms alone, were also seen in ten (8%) of the papers. Thirty five papers (28%) used a definition which was a combination of symptoms, (doctor's) diagnosis of asthma and asthma medication use. Two papers (2%) did not mention any definition.

The three most prevalent operationalisations were: a questionnaire filled in by the parents and or child (58%), interview with the parents and/or child (20%), and a clinical examination by a health professional (7%). In 2% of the definitions it was unclear which operationalisation was used.

Prevalence estimates and prediction models' performances using different definitions

Table 2 shows the four operational definitions which were used to estimate prevalences and predictive performances of prediction models (see table 3).

For the definition “doctor’s diagnosis of asthma ever” (Definition 1 “Dr-ever”) it did not seem logical to construct a prediction model since this definition covers the whole period back to birth which defies the purpose of prediction. Therefore, we did not determine the prediction model performance for the definition *Dr-ever*.

Prevalence estimates

Table 3 (2nd column) shows that prevalence estimates using different definitions ranged from 15.1% (definition 2 “Dr-ever&whe”) to 51.1% (definition 4 “BHR&sym/med”). The prevalence estimate for definition 2 (Dr-ever&whe) is smaller than that according to definition 1 (Dr-ever) since the former requires wheezing and is, therefore, more stringent.

Although a methacholine challenge test was a component of two definitions (definition 3 “BHR&whe” and definition 4 “BHR&sym/med”), prevalence estimates between them varied greatly, difference of -25.3% (95% CI from -31.7 to -19.4).

Figure 1 shows that 15% (28/186) and 46% (86/186) of the children were defined as having and not having asthma by all definitions, respectively (overall agreement 61%). This figure shows also that almost all children (95/100) who were defined as having asthma by definition 1 (Dr-ever), definition 2 (Dr-ever&whe), or definition 3 (BHR&whe) had asthma according to definition 4 (BHR&sym/med).

Table 3 (2nd column) shows that the prevalence estimates for definition 1 (Dr-ever) and definition 3 (BHR&whe) were similar, 47/186 (25.3%) and 48/186 (25.8%) respectively. However, figure 1 also shows that definitions 1 and 3 nevertheless disagree in 39/186 (21%) of children.

Posterior probability distribution

Table 3 (3rd column) also shows the posterior probability distributions of the three prediction models (as mentioned before definition 1: Dr-ever was omitted from this analysis).

Definitions 2 (Dr-ever&whe) and 3 (BHR&whe) showed a similar posterior probability distribution. The posterior probabilities for definition 4 (BHR&sym/med) differed greatly from

the other two definitions. In particular, the 90th centile of definition 2 (Dr-ever&whe; 37.4%) is similar the 50th centile of definition 4 (BHR&sym/med; 40.7%).

Predictive performances of prediction models (thresholds)

Table 3 (4th column) shows the predictive performance of the models using the proportion of children who remained in the area of clinical indecision. The percentage of children in this area varied from 14.9% (definition 2 “Dr-ever&whe”) to 65.3% (definition 4 “BHR&sym/med”).

Areas under the ROC Curve ($AUCs_{ROC}$)

The AUC_{ROC} may be interpreted as the probability that from two randomly drawn children, one with asthma and one without, the one with asthma is assigned a higher probability [20]. The $AUCs_{ROC}$ varied from 0.67 for definition 4 (BHR&sym/med) to 0.76 for definition 2 (Dr-ever&whe) with their differences varying from 4 to 9%.

DISCUSSION

Main findings

In 122 papers, we found 60 different operational definitions. Applied in a single cohort, we found that prevalence estimates and posterior probabilities varied substantially with the operational definition used. Similarly, the proportion of children that remained in an area of clinical indecision varied greatly with the definition chosen. Although the $AUCs_{ROC}$ between the models were fairly similar, the predictive performances of the models clearly were not.

Strength and limitations

A strength of this study is that the comprehensive search of the literature for all published cohort studies on asthma in 6-18 year old children in the previous 10 years is unlikely to have missed many operational definitions. In addition, the use of a single cohort, thus fixing time, region, study population and study design allowed us to isolate the effect of definitions. We see the following limitations. First, 10% of the papers excluded by the first author were checked by a second author. Although, there were no discordances between first and second author, we cannot exclude that on a total of 1126 papers, up to 30 discordances between the authors might have occurred. However, it is unlikely that these papers contained different asthma definitions and would affect our findings and message to any important degree. Second, different prediction models were compared using a fixed set of three plausible predictors of asthma. Prediction models using other predictors might show different results. Third, we were unable to construct prediction models based on clinical examination or medical records since the ARCADE data do not contain such data. Fourth, 44 observations for specific IgE were missing. In general, test results that cannot be obtained reliably in clinical practice should not be used in prediction models. However, we believe that, the missings were due to the research situation. In a number of children IgE measurements were not obtained because parents did not make an appointment with the GP's surgery. Second, specific IgE outcome were missing due to GP assistants taking insufficient blood for

* that is the upper 95% percent confidence limit of an exact confidence interval around the proportion of zero found in a sample of size 112

analysis. Although GP assistants had received instruction to perform the measurements, occasionally they failed to collect enough blood. We believe that these problems will seldom occur in practice where most likely, parents will follow their doctor's advice and visit a dedicated laboratory. And finally, to illustrate potential clinical consequences we introduced two thresholds. These thresholds were pre-selected by us and not determined by formal cost-effectiveness analysis or cost-utility analysis. In reality, physicians may use different thresholds in their decision-making although we think that they will lie within the proximity of those we selected.

Relation to other studies

Literature search

Sixty two (51%) of the papers were (partly) based on the definition 'a doctor's diagnosis of asthma ever'. This definition is based on ISAAC questionnaire's core questions 'Has your child ever had asthma?' combined with 'Has your child's asthma been confirmed by a doctor?' Standardized questionnaires are easy to use and allow prevalence comparisons and trends worldwide, but they can also be subjective and highly dependent on the interpretation and judgement of the person responding to the questionnaire [21].

Only 7% of the definitions used objective criteria for 'signs and symptoms' as documented by a physician. Since prevalence estimates based on symptoms of wheezing determined by questionnaire differed from definitions based on clinical examination this point deserves our attention [21,22]. Eleven percent of definitions were also based on more objective criteria, such as spirometry or severity of bronchial hyperresponsiveness. Measuring bronchial hyperresponsiveness is less influenced by variation in symptom perception. Our literature search made clear that various symptoms are being used in combination with bronchial hyperresponsiveness. Wheezing appeared to be the symptom most often used.

Prevalence estimates

Our findings that prevalence estimates are affected by the choice of definition is confirmed in other studies, although these focused on crude prevalence estimates, not on prediction models. Overall, with exception of Greenlee, lower prevalence estimates were found for definitions based on medical records versus parental reports based on the ISAAC questionnaire [23-25]. Lower prevalence estimates found in medical records than with parental report may be due to incompleteness of records, physicians mentioning the term asthma to parents without believing strongly enough in the diagnosis to document it, errors in parental memory, or combinations of these factors. Even parental forgetfulness in combination with more severe incompleteness of records may be compatible with these figures.

The prevalence estimates we report may strike as high. Prevalences varied from 15.1% (definition 2 “Dr-ever&whe”) to 51.1% (definition 4 “BHR&sym/med”). The dataset of the ARCADE study, however, purposefully consists of data of children in whom a general practitioner is likely to consider a diagnosis of asthma. Furthermore, the majority (158/186) of the children in the analysis were enrolled at age 3 and 4 years old, that is, mostly past the stage of transient wheezing, and therefore a higher prevalence could be expected [26,27]. This also makes comparison with studies such as the one by Wördemann et al. difficult [22].

The definition used in the ARCADE study (definition 4 “BHR&sym/med”) yielded a much higher prevalence than definition 3 (BHR&whe) (51.1% versus 25.8%). Although both definitions were based on bronchial hyperresponsiveness 25.3% of the bronchially hyperresponsive children had no symptoms of wheezing during the previous 12 months but experienced symptoms of coughing and/or shortness of breath and/or had recently used asthma medication.

The similar prevalence estimates (25.8% with definition 1 “Dr-ever”, and 25.3% with definition 3 “BHR&whe”) did not label the same children with asthma highlighting that different definitions result in similar prevalence estimates but does not imply that the same children are labelled as having asthma.

Prediction model's performance

As far as we know, this is the first time that the variation in the performance of a prediction model associated with the use of different definitions was studied. Although Miller selected predictors of asthma by logistic regression analysis using different definitions, she did not compare different prediction models using different definitions [25].

We chose as the outcome of main interest the proportion of patients remaining in the area of clinical indecision. This method is related to reclassification methods that are currently gaining ground [28] and are, in our opinion, more informative than the area under the ROC curve (AUC) method to express how well a prediction model tells the ill from the healthy. In addition, this reclassification method more directly relates to clinical decision making.

Illness definition issues are not restricted to asthma, nor to prediction models [29] and we believe that many areas of medicine may benefit from scrutinizing illness definitions and the variation of operationalisations in research and practice [30].

Conclusion

Although much has been written on how we should select and statistically deal with predictors used in prediction models, the role of the dependent variables in such models seems to have received less emphasis. We have shown that measurement choices underlying the construction of the outcome or dependent variable may have large impact on estimates of prevalence as well as on predictive probabilities as provided by a prediction model. This variation in posterior probabilities is likely to have its impact on clinical management with both over- and undertreatment as a consequence. Nevertheless, achieving agreement on illness operational illness definitions will remain a challenge.

Acknowledgements

We would like to thank the general practitioners from HAG-net-AMC, Zorggroep Almere and Prinsenhof, Leemhuis and Buitenhuis for their time and effort in this study. We would further like to thank all children with their parents for participating in the study. We also would like to thank Pauline van Steenwijk, MSc, for her invaluable help and Machteld IJff, MSc, Alice Karsten, and Albertien Buijs for their assistance.

This study was financially supported by the Netherlands Asthma Foundation (3.4.02.20 and 3.4.06.078) and by the Stichting Astma Bestrijding.

Details of contributors

GtR conceived the idea. All authors designed the study, and drafted and critically revised the manuscript. KvW, LvdM, and JM extracted the data. GtR, KvW, LvdM, and JM analysed and interpreted the data. GtR and KvW did the statistical analysis.

REFERENCES

1 British Thoracic Society SIGN. British guideline on the management of asthma 2005

2 The Global Initiative for Asthma 2008. GINA Report, Global Strategy for Asthma Management and Prevention 2008.

3 National Asthma Education and Prevention Program. Expert Panel Report 3: Clinical Practice Guidelines. Guidelines for the Diagnosis and Management of Asthma 2007. Report No.: http://www.nhlbi.nih.gov.

4 Moons KGM, Royston P, Vergouwe Y, *et al.* Prognosis and prognostic research: what, why, and how? *BMJ* 2009;338:b375.

5 Royston P, Moons KGM, Altman DG, *et al.* Prognosis and prognostic research: Developing a prognostic model. *BMJ* 2009;338:b604.

6 Altman DG, Vergouwe Y, Royston P, *et al.* Prognosis and prognostic research: validating a prognostic model. *BMJ* 2009;338:b605.

7 Moons KGM, Altman DG, Vergouwe Y, *et al.* Prognosis and prognostic research: application and impact of prognostic models in clinical practice. *BMJ* 2009;338:b606.

8 World Health Organization. Prevalence of asthma and allergies in children. *http://www euro who int accessed* 2007;accessed 19 January 2009.

9 Beasley R. Worldwide variation in prevalence of symptoms of asthma, allergic rhinoconjunctivitis, and atopic eczema: ISAAC. *The Lancet* 1998;351:1225-32.

10 van Wonderen K, Mark L, Mohrs J, *et al.* Prediction and treatment of asthma in preschool children at risk: study design and baseline data of a prospective cohort study in general practice (ARCADE). *BMC Pulmonary Medicine* 2009;9:13.

11 Jackson DJ, Gangnon RE, Evans MD, *et al.* Wheezing Rhinovirus Illnesses in Early Life Predict Asthma Development in High-Risk Children. *American Journal of Respiratory and Critical Care Medicine* 2008;178:667-72.

12 Eysink PE, ter Riet G, Aalberse RC, *et al.* Accuracy of specific IgE in the prediction of asthma: development of a scoring formula for general practice. *Br J Gen Pract* 2005;55:125-31.

13 Celedón JC, Wright RJ, Litonjua AA, *et al.* Day care attendance in early life, maternal history of asthma, and asthma at the age of 6 years. *Am J Respir Crit Care Med* 2003;167:1239-43.

14 Castro-Rodriguez JA, Holberg CJ, Wright AL, *et al.* A clinical index to define risk of asthma in young children with recurrent wheezing. *Am J Respir Crit Care Med* 2000;162:1403-6.

15 Stapel SO, Eysink PED, de Vrieze J, *et al.* IgE testing in capillary blood. *Pediatric Allergy and Immunology* 2004;15:230-3.

16 Cockcroft DW, Killian DN, Mellon JJ, *et al.* Bronchial reactivity to inhaled histamine: a method and clinical survey. *Clin Allergy* 1977;7:235-43.

17 Cockcroft DW. Bronchoprovocation methods - Direct challenges. *Clinical Reviews in Allergy & Immunology* 2003;24:19-26.

18 van Buuren S. Multiple imputation of discrete and continuous data by fully conditional specification. *Statistical Methods in Medical Research* 2007;16:219-42.

19 Sterne JAC, White IR, Carlin JB, *et al.* Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ* 2009;338:b2393.

20 Hanley JA. Receiver Operating Characteristic (Roc) Methodology - the State of the Art. *Critical Reviews in Diagnostic Imaging* 1989;29:307-35.

21 Martinez FD. Trends in asthma prevalence, admission rates, and asthma deaths. *Respiratory Care* 2008;53:561-5.

22 Wordemann M, Polman K, Diaz RJ, *et al.* The challenge of diagnosing atopic diseases: outcomes in Cuban children depend on definition and methodology. *Allergy* 2006;61:1125-31.

23 Greenlee RT, Chyou PH, Kieke A, *et al.* Prevalence of asthma in a general population cohort of farm children: comparison of estimates based on parental report and medical record review. *J Agromedicine* 2008;13:225-31.

24 Hederos CA, Hasselgren M, Hedlin G, *et al.* Comparison of clinically diagnosed asthma with parental assessment of children's asthma in a questionnaire. *Pediatric Allergy and Immunology* 2007;18:135-41.

25 Miller JE. Predictors of asthma in young children: Does reporting source affect our conclusions? *Am J Epidemiol* 2001;154:245-50.

26 Henderson J, Granell R, Heron J, *et al.* Associations of wheezing phenotypes in the first 6 years of life with atopy, lung function and airway responsiveness in mid-childhood. *Thorax* 2008;63:974-80.

27 Martinez FD, Wright AL, Taussig LM, *et al.* Asthma and wheezing in the first six years of life. The Group Health Medical Associates. *N Engl J Med* 1995;332:133-8.

28 Pencina MJ, D'Agostino RB, D'Agostino RB, *et al.* Evaluating the added predictive ability of a new marker: From area under the ROC curve to reclassification and beyond. *Statistics in Medicine* 2008;27:157-72.

29 Ninomiya T, Donnan G, Anderson N, *et al.* Effects of the End Point Adjudication Process on the Results of the Perindopril Protection Against Recurrent Stroke Study (PROGRESS). *Stroke* 2009;40:2111-5.

30 Miettinen OS, Flegel KM. Elementary concepts of medicine: III. Illness: somatic anomaly with.. *Journal of Evaluation in Clinical Practice* 2003;9:315-7.

Table 1: Operational definitions of asthma in 122 cohort studies published in English between 1998 and 2008

Asthma definition	N	Refs
Bronchial hyperresponsiveness AND current wheeze	1	12
Bronchial hyperresponsiveness AND (current wheeze or current nocturnal cough)	1	23
Bronchial hyperresponsiveness AND history of asthma and at least one of the following: reported dyspnoea, chest tightness, or wheezing in the previous 12 months	1	17
Bronchial hyperresponsiveness AND ≥1 episode of wheezing in the previous 12 months	3	34*, 64, 87
Bronchial hyperresponsiveness AND (current complaints or complaints during the previous 12 months AND/OR use of asthma medication (use of β2 agonists or inhaled corticosteroids currently or in the last 12 months))	1	40
Bronchial hyperresponsiveness AND doctor diagnosed asthma AND (at least two respiratory symptoms (cough, wheezing, dyspnoea, and nocturnal cough, wheezing, dyspnoea) or a history of recurrent asthma attacks)	1	83
Bronchial hyperresponsiveness AND doctor's diagnosis of asthma ever and asthma symptoms in the previous 12 months	1	48
Bronchial hyperresponsiveness AND current wheeze (wheezing, whistling, nocturnal cough or exercise induced wheeze in the last 12 months)	1	70
At least a 15% increase in peak flow after inhaled salbutamol AND ((episodic wheezing or dyspnoea) or successive 3-day nocturnal cough))	1	82
Based on lung function results, history of the disease, and physical examination	3	110, 45, 69
Doctor's diagnosis of asthma ever	12	51, 120, 86, 103, 77, 99. 98, 104, 32, 88, 95, 27
Doctor's diagnosis of asthma at this moment	4	61, 71, 38, 5
Doctor's diagnosis of asthma (unspecified when)	9	4, 28, 108, 111, 119, 59, 37,15, 94
Doctor diagnosed asthma (in a certain period)	2	67, 68
Doctor diagnosed asthma ever AND current wheeze (in the last three years)	1	74
Doctor's diagnosis of asthma ever and asthma symptoms in the previous 12 months	5	13, 33, 53, 85, 106*
Doctor's diagnosis of asthma ever AND ≥1 episode of wheezing in the previous 12 months	4	6, 50, 41, 89

Doctor's diagnosis of asthma (unspecified when) and ≥1 episode of wheezing in the previous 12 months	3	14, 52, 76
Doctor diagnosed asthma AND ≥ 1 exacerbation in previous 12 months	2	114, 107
Doctor diagnosed asthma AND (asthma symptoms and/or wheezing ever)	1	101
Doctor diagnosed asthma AND (at least one episode of asthma during the previous year or more than three episodes of wheezing during the previous year)	1	105
Doctor diagnosed asthma AND wheeze in the last year AND sleep disturbance due to wheeze in the last year AND objective atopy	1	100
Doctor diagnosis of asthma ever AND having asthma in a given time period (16-18 years) AND experiencing any asthma symptoms during the period	1	39
Doctor diagnosed asthma AND wheeze (unspecified when) AND current use of asthma medication	1	66
Doctor diagnosis of asthma AND asthma symptoms (wheeze and/or nocturnal cough in the absence of an obvious respiratory infection) in the past 12 months AND use of asthma medication in the past 12 months	1	62
Doctor diagnosis of asthma ever AND (symptoms of asthma or use of asthma medication in the previous 12 months)	6	1, 81, 56, 75, 84, 96*
Doctor diagnosed asthma (ever) AND more than one illness in the previous 12 months or one illness and 1. ever interrupted sleep or 2. any medication in the last 12 months or 3. overnight hospital stay in the last 12 months	1	118
Physician diagnosed asthma, observed wheezing, and/or prescription of asthma medication during the time period when the child was between 6 and 8 years of age (from medical records). Parental reports of occurrence of asthma collaborated the medical record data	1	91
Physician diagnosed asthma AND wheeze or asthma symptoms reported on >=2 questionnaires	1	97
Doctor diagnosed asthma AND (wheeze in the past 12 months or use of asthma medication in the past 12 months)	2	2, 25*
((Answers "Yes" to "Has your child had asthma?" AND to "Was asthma diagnosed or treated by a physician?") AND (asthma symptoms ('usual cough', 'chest wheezy or whistling' or 'attacks of wheezing with shortness of breath') during the past year or have used asthma medication during the 3 previous months).	1	60
(Two or more episodes of wheezing accompanied by dyspnoea that had ever been given the diagnosis of asthma by a physician) AND (the occurrence of asthmatic attacks or the need for any medication for asthma during the past two years).	1	72

Asthma at this moment	1	22
Asthma ever	7	8, 20, 49, 30, 54, 46, 92*
Asthma in the previous 12 months	4	43, 90, 11, 29*
Asthma in the previous 24 months	1	112

Does your child have long-term illnesses?' Asthma was one of the main concerns listed. OR children that had ever been hospitalized due to asthma	2	113, 102
Asthma ever AND asthma in the last 6 months	2	10, 35
Asthma (unspecified when) AND symptoms in the previous 12 months	1	79
Asthma ever AND wheezing or whistling in the chest in the past 12 months	1	115
Admitted to hospital with primary diagnosis of asthma	1	55
Diagnosis of asthma AND (wheezy symptoms (i.e. child's chest sounding wheezy or whistling occasionally apart from colds, most days or nights or during the past year) or if asthma medication was used in the past 3 months)	1	47

Symptoms of wheezing at this moment and symptoms of wheezing in the previous 12 months OR doctors diagnoses asthma	1	18
At least three episodes of bronchial obstruction verified by a physician	1	109
Wheezing in the previous 12 months in absence of an upper respiratory infection	1	93
Symptoms of wheezing in the previous 12 months	4	58, 24, 73, 42
More than 6 attacks of wheezing in the previous 12 months	1	65
Use of bronchodilators within the previous year for attacks of wheezing	1	26
Confirmative answers to the following questions: question 1 and one or more of questions 3, 4, 5, 6, and 7 and one or both questions 8 and 9; question 1 and at least question 11 and confirmed by the medical record of the child; question 1 and at least question 8,9,10 and confirmed by the medical record of the child; question 11 and confirmed by the medical record of the child. 1. Has your child ever had wheezing or whistling in the chest? No/Yes; 2. Has your child had wheezing or whistling in the chest at any time during the last 12 months No/Yes; 3. How many episodes with wheezing has your child had during the last 12 months? none/1-3/4-12/>12; 4. during the last 12 months, how often, on average, has your child been disturbed by wheezing? never/less than 1 night per week/1 or more nights a week; 5. during the last 12 months, had the wheezing of your child ever been so severe that he or she only could say 1 to 2 words between the breathings? no/yes; 6. during the last 12 months, has your child had wheezing in the chest during or after exercise? no/yes; 7. during the last 12 months, has your child had dry cough in the nights without having a cold or an infection? no/yes; 8. has your child ever had wheezing at any time after 2 years of age? no/yes; 9. has your child ever had 3 diagnosed episodes of bronchitis before 2 years of age? no/yes; 10. had your child ever had treatment with inhaled lomudal (cromolyn sodium) or inhaled steroids? no/yes; 11. has your child ever had a diagnosis of bronchial according to a physician? no/yes	1	63
Treatment for acute wheezing at a healthcare centre or in hospital in previous 12 months	1	31*
Three or more recurring attacks of bronchial obstruction causing wheezing, coughing or heavy breathing due to external factors such as animal dander, pollen, house dust or food	1	44
Recurrent wheezing episodes in the child's outpatient medical record. At least one episode of wheezing illness must have occurred in the absence of a respiratory infection	1	121

Three episodes of bronchial obstruction in the previous 12 months verified by a physician	1	78
Attacks of shortness of breath with wheezing within the preceding 12 months in addition to positive responses to question 2, 3, 4 and/or 5. Question (2): Does your breathing ever sound wheezy or whistling? Question (3): Do you ever have attacks of shortness of breath with wheezing? Question (4): Do you experience wheezing, chest tightness, cough, or breathlessness with any of the following: at rest, with exertion, with emotional stress, with exposure to cold air, or with chest infections, or head cold? Question (5): Do you experience wheezing after exposure to: dust, fumes, moulds, pollen, food, pets or drugs?	1	57*
(Doctor diagnosed asthma (ever) AND symptoms of asthma in previous 12 months) OR Three separate episodes of persistent wheezing (>=3 days duration) in the previous 12 months	1	16
Doctor diagnosed asthma at least once or 'asthmatic, spastic or obstructive bronchitis' more than once	1	19
≥ 2 doctor visits for asthma in the previous year OR two prescriptions for any asthma drug (β-agonists, inhaled corticosteroids, cromones of leukotriene receptor antagonists) in the previous year OR one hospitalisation in the previous year	1	9
Cough at night, ever wheeze, wheeze in the last year, asthma ever diagnosed or asthma treated in the previous 4 weeks	1	116
At least two of the following criteria: 1. dyspnoea ever, chest tightness ever and/or wheezing ever 2. doctor's diagnosis of asthma 3. use of medication (β2-agonist, sodium cromoglycate, corticosteroids, leukotriene antagonists and/or aminophylline) ever	1	21
Code 493/j45-46	1	117
No definition of asthma given	2	3, 7

* more than one outcome of asthma was given. The most stringent definition is cited in this table. For papers 36, 80 and 122 it was not possible to define the most stringent one. Therefore, these definitions are not cited in the table.

Table 2: Four operational definitions of asthma used to estimate prevalences and prediction models' performances

Asthma definition		Operationalisation	
	Name used in text	**Instrument**	**Source**
Doctor's diagnosis of asthma ever	*Dr-ever*	Questionnaire sent to the child's home	Parents
Doctor's diagnosis of asthma ever and ≥1 episode of wheezing in the previous 12 months	*Dr-ever & whe*	Questionnaire sent to the child's home	Parents
Bronchial hyperresponsiveness and ≥1 episode of wheezing in the previous 12 months	*BHR & whe*	Methacholine challenge test (PC_{20} $FEV_1 \leq 8.0$ mg/ml)[#] and questionnaire	Hospital (lung function laboratory) and information about wheezing by parents
Bronchial hyperresponsiveness and ((wheezing and/or shortness of breath and/or recurrent coughing) and/or (use of β_2 agonists and/or inhaled corticosteroids)) during the previous 12 months	*BHR & sym/med*[¶]	Methacholine challenge test (PC_{20} $FEV_1 \leq 8.0$ mg/ml)[#] and questionnaire	Hospital (lung function laboratory) and information about symptoms and medication by parents

[#] PC_{20} is concentration of methacholine that induced a 20% fall in FEV_1 (forced expiratory volume in one second)
[¶] Definition used in the ARCADE study

Table 3: Prevalence estimates and performance of prediction models given a fixed set of three predictors [~] **and four different asthma definitions (n=186)**

	Outcome (asthma at age 6)									
	Asthma prevalence (%)		**Posterior probability distribution (%)**			**Percentage of children with probability between 25% and 60%**		**AUC_{ROC}**	
Prediction model based on:			**10***	**50***	**90***				
Definition 1: Dr-ever[#]	25.3		n/a	n/a	n/a	n/a		n/a	
Definition 2: Dr-ever&whe[^]	15.1		4.6	8.2	37.4	14.9		0.76	
Definition 3: BHR&whe[§]	25.8		11.8	14.0	41.9	36.2		0.72	
Definition 4: BHR&sym/med[†¶]	51.1		30.2	40.7	70.4	65.3		0.67	
Contrasts[&]	**Δ**	**95% CI**	**Δ[$]**		**95% CI**	**Δ**	**95% CI**	**Δ**	**95% CI**
Dr-ever vs Dr-ever&whe	10.2	6.5 to 15.1	n/a			n/a	n/a	n/a	n/a
Dr-ever vs BHR&whe	-0.5	-6.9 to 5.9	n/a			n/a	n/a	n/a	n/a
Dr-ever vs BHR&sym/med	-25.8	-32.8 to -18.8	n/a			n/a	n/a	n/a	n/a
Dr-ever&whe vs BHR&whe	-10.8	-15.1 to -6.5	p< 0.0001	[$]		-21.3[‡]	-75.8[‡] to 21.5[‡]	0.05	-0.06 to 0.16
Dr-ever&whe vs BHR&sym/med	-36.0	-43.0 to -29.0	p< 0.0001	[$]		-50.3	-94.6 to -0.5	0.09	-0.04 to 0.21
BHR&whe vs BHR&sym/med	-25.3	-31.7 to -19.4	p< 0.0001	[$]		-29.0	-75.3 to 3.2	0.04	-0.06 to 0.14

[#] Doctor's diagnosis of asthma ever

[^] Doctor's diagnosis of asthma ever and ≥1 episode of wheezing in the previous 12 months

[§] Bronchial hyperresponsiveness and ≥1 episode of wheezing in the previous 12 months
Bronchial hyperresponsiveness defined as PC_{20} $FEV_1 \leq 8.0$ mg/ml.
PC_{20} is concentration of methacholine that induced a 20% fall in FEV_1 (forced expiratory volume in one second)

[†] Bronchial hyperresponsiveness and ((wheezing and/or shortness of breath and/or recurrent coughing) and/or (use of β2 agonists and/or inhaled corticosteroids)) during the previous 12 months.
Bronchial hyperresponsiveness defined as PC_{20} $FEV_1 \leq 8.0$ mg/ml.
PC_{20} is concentration of methacholine that induced a 20% fall in FEV_1 (forced expiratory volume in one second)

¶ Definition used in the ARCADE study

~ Fixed set of predictors are: specific IgE (directed against house dust mite, cat and dog dander), eczema (during the previous year), and wheezing (apart from colds during the previous year)

‡ Example: proportion of children with probability between 25% and 60%
A difference of -21.3% between the models using Dr-ever&whe (definition 2) versus BHR&whe (definition 3) means that using the definition of BHR&whe results in 21.3% more children that stay in the area of clinical indecision as compared to using definition Dr-ever&whe

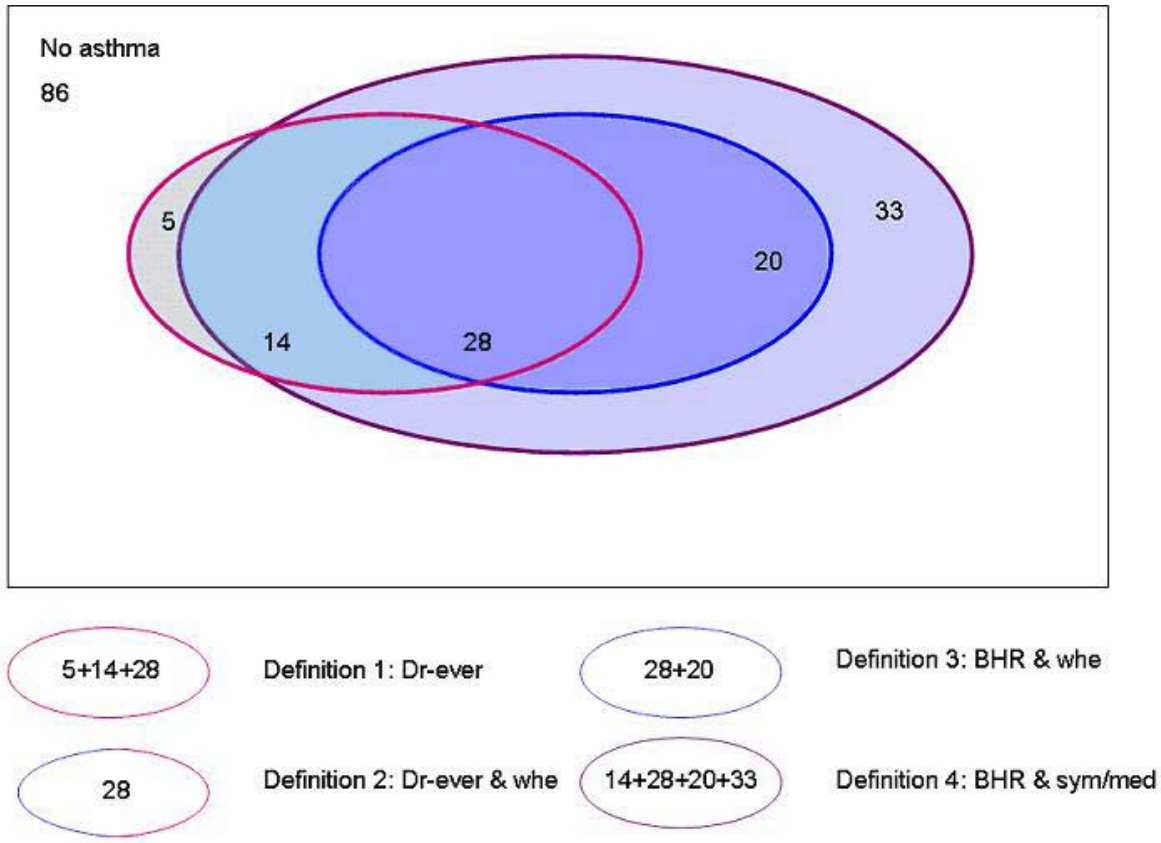
* 10th, 50th and 90th centiles

n/a Not available since this definition covers the period back to a child's birth which, strictly speaking, defeats the purpose of prediction

& All differences and their 95% confidence intervals were calculated using bootstrapping procedures (1000 times)

$ Differences between the posterior probability distributions are expressed as p-values and were calculated by bootstrapping the Mann-Whitney ranksum statistics 1000 times and averaging across the 5 imputed data sets. In none of the 5 data sets the upper 95% confidence limit was greater than 0.0001.

Figure 1: Agreement between four different definitions of asthma in a cohort of 186 six-year-old children



Numbers are n

Figure 1 shows that 15% (28/186) and 46% (86/186) of the children were defined as having and not having asthma by all four definitions, respectively (overall agreement 61%). This figure shows also that almost all children (95/100) who were defined as having asthma by definition definition 1 (Dr-ever), definition 2 (Dr-ever&whe), or definition 3 (BHR&whe) had asthma according to definition 4 (BHR&sym/med). Although prevalence estimates for definition 1 (Dr-ever) and definition 3 (BHR&whe) were similar, 47/186 (25.3%) and 48/186 (25.8%) respectively, figure 1 shows that definitions 1 and 3 nevertheless disagree in 39/186 (21%) of children.