Received: 25 May 2018 | Revised: 27 September 2018 | Accepted: 11 October 2018

DOI: 10.1111/ppe.12525

ORIGINAL ARTICLE

WILEY Paediatric and Perinatal Epidemiology

Diagnosing atopic dermatitis in infancy: Questionnaire reports vs criteria-based assessment

Christoffer Dharma[1] | Diana L. Lefebvre[1] | Maxwell M. Tran[1] | Zihang Lu[2,3] | Wendy Y. W. Lou[2] | Padmaja Subbarao[3] | Allan B. Becker[4] | Piush J. Mandhane[5] | Stuart E. Turvey[6] | Theo J. Moraes[3] | Meghan B. Azad[4] | Malcolm R. Sears[1] | for the CHILD Study Investigators*

[1]Department of Medicine, McMaster University, Hamilton, Canada

[2]Dalla Lana School of Public Health, University of Toronto, Toronto, Canada

[3]Department of Pediatrics, University of Toronto & Hospital for Sick Children, Toronto, Canada

[4]Department of Pediatrics & Child Health, University of Manitoba, Winnipeg, Canada

[5]Department of Pediatrics, University of Alberta, Edmonton, Canada

[6]Department of Pediatrics, University of British Columbia, Vancouver, Canada

Correspondence
Malcolm R. Sears, Firestone Institute for Respiratory Health, St. Joseph's Healthcare and McMaster University, Ontario, Canada.
Email: searsm@mcmaster.ca

Funding information
Canadian Institutes of Health Research (CIHR); Allergy Genes and Environment Network of Centers of Excellence (AllerGen NCE)

Abstract

Background: Persisting atopic dermatitis (AD) is known to be associated with more serious allergic diseases at later ages; however, making an accurate diagnosis during infancy is challenging. We assessed the diagnostic performance of questionnaire-based AD measures with criteria-based in-person clinical assessments at age 1 year and evaluated the ability of these diagnostic methods to predict asthma, allergic rhinitis and food allergies at age 5 years.

Methods: Data relate to 3014 children participating in the Canadian Healthy Infant Longitudinal Development (CHILD) Study who were directly observed in a clinical assessment by an experienced healthcare professional using the UK Working Party criteria. The majority (2221; 73.7%) of these children also provided multiple other methods of AD ascertainment: a parent reporting a characteristic rash on a questionnaire, a parent reporting the diagnosis provided by an external physician and a combination of these two reports.

Results: Relative to the direct clinical assessment, the area under the Receiver Operating Characteristic curve for a parental report of a characteristic rash, reported physician diagnosis and a combination of both were, respectively, 0.60, 0.69 and 0.70. The strongest predictor of asthma at 5 years was AD determined by criteria-based in-person clinical assessment followed by the combination of parental and physician report.

Conclusions: These findings suggest that questionnaire data cannot accurately substitute for assessment by experienced healthcare professionals using validated criteria for diagnosis of atopic dermatitis. Combining the parental report with diagnosis by a family physician might sometimes be appropriate (eg to avoid costs of a clinical assessment).

KEYWORDS

accuracy, atopic dermatitis, diagnosis, eczema, epidemiology, questionnaire reports

*Canadian Healthy Infant Longitudinal Development Study Investigators are listed in Appendix 1

1 | INTRODUCTION

Atopic dermatitis/eczema (AD) affects 15%-20% of children worldwide. The risk for progression to more severe allergic diseases varies depending on the trajectories of AD; those whose AD persisted from age 1 year were more likely to develop asthma and allergic rhinitis than those with a later onset.[1,2] Currently, reliable criteria for distinguishing between transitory rashes and AD in infancy are not well established. Young children cannot describe or report itchiness,[3] leading to problems in diagnoses and nomenclature of AD.[4] Furthermore, some two-thirds of children diagnosed with AD are not truly atopic, ie not sensitised to any allergens.[5,6]

The literature relating to clinical trials and cohorts reports several methods for ascertaining AD. Most often, parents are asked about rashes that the child has had in typical locations such as on the face, neck, elbow, behind the knees, hands or feet.[1] Other studies have utilised electronic medical records[7,8] seeking a diagnosis by an attending physician, while others relied on parent-reported physician diagnosis, where parents were typically asked "Has your child been diagnosed with eczema?".[9–12] Some studies have employed direct examination of the child by a health care professional,[13–16] while others have combined more than one of these methods.[13,15] It is unknown if these methods are comparable. Ideally, in a clinical study, infants would all be assessed by the same experienced paediatrician, but this may not be feasible especially in large multicentre studies. Studies in older children have shown that electronic medical records allow identification of cases as accurately as a physician;[7] however, this was not examined among infants.

Using data from the Canadian Healthy Infant Longitudinal Development (CHILD) Study, we have analysed different methods of diagnosing AD in infancy to determine whether a parental report of rashes in classical locations, or parental report of physician-diagnosed AD, are as reliable as diagnoses made by an experienced healthcare professional using established criteria developed by the UK working party.[17] We further determined which diagnostic approaches were most associated with future relevant clinical outcomes such as allergic diseases, sensitisation and wheezing symptoms at 5 years.

2 | METHODS

The CHILD study is a longitudinal birth cohort study following 3455 children recruited at 4 different sites in Canada (Edmonton, Toronto, Vancouver and Manitoba). The study involves in person assessments in the home at 3-4 months of age, and during clinic visits at age 1, 3 and 5 years, as well as health questionnaires completed at 3, 6, 12, 18, 24, 30, 36, 48 and 60 months. Study procedures have been published elsewhere; all study sites received approval from their local Research Ethics Board.[18] The current analysis includes 3014 children who were evaluated for AD at the 1-year clinical assessment; those who were missing information from the questionnaires had data imputed using Multiple Imputation (see Statistical Analysis methods) (Figure 1).

2.1 | Measures of atopic dermatitis

2.1.1 | Diagnosis based on the parental report of a rash (Parental Report)

This was based on an algorithm using data recorded in the child health questionnaires completed by the parent (usually mother) at 3, 6 and 12 months. The wording of the questions for all time points was very similar; parents were asked "Has your child had ANY rash

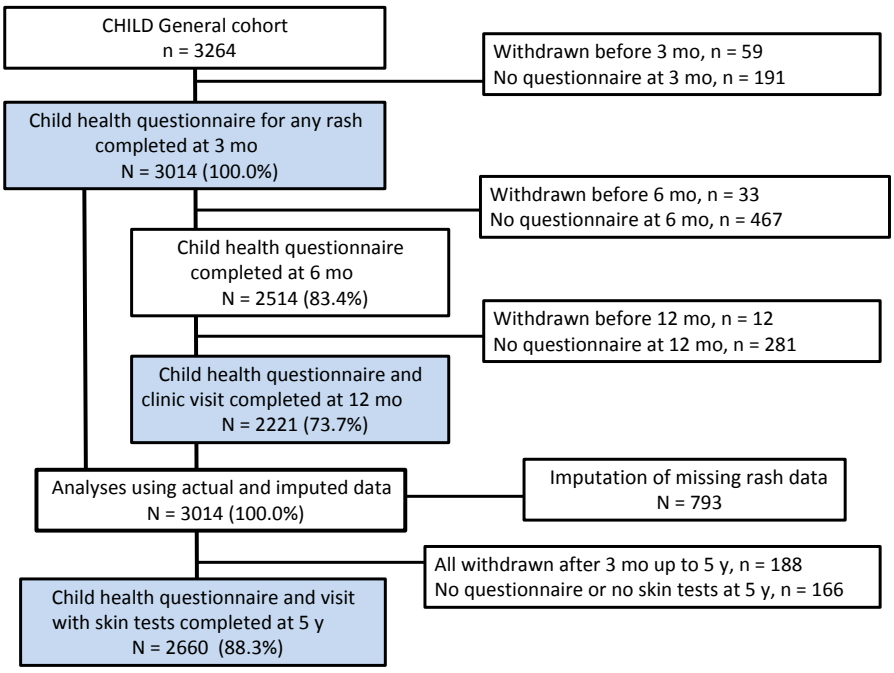


FIGURE 1 Flowchart of participants, showing withdrawals and completion of health questionnaires by age to 1 y and follow-up at 5 y

in the last 3 or 6 months?" If "yes," parents were asked "Where was the WORST rash located?" with options: face, inside elbow(s), diaper area, wrist/hand(s), back of knee(s), scalp, ankle(s) and other. The next question asked the parent to "Describe the rash," with options: wet, red, dry and scaly. Parents could select more than one option for each question. A child was considered to have AD by this measure if the worst rash was in a classical or typical location for infants (on the face, inside of elbow, wrist/hands or back of knees) AND was described as either red or wet (see Appendix 2 for algorithm).

2.1.2 | Diagnosis made by a family physician or non-study paediatrician (External Physician)

If the parent reported the child had had a rash, two additional questions were asked. First, "Was this rash EVER seen by a doctor or healthcare professional?" If yes, "What was the diagnosis for this WORST rash?" with options of eczema/atopic dermatitis, hives or other. A child was classified to have AD by this measure if parents indicated a physician diagnosis of eczema/atopic dermatitis. Children who had not consulted a physician for a rash, or reported no rash, were coded to have no AD diagnosis from an external physician.[9]

2.1.3 | Combining Parental Report with External Physician (Combined Reports)

We evaluated this measure in which a child was considered to have AD only if they were ascertained to have AD by both Parental Report and External Physician Report.

2.1.4 | Diagnosis based on criteria-based clinical assessment (Clinical Assessment) at the 1-year visit

Experienced healthcare professionals (paediatricians or study staff trained and supervised by these physicians) assessed each child at age 1 year for AD using the UK Working Party definition.[17] This required "an itchy skin condition (or a parental report of scratching or rubbing in a child)," and at least one of the following three criteria: "history of involvement of the skin creases of elbows, behind knees, front of ankles or around neck," "history of general dry skin in the last year," and "visible flexural eczema or eczema involving the cheeks/ forehead and outer limbs". On a few occasions (<2% of cases), physicians who were certain that the child had AD despite a lack of reported itch still made a diagnosis of AD. If AD was diagnosed, severity was recorded as "mild," "moderate" or "severe" (see Appendix 3 for definitions of severity).

2.2 | Definition of Other Measures

Allergic sensitisation at 5 years was assessed using skin prick tests (SPTs) in 2660 of these children (88.3% of the 3014 with data at 3 months) (Figure 1). A child was considered to be sensitised if SPT showed ≥2 mm wheal to any of 4 foods (peanut, milk, egg white or soy) or 13 inhalants (alternaria tenuis, cat hair, dog epithelium, house dust mites (Der.p and Der.f), cockroach, penicillium, cladosporium, aspergillus fumigatus, trees, grasses, weeds and ragweed). Details of skin tests procedure have been described previously.[19]

Parental sensitisation was assessed using the same inhalant allergens as used in the children but only one food (peanut) was tested. Mothers were usually tested during the child's clinic visit at age 1 year or later, rather than during pregnancy. Fathers were tested at the time of recruitment to the study.

Parental histories of allergic diseases were self-reported. Mothers and fathers were asked to indicate their history of asthma, allergic rhinitis (hay fever), skin allergy symptoms (eczema, hives or allergic rash) and food allergies.

Diagnoses of asthma, allergic rhinitis and food allergy were made during the clinic visit at age 5 years. Experienced paediatricians in the CHILD Study or highly trained health care professionals assessed all children for these conditions through a structured parental interview, eliciting a history of allergic symptoms. For all allergic diseases, diagnoses were recorded as "yes," "possible" or "no" (see details in Appendix 4); only those assigned "yes" were considered as cases.

2.3 | Statistical analysis

We calculated the sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and Area under the Receiver Operating Curve (AUROC) of an AD diagnosis based on the Parental Report and diagnosis by an External Physician, using the Clinical Assessment as the "gold standard ". We calculated Kappa statistics or chance-corrected agreement between the three measures (including the Combined Reports) with the Clinical Assessments; as well as crude agreements between the Parental Report and the External Physician diagnoses within each severity group.

We have previously reported that the clinical assessment of AD using the UK Working party criteria provided the strongest correlation with clinical outcomes of allergic diseases at 3 years.[14] We have now extended follow-up time, and evaluated which AD diagnostic measure was most associated with clinical outcomes (atopy, asthma, allergic rhinitis, and food allergy) at age 5 years. We applied unadjusted Poisson regression models to calculate the risk ratios of exhibiting these characteristics associated with AD diagnosed by each method.[20]

To maximise power, we imputed missing questionnaire information in two steps. First, for all children who had been diagnosed as having AD by a specific method at a particular age (eg a rash that matched the criteria required by the Parental Report at 3 months), any missing values at other ages in the first year were also assumed to be positive. Similarly, children who were missing data for a period of only 3 months (ie missed the questionnaire only once at either 3 or 6 months, but completed the questionnaires at every other time) and reported no AD symptoms at both times, were considered as negative.

These first steps of imputation provided 2511 complete cases, leaving 503 to be imputed statistically. We performed multiple imputation using the MICE package in R 3.3.1 in these remaining cases.[21,22]

Multiple imputation was undertaken using all available information from the questionnaires; 50 imputed data sets over 10 iterations were created to give the most reliable estimation (see Appendix 5 for details).[21,23] Regression estimates were pooled over the 50 separate data sets; confidence intervals for sensitivity, specificity and other diagnostic measures were calculated using the bootstrap method with the Bias-corrected and accelerated (BCa) interval.[24] We did not impute missing values for the 1-year clinical assessment, nor the 5-year allergic outcomes predicted by each AD measure. We also conducted a sensitivity analysis which included only the 2221 children who had completed all three questionnaires pertaining to AD at 3, 6 months and 1 year. Finally, to ensure robustness of results regarding agreement among the four measures, we determined best and worst case scenarios for all 503 missing values imputed to assess how estimates would change if all missing cases were in agreement with the gold standard or if all were in disagreement.

3 | RESULTS

3.1 | Demographics of participants and distribution of AD measures

In the overall sample of 3014 children assessed at age 1 year, there were slightly more males (52.9%) than females. Mothers were primarily white Caucasian (76.2%) with a relatively high level of atopic sensitisation (57.8%) and history of any allergic disease (79.2%) (Table 1).

Of the 2221 children with all questionnaires completed throughout the first year of life, 1882 (84.4%) reported a rash (in any location) between birth and age 1 year but in 429 (22.8%) this was diaper rash only; these children were not considered to have AD, leaving 1453 reporting any rash other than diaper rash. The algorithm based on the Parental Report (namely the site and the characteristics of any reported rash) considered 926 of these (41.7% of the total children assessed) to have atopic dermatitis, while fewer (536; 24.1%) were diagnosed by an External Physician, and only 248 (11.2%) infants were so diagnosed at the criteria-based Clinical Assessment. There were 1100 (49.5%) children diagnosed with AD by at least one of the three measures, while only 156 (7.0%) infants were diagnosed by all three measures (Figure 2).

TABLE 1 Participant demographics (n = 3014)

Demographics	N (%)	
Sex		
Female	1420 (47.1)	
Male	1594 (52.9)	
Study centre		
Edmonton	666 (22.1)	
Toronto	684 (22.7)	
Vancouver	689 (22.9)	
Manitoba	975 (32.3)	
Parental history*	Mother (n = 2979)	Father (n = 2528)
Current skin allergy symptoms	961 (32.3)	606 (24.0)
Current allergic rhinitis (hay fever) symptoms	1490 (50.0)	1226 (49.0)
Any history of asthma	658 (22.1)	496 (19.6)
Any history food allergies	654 (22.0)	429 (17.0)
Sensitised (≥1 positive skin test)	1723 (57.8)	1610 (63.9)
Parental ethnicity		
First Nation	119 (4.0)	115 (4.5)
South-East Asian	143 (4.8)	115 (4.5)
East Asian	171 (5.7)	130 (5.1)
South Asian	88 (3.0)	108 (4.3)
Black or Hispanic	106 (3.6)	134 (5.3)
White	2213 (74.3)	2240 (88.6)
Other (including mixed ethnicities)	145 (4.9)	143 (5.7)
Unknown or skipped	27 (0.9)	29 (1.1)

*May not add up to 3014 due to missing data.

3.2 | Comparing Parental Report and External Physician diagnosis of AD with the Clinical Assessment

Based on data for all 3014 children, including those with imputed questionnaire values, the Parental Report algorithm had a sensitivity of 65.2% and specificity of 75.9% using the Clinical Assessment as the "gold standard". This means that among infants who were diagnosed with AD during the clinic visit, 65.2% would have been diagnosed by the Parental Report; however, among infants not diagnosed with AD during the clinic visit, 75.9% were correctly classified as not having AD in the Parental Report. Likewise, only 70.2% of infants diagnosed with AD during the clinic visit were also diagnosed with AD by an External Physician report, while 87.0% of infants without AD by Clinical Assessment were not diagnosed with AD by an External Physician. Combining the Parental and External Physician reports gave a lower sensitivity of 53.2%, but an improved specificity of 92.1%. PPV and NPV for these measures can be interpreted in similar ways. The AUROC for Parental Report, External Physician and Combined Report diagnoses were as follows: 0.60, 0.69 and 0.70, respectively, suggesting a similarly acceptable level of accuracy between the three questionnaire-based methods. The chance-corrected agreements (Kappa statistics) between the Parental Report, External Physician diagnosis and the Combined Report with the Clinical Assessment were 0.25, 0.45 and 0.43 respectively (Table 2). When we considered the worst- and best-case scenario for the missing data, worst case scenario gave an AUROC as low as 55.0 for Parental Report, while the best-case

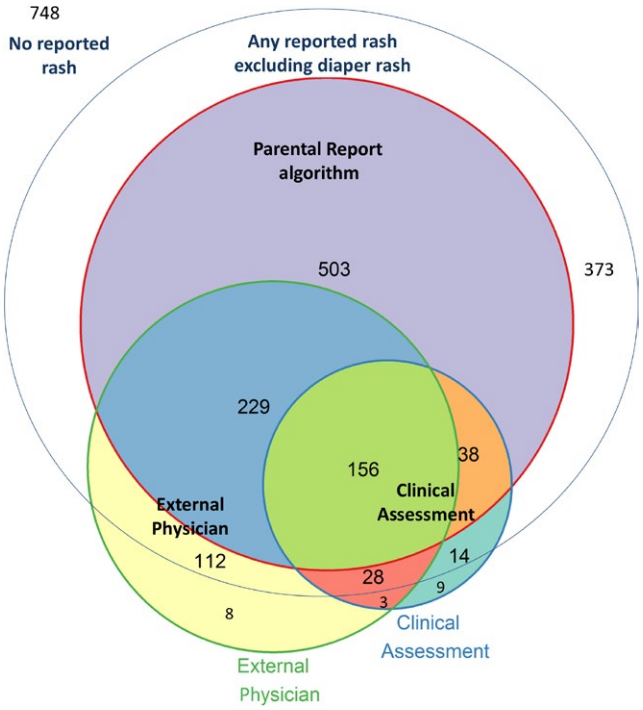


FIGURE 2 Distribution and overlap of various diagnostic methods of identifying AD in the sample. Since multiple imputation method was pooled over 50 iterations of data sets, we are unable to determine to which group the children with imputed data belong. This Venn diagram applies only to the 2221 children who completed all questionnaires.

scenario gave an AUROC as high as 78.1 for the Combined Reports (see Appendix 6A, 6B).

When we considered only the 2221 children who completed all AD questionnaire measures in the first year, similar conclusions were reached, although the sensitivity for Parental Report was lower than specificity (see Appendix 6C). There was a gradient in the crude agreement between the two questionnaire measures of AD with the severity of the AD as determined at the clinical assessment among the 2221 children (Table 2). Only 6 children were rated as "severe," all 6 (100%) were diagnosed to have AD by all measures. Among those with "moderate" and "mild" severity, 74.5% and 59.0%, respectively, were diagnosed by both Parental Report and an External Physician.

3.3 | Predicting future clinical characteristics using each measure of AD in infancy

The risk of acquiring any allergic sensitisation at age 5 years was highest among infants diagnosed with AD during the criteria-based Clinical Assessment (RR: 2.92, 95% CI: 2.40, 3.55). Infants identified with AD by the other two methods were still at a significantly increased risk, but lower than those identified from the clinical assessments. For diagnosis using Combined Reports, the risk for developing allergic sensitisation was higher than the other two individual methods (RR: 2.42, 95% CI: 1.96, 2.97), although still lower than the risk for infants diagnosed with AD by the Clinical Assessments.

AD diagnosis from the Clinical Assessment in infancy significantly increased the risk of being diagnosed with allergic rhinitis (RR: 3.66; 95% CI: 2.67, 4.99), food allergy (RR: 7.22; 95% CI: 5.13, 10.15) and asthma (RR: 2.63; 95% CI: 1.85, 3.74) during the clinical assessments at 5 years. Risks for all the 5-year outcomes were lower for infants who were identified with AD by an External Physician as well as by the Parental Report. When the latter two reports were combined, the risks were closer to the risks obtained for these outcomes by the Clinical Assessments, except for allergic rhinitis, which had a substantially lower risk. These risk ratios for each AD measure are summarised in Figure 3 (see Appendix 7A for details). Results were maintained when we only considered the 2221 children who have complete data for all questionnaires (see Appendix 7B for details).

4 | COMMENT

4.1 | Principal findings

We have compared four methods of diagnosing AD in infancy that are commonly used in clinical and epidemiological studies. In general, taking the Clinical Assessment based on the UK Working Party criteria as the "gold standard," sensitivities and PPV for all three questionnaire-based methods (with imputation where needed) were poor (sensitivities ≤70% for all measures), suggesting that they resulted in many false negatives. In both imputed and complete data sets, while not in perfect agreement with the standardised clinical assessments, diagnosis from an External Physician as reported by parents had a sensitivity, specificity, and NPV of above 70% and Kappa statistic of up to 0.45, indicating moderate agreement.[25] When the Parental Report and External Physician's diagnosis were combined, specificity was the highest, with a similar moderate agreement as shown by the Kappa statistic.[25] We also found a clear gradient in the ability of these four diagnostic measures to predict important allergic outcomes. Infants identified with AD by the Clinical Assessment had a significantly higher likelihood of allergic sensitisation at 5 years compared to those diagnosed by an External Physician or Parental Report, as well as higher risks of 5-year asthma diagnosis compared to infants diagnosed by the External Physician. Combining the two questionnaire-based reports of AD gave an improved predictive ability; the odds of developing these outcomes were closest to that of the children diagnosed by the Clinical Assessment.

4.2 | Strengths of the study

Strengths of the current study include the longitudinal nature of the data with multiple measures of AD within the same child and a standardised definition of allergic diseases. Although these standardised definitions enhance the reliability of diagnosis, the fact that they were seen by the same CHILD Study physicians at age 5 years may contribute to our finding that favours AD at clinical assessment as the best measure to predict allergic outcomes at later

TABLE 2 Diagnostic values and quantitative agreement between the Parental Reports, External Physician's Diagnosis and Combined Reports of AD compared to criteria-based Clinical Assessment diagnosis

	Estimate (%, 95% confidence interval), using the criteria-based Clinical Assessment as the "gold standard"		
All data (n = 3014)	**Parental Report**	**External Physician**	**Combined Reports**
Sensitivity	65.2 (64.9, 65.4)	70.2 (70.0, 70.4)	53.2 (52.9, 53.4)
Specificity	75.9 (75.7, 76.0)	87.0 (86.9, 87.1)	92.1 (92.0, 92.2)
Positive predictive value (PPV)	26.9 (26.7, 27.0)	42.4 (42.3, 42.6)	47.8 (47.6, 48.1)
Negative predictive value (NPV)	94.1 (94.0, 94.2)	95.5 (95.5, 95.6)	93.5 (93.5, 93.6)
Area under the Receiver Operating Curve (AUROC)	0.60 (0.60, 0.61)	0.69 (0.68, 0.70)	0.70 (0.70, 0.71)
	Kappa (95% CI); N = 3014		
	Parental Reports	External Physician	Combined Reports
Agreement with Clinical Assessments	0.25 (0.25, 0.26)	0.45 (0.44, 0.45)	0.43 (0.43, 0.44)
	Number of positive diagnoses from each measure (N = 2221)		
Severity determined at the Clinical Assessment	Parental Report	External Physician	Agreement[a]
Mild (n = 195)	145	144	115 (59.0)
Moderate (n = 47)	43	37	35 (74.5)
Severe (n = 6)	6	6	6 (100)

[a]Agreement refers to the proportion of children diagnosed by both measures within each severity group. Note that we only present results for children with complete data when comparing the severity ratings within each group, since the multiple imputation method iterates over 50 separate data sets and does not provide the exact number of children assigned to each group.

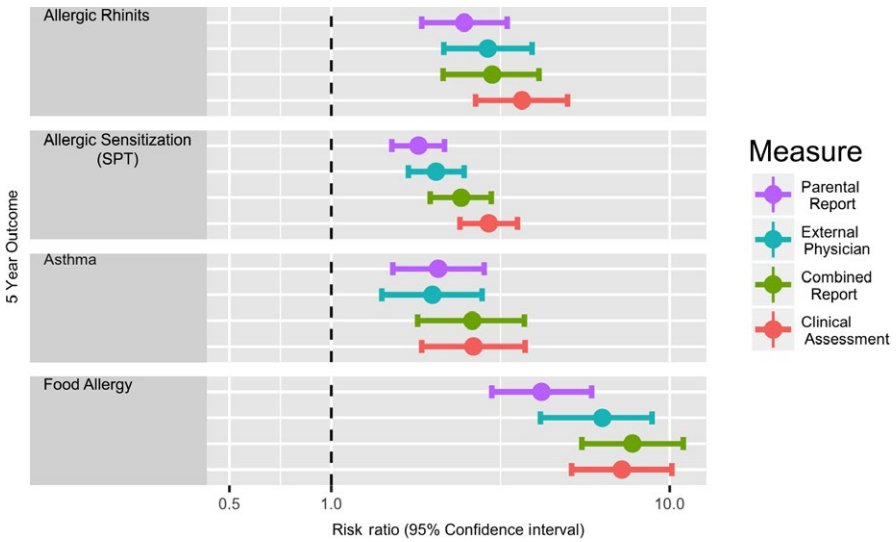


FIGURE 3 Risk ratios for each AD measure in predicting allergic outcomes at age 5 y, comparing each with the criteria-based Clinical Assessments.[a]See Appendix 7A and 7B for a tabular format

ages. However, AD diagnosed at clinical assessment is also the most strongly associated with the objective measure of allergic sensitisation based on skin prick tests; hence any biases that arise from CHILD Study physicians' diagnoses of allergic diseases would appear to be minimal.

4.3 | Limitations of the data

One limitation of the study is reflected in some inconsistencies within the data, namely that some children whose parent had never reported a rash were diagnosed with AD at the clinical

assessments (Figure 1); this may have been due to parents missing rashes which were observed by the study physician during the clinic visit or mistakes made in answering the questionnaire. Furthermore, some parents do not consistently check with a physician whenever the child has a rash; parents whose child was not diagnosed with AD at an earlier physician visit (eg 3 months) may not check with their physician at a later visit. Nevertheless, both of these limitations emulate real-world scenarios of studies that have to solely rely on questionnaires completed by the parents; these parents would have also reported that the child was never diagnosed with AD. Our classification of severity as mild, moderate and severe, while including the clinical criteria of severity included in the Eczema Area and Severity Index (EASI), namely erythema, induration/papulation, excoriation and lichenification, did not estimate the total area of involvement and so we could not calculate an EASI score.[26] Likewise, we do not have biomarker signatures which have been shown to improve precision measurement of disease severity.[27]

Finally, many parents did not complete the questionnaires at all time points, which may bias our results since those completing all questionnaires may have been more concerned over their child's health and so provided more accurate and complete information. Greater discrepancies leading to less accurate diagnosis could be observed in the general population utilising only survey reports. We addressed this potential problem by introducing multiple imputation methods to the data as well as conducting a best and worst case scenario in terms of the agreement with the clinical assessment diagnosis. Although a few differences in results occur comparing the complete and full including imputed data, our results appear robust to these many different scenarios of different plausible imputation values.

4.4 | Interpretation

There are variations in the early presentation of atopic dermatitis and in the patterns of persistence or remission, severity and comorbidity. In a latent class analysis involving two birth cohorts, early-onset and early-resolving atopic dermatitis could be distinguished from early-onset-persistent or early-onset-late resolving atopic dermatitis, the latter being most strongly associated with genetic risk and a personal and parental history of atopic disease.[28] In another European cohort, among those destined to develop atopic dermatitis in the first 4 years, almost 60% presented with AD in the first year of life.[29]

Currently, it is known that AD that persists from age 1 to 3 years results in the highest risk for continuing the "atopic march."[1,12,19] For studies aimed at assessing the risk of developing severe allergic outcomes at later ages, a method that leads to over-diagnosis (such as Parental Reports, as shown in current study) during infancy might be acceptable, as AD in many of these children will resolve at later ages, negating the risk. Combining Parental Reports with the diagnosis from an External Physician might be a suitable alternative in some circumstances (eg to avoid costs of a clinical assessment) since it improves the ability to predict future allergic outcomes compared to using one method of reporting alone. To minimise the cost of a clinical assessment, future studies may also want to consider the feasibility of implementing a more layman friendly question that is based on the UK Working Party criteria to be directly answered by parents, which may help improve diagnostic accuracy of a questionnaire-based report. Research questions that seek to accurately identify infants with AD in infancy ideally require a direct clinical assessment, utilising established criteria for diagnosis.

5 | CONCLUSIONS

These findings indicate that AD diagnosed by an experienced healthcare professional during a clinical assessment in infancy using the UK Working Party criteria provided the best prognostic marker of all allergic outcomes at age 5 years. Examining the child, especially at this early age, is important for an accurate and reliable diagnosis. While direct healthcare professional assessment will be the most costly, researchers that utilise parental reports of AD and physician's diagnosis will need to be cognizant of this difference and interpret their results cautiously, noting the methods of AD ascertainment.

ACKNOWLEDGEMENTS

We are grateful to all the families who took part in this study, and the whole CHILD team, which includes interviewers, nurses, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers and receptionists. The Canadian Institutes of Health Research (CIHR) and the Allergy, Genes and Environment (AllerGen) Network of Centres of Excellence (NCE) provided core funding for CHILD. Additional support has been provided by Health Canada, Environment Canada, Canada Mortgage and Housing Corporation, the Sick Children's Hospital Foundation, Don & Debbie Morrison, the Silver Thread Foundation and the Childhood Asthma Foundation. We also acknowledge the generosity of ALK-Abello, Mississauga, ON, Canada, in supplying all allergens for the study, and Lincoln Diagnostics Inc., Decatur, IL USA, for supplying the Duotip-Test II devices and skin testing kits. M Sears holds the AstraZeneca endowed chair in Respiratory Epidemiology. Funders had no role in the design and analysis of the study.

CONFLICT OF INTEREST

All authors have no conflict of interests to declare.

ORCID

Malcolm R. Sears http://orcid.org/0000-0002-7113-4917

REFERENCES

1. Roduit C, Frei R, Depner M, et al. Phenotypes of atopic dermatitis depending on the timing of onset and progression in childhood. *JAMA Pediatr*. 2017;171:655-662. https://doi.org/10.1001/jamapediatrics.2017.0556.
2. Fennessy M, Coupland S, Popay J, Naysmith K. The epidemiology and experience of atopic eczema during childhood: a discussion paper on the implications of current knowledge for health care, public health policy and research. *J Epidemiol Community Health*. 2000;54:581-589.
3. Dharmage SC, Lowe AJ, Matheson MC, Burgess JA, Allen KJ, Abramson MJ. Atopic dermatitis and the atopic march revisited. *Allergy*. 2014;69:17-27.
4. Silverberg JI, Thyssen JP, Paller AS, et al. What's in a name? Atopic dermatitis or atopic eczema, but not eczema alone. *Allergy*. 2017;72:2026-2030.
5. Williams H, Flohr C. How epidemiology has challenged 3 prevailing concepts about atopic dermatitis. *J Allergy Clin Immunol*. 2006;118:209-213.
6. Flohr C, Johansson S, Wahlgren CF, Williams H. How atopic is atopic dermatitis? *J Allergy Clin Immunol*. 2004;114:150-158.
7. Abuabara K, Magyari AM, Hoffstad O, et al. Development and validation of an algorithm to accurately identify atopic eczema patients in primary care electronic health records from the UK. *J Invest Dermatol*. 2017;137:1655-1662.
8. Balekian DS, Linnemann RW, Castro VM, Perlis R, Thadhani R, Camargo CA. Pre-birth cohort study of atopic dermatitis and severe bronchiolitis during infancy. *Pediatr Allergy Immunol*. 2016;27:413-418.
9. Loo E, Sim J, Goh A, et al. Predictors of allergen sensitization in Singapore children from birth to 3 years. *Allergy Asthma Clin Immunol*. 2016;12:26.
10. Lee E, Lee S, Kwon J, et al. Atopic dermatitis phenotype with early onset and high serum IL-13 is linked to the new development of bronchial hyperresponsiveness in school children. *Allergy*. 2016;71:692-700.
11. Wan J, Mitra N, Hoffstad OJ, Gelfand JM, Yan AC, Margolis DJ. Variations in risk of asthma and seasonal allergies between early- and late-onset pediatric atopic dermatitis: a cohort study. *J Am Acad Dermatol*. 2017;77:634-640. https://doi.org/10.1016/j.jaad.2017.06.013.
12. Paternoster L, Savenije OEM, Heron J, et al. Identification of atopic dermatitis subgroups in children from 2 longitudinal birth cohorts. *J Allergy Clin Immunol*. 2017;141:964-971. https://doi.org/10.1016/j.jaci.2017.09.044.
13. Martin PE, Eckert JK, Koplin JJ, et al. Which infants with eczema are at risk of food allergy? Results from a population-based cohort. *Clin Exp Allergy*. 2015;45:255-264.
14. Tran MM, Lefebvre DL, Dharma C, et al. Predicting the atopic march: results from the Canadian Healthy Infant Longitudinal Development Study. *J Allergy Clin Immunol*. 2017;141:601-607.
15. Berents TL, Carlsen KCL, Mowinckel P, et al. Weight-for-length, early weight-gain velocity and atopic dermatitis in infancy at two years of age: a cohort study. *BMC Pediatr*. 2017;17:141. https://doi.org/10.1186/s12887-017-0889-6.
16. Kamer B, Pasowska R, Dółka E, Blomberg A, Rotsztejn H. Prevalence of atopic dermatitis in infants during the first six months of life: authors' observations. *Postepy Dermatol Alergol*. 2013;30:277-281. https://doi.org/10.5114/pdia.2013.38355.
17. Williams H, Jburney P, Pembroke A, Hay R. The UK Working Party's diagnostic criteria for atopic dermatitis. III. Independent hospital validation. *Br J Dermatol*. 1994;131:406-416.
18. Subbarao P, Anand S, Becker A, et al. The Canadian Healthy Infant Longitudinal Development (CHILD) Study: examining developmental origins of allergy and asthma. *Thorax*. 2015;70:998-1000.
19. Dharma C, Lefebvre DL, Tran MM, et al. Patterns of allergic sensitization and atopic dermatitis from 1 to 3 years: effects on allergic diseases. *Clin Exp Allergy*. 2017;48:48-59. https://doi.org/10.1111/cea.13063.
20. Zou GY. A modified poisson regression approach to prospective studies with binary data. *Am J Epidemiol*. 2004;159:702-706.
21. Buuren SV, Groothuis-Oudshoorn K. MICE: multivariate imputation by chained equations in R. *J Stat Softw*. 2010:1-68. https://www.jstatsoft.org/.
22. R Core Team. R: A Language and Environment for Statistical Computing. Austria: Vienna; 2016.
23. Klebanoff MA, Cole SR. Use of multiple imputation in the epidemiologic literature. *Am J Epidemiol*. 2008;168:355-357.
24. Canty A, Ripley B. Boot: Bootstrap R (S-Plus) Functions. R package version 1.3-2.0, 2017.
25. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33:159-174.
26. Hanifin JM, Thurston M, Omoto M, et al. The eczema area and severity index (EASI): assessment of reliability in atopic dermatitis. *Exp Dermatol*. 2001;10:11-18.
27. Thijs JL, Drylewicz J, Fiechter R, et al. EASI p-EASI: utilizing a combination of serum biomarkers offers an objective measurement tool for disease severity in atopic dermatitis patients. *J Allergy Clin Immunol*. 2017;140:1703-1705.
28. Paternoster L, Savenije OE, Heron J, et al. Identification of atopic dermatitis subgroups in children from 2 longitudinal birth cohorts. *J Allergy Clin Immunol*. 2018;141:964-971.
29. Roduit C, Frei R, Loss G, et al. Development of atopic dermatitis according to age of onset and association with early-life exposures. *J Allergy Clin Immunol*. 2012;130:130-136.

How to cite this article: Dharma C, Lefebvre DL, Tran MM, et al. Diagnosing atopic dermatitis in infancy: Questionnaire reports vs criteria-based assessment. *Paediatr Perinat Epidemiol*. 2018;32:556–567. https://doi.org/10.1111/ppe.12525

APPENDIX 1

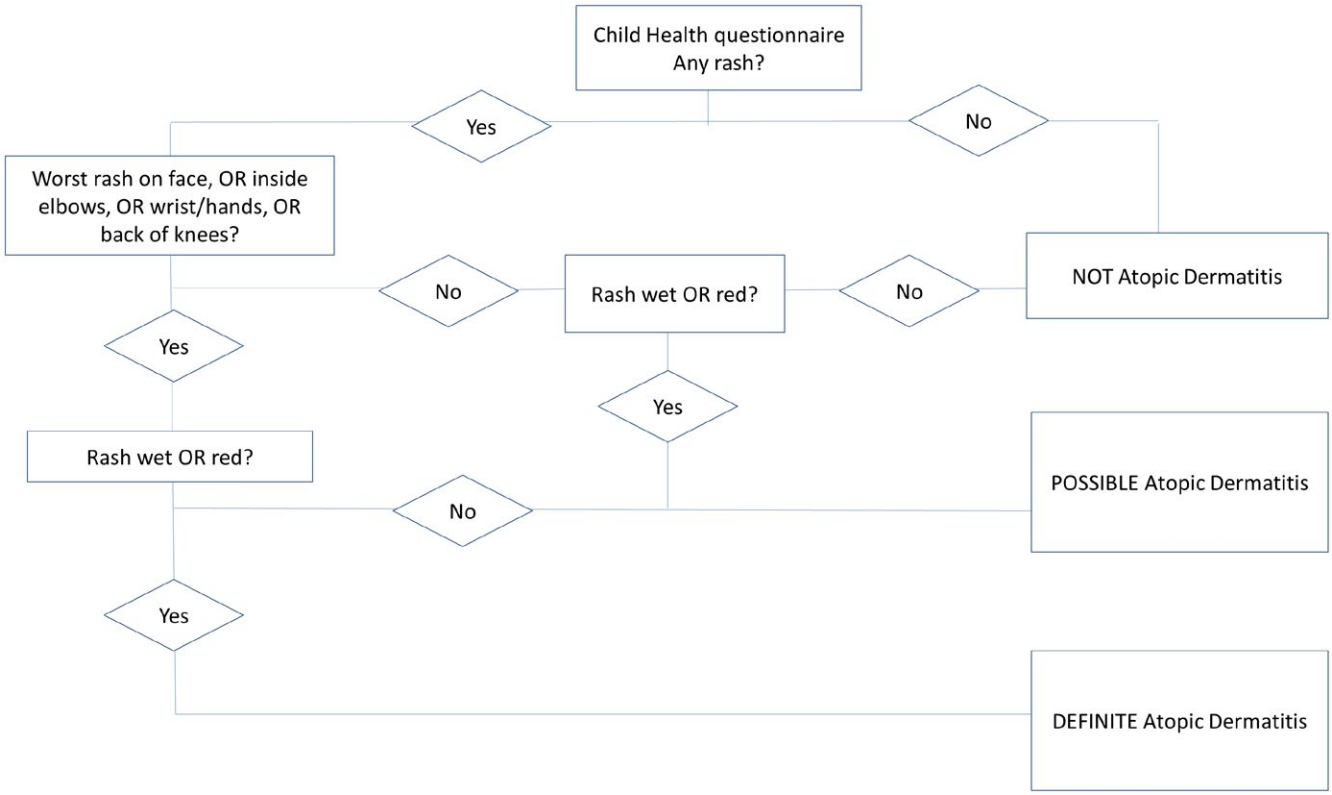
CHILD Study Investigators

Subbarao P (Director), The Hospital for Sick Children & University of Toronto; Turvey SE (co-Director), University of British Columbia; Sears MR, (Founding Director), McMaster University; Anand SS, McMaster University; Azad MB, University of Manitoba; Becker AB, University of Manitoba; Befus AD, University of Alberta; Brauer M, University of British Columbia; Brook JR, University of Toronto; Chen E, Northwestern University, Chicago; Cyr MM, McMaster University; Daley D, University of British Columbia; Dell SD, The Hospital for Sick Children & University of Toronto; Denburg JA, McMaster University; Duan QL, Queen's University; Eiwegger T, The Hospital for Sick Children & University of Toronto; Grasemann H, The Hospital for Sick Children & University of Toronto; HayGlass K, University of Manitoba; Hegele RG, The Hospital for Sick Children & University of Toronto; Holness DL, University of Toronto; Hystad P, Oregon State University; Kobor M, University of British Columbia; Kollmann TR, University of British Columbia; Kozyrskyj AL, University of Alberta; Laprise C, Université du Québec à

Chicoutimi; Lou WYW, University of Toronto; Macri J, McMaster University; Mandhane PJ, University of Alberta; Miller G, Northwestern University, Chicago; Moraes TJ, The Hospital for Sick Children & University of Toronto; Paré P, University of British Columbia; Ramsey C, University of Manitoba; Ratjen F, The Hospital for Sick Children & University of Toronto; Sandford A, University of British Columbia; Scott JA, University of Toronto; Scott J, University of Toronto; Silverman F, University of Toronto; Simons E, University of Manitoba; Takaro T, Simon Fraser University; Tebbutt SJ, University of British Columbia; To T, The Hospital for Sick Children & University of Toronto.

APPENDIX 2

Algorithm used to diagnose AD based on questionnaire data (Parental Report)



APPENDIX 3

Definition of AD severity ratings during the clinical assessment

1 Mild—Single site or no more than 2 sites, minor symptoms (little itching/rubbing), minor crusting and papules, not excoriated or oozing, not needing frequent medical attention
2 Moderate—Neither mild nor severe
3 Severe—Multiple sites, with extensive crusting or papules or excoriations or oozing or lichenification, sleep loss, needing frequent medical attention, major concern to parents

APPENDIX 4

Diagnostic criteria for asthma, allergic rhinitis and food allergy

Asthma is difficult to diagnose with absolute certainty at age 3 years. A paediatric asthma specialist or a highly trained health care professional working under their supervision conducted a structured interview with the accompanying parent or guardian identifying symptoms consistent with asthma, namely recurrent wheeze and coughing without a cold, and noted any physical findings. Asthma was considered definite if the parent reported physician-diagnosed asthma, or use of a bronchodilator prescribed by a physician for coughing or wheezing episodes, or use of a prescribed daily controller medication, or frequent wheezing (3 or more distinct episodes over the previous year) with no alternative diagnosis. Atopy was not essential to the diagnosis but, together with parental history, weighted the likelihood of diagnosing recurrent wheezing as definite asthma. Possible asthma was recorded if there were less frequent episodes of wheeze or coughing without colds and no report of medication use.

Allergic rhinitis was diagnosed based on the questions validated in the ISAAC study, namely nasal symptoms (itching, runny nose, sneezing) without a cold. The combination of symptoms related to particular exposures with a matching positive skin test increased the likelihood of a diagnosis of definite allergic rhinitis, but a positive response to one or more of our representative but limited range of allergens was not considered essential to the diagnosis. Differentiation between definite and possible rhinitis also reflected assessment of the frequency and severity of symptoms, as well as response to therapy with antihistamines or nasal corticosteroids.

Food allergy was considered definite if there was a substantial history consistent with an IgE-mediated response with a matching positive skin test, or previous food allergy testing and diagnosis by a paediatric allergist. Less convincing histories of food allergy were considered possible. Only 4 foods were skin tested in this cohort, and although these represented the most common food allergens in this age group they may not have identified the food allergen involved.

APPENDIX 5

Imputation procedures

The current analysis includes 3014 children who were assessed for AD at the 1-year clinical assessment. Those who did not attend were excluded from analysis as we do not want to impute a value for the "gold standard" measure. The first set of analyses of the sensitivity, specificity and other characteristics of the three evaluated AD measures were conducted on 3014 children including imputed data. The second set of analyses were conducted on 2660 children who attended the clinical assessment at age 5 years. Missing AD information from questionnaires was also imputed in this second analysis, but not values for the 5-years outcomes.

Imputations for missing data for AD in questionnaires were undertaken as follows:

1. Firstly, for children who had a positive response to the Parental Report algorithm (Appendix 2) using any of the 3, 6 months or 1-year questionnaires were immediately considered to have AD based on the Parental Report, even if they had not completed questionnaires at all three-time points. This was similarly undertaken for the External Physician report. This action was taken based on the rationale that even if we had complete data at all three-time points, these children would have been classified as positive regardless of whether the missing data were positive or negative.
2. Secondly, if a child had AT LEAST two data points with "No" responses and missed only one question at EITHER age 3 or 6 months, they were considered as "No". However, if the child missed the questionnaire at age 1 year, they would have missed a period for at least 6 months, hence, we did not impute these cases using this rationale.

Applying the two procedures above gave complete data on 2511 children. The remaining missing data were imputed using the multiple imputation procedure from the MICE package on R, which is based on a logistic regression method. Whenever a data point was missing, we imputed sequentially for each questionnaire at each ages 3, 6 months and 1 year. Since the point was to predict the value, we used all available data to the child, including all reports on AD reported on questionnaires from age 3 months to age 5 years. We created 50 imputed data sets run over 10 iterations. The imputed values at all ages 3, 6 months and 1 year were collated to create single measures for AD Parental Report, AD External Physician, and AD Combined Report, respectively, at age 1 year.

In the first set of analyses, all 3014 children with imputed values were maintained. We calculated sensitivity and specificity over the 50 created data sets. Since there was no direct way to calculate 95% confidence interval over the 50 data sets, we averaged them to get a point estimate for the sensitivity and specificity; we then used the bootstrap procedure over the 50 point estimates using 500 replications with the Bca method to obtain a 95% CI for each of the point estimates. The same method was used for the Kappa statistics and AUROC.

In the second set of analyses, we removed children not attending the assessment at 5 years, leaving 2660 children. The regression estimates were obtained using the pooled estimates package from MICE, which directly pooled all results from the 50 data sets and provide the 95% CI.

Since multiple imputation does not give a single imputed value, but rather 50 different sets, we were able to calculate pooled regression estimates and provide estimates of sensitivity, specificity, PPV and NPV from the pooled results of 50 data sets. However, we are unable to present the exact frequencies of children diagnosed by each measure of AD from the set of 3014, hence, we can only present results for children with complete data (n = 2221) in Figure 2 (Venn diagram) and Table 2 of the agreement with severity ratings, as both require an exact frequency of children diagnosed by each method.

APPENDIX 6A

Diagnostic values and quantitative agreement between the Parental Reports, External Physician's diagnosis and Combined Reports of AD compared to criteria-based Clinical Assessment diagnosis for all 3014 children under different scenarios, invoking the worst case scenario, where all missing data disagree with the Clinical Assessments

	Estimate (95% CI), using the criteria-based Clinical Assessment as the "gold standard"		
All data (n = 3014)	**Parental Report**	**External Physician**	**Combined Reports**
Sensitivity	202/361 0.56 (0.51, 0.61)	215/361 0.60 (0.54, 0.65)	150/361 0.42 (0.36, 0.47)
Specificity	1754/2653 0.66 (0.64, 0.68)	1987/2653 0.75 (0.73, 0.77)	2096/2653 0.79 (0.77, 0.81)
Positive predictive value (PPV)	202/1101 0.18 (0.16, 0.21)	215/881 0.24 (0.22, 0.27)	150/707 0.21 (0.18, 0.24)
Negative predictive value (NPV)	1754/1913 0.92 (0.90, 0.93)	1987/2133 0.93 (0.92, 0.94)	2096/2307 0.91 (0.90, 0.92)
Area under the Receiver Operating Curve (AUROC)	55.0 (53.7, 56.3)	58.8 (57.3, 60.3)	56.0 (54.4, 57.7)
	Kappa (95% CI)		
	Parental Reports	**External Physician**	**Combined Reports**
Agreement with Clinical Assessments	0.12 (0.09, 0.15)	0.21 (0.18, 0.25)	0.15 (0.11, 0.18)
	Number of positive diagnoses from each measure (N = 3014)		
Severity determined at the Clinical Assessment	Parental Report	External Physician	Agreement*
Mild (n = 265)	169	150	97 (36.6)
Moderate (n = 69)	56	44	36 (52.2)
Severe (n = 13)	9	10	8 (61.5)

*Agreement refers to the proportion of children diagnosed by both measures within each severity group.

APPENDIX 6B

Diagnostic values and quantitative agreement between the Parental Reports, External Physician's diagnosis and Combined Reports of AD compared to criteria-based Clinical Assessment diagnosis for all 3014 children under different scenarios, invoking the best case scenario, where all missing data agrees with the Clinical Assessments

	Estimate (95% CI), using the criteria-based Clinical Assessment as the "gold standard"		
All complete data (n = 3014)	**Parental Report**	**External Physician**	**Combined Reports**
Sensitivity	249/361 0.69 (0.64, 0.74)	267/361 0.60 (0.54, 0.65)	212/361 0.59 (0.53, 0.64)
Specificity	2128/2653 0.80 (0.79, 0.82)	2382/2653 0.75 (0.73, 0.77)	2522/2653 0.95 (0.94, 0.96)
Positive predictive value (PPV)	249/774 0.32 (0.29, 0.36)	267/538 0.50 (0.45, 0.54)	212/343 0.62 (0.56, 0.67)
Negative predictive value (NPV)	2128/2240 0.95 (0.94, 0.96)	2382/2476 0.96 (0.95, 0.97)	2096/2307 0.91 (0.90, 0.92)
Area under the Receiver Operating Curve (AUROC)	63.6 (61.9, 65.3)	72.9 (70.8, 75.1)	78.1 (75.6, 80.7)
	Kappa (95% CI)		
	Parental Reports	**External Physician**	**Combined Reports**
Agreement with Clinical Assessments	0.33 (0.29, 0.37)	0.53 (0.48, 0.57)	0.55 (0.50, 0.60)
	Number of positive diagnoses from each measure (N = 3014)		
Severity determined at the Clinical Assessment	**Parental Report**	**External Physician**	**Agreement***
Mild (n = 265)	169	185	138 (52.1)
Moderate (n = 69)	56	57	51 (73.9)
Severe (n = 13)	13	13	13 (100)

*Agreement refers to the proportion of children diagnosed by both measures within each severity group.

APPENDIX 6C

Diagnostic values and quantitative agreement between the Parental Reports, External Physician's diagnosis and Combined Reports of AD compared to criteria-based Clinical Assessment diagnosis, considering only children with complete data (n = 2221)

	Estimate (95% CI), using the criteria-based Clinical Assessment as the "gold standard"		
All complete data (n = 2221)	**Parental Report**	**External Physician**	**Combined Reports**
Sensitivity	194/248 78.2 (73.1, 83.4)	187/248 75.4 (70.0, 80.8)	156/248 62.9 (56.9, 68.9)
Specificity	1241/1973 62.9 (60.8, 65.5)	1624/1973 82.3 (80.6, 84.0)	1744/1973 88.4 (87.0, 89.8)
Positive predictive value (PPV)	194/926 21.0 (18.3, 23.6)	187/536 34.9 (30.9, 38.9)	156/385 40.5 (35.6, 45.4)
Negative predictive value (NPV)	1241/1295 95.8 (94.7, 96.9)	1624/1685 96.4 (95.5, 97.3)	1744/1836 95.0 (94.0, 96.0)
Area under the Receiver Operating Curve (AUROC)	0.71 (0.68, 0.73)	0.79 (0.76, 0.82)	0.76 (0.73, 0.79)
	Kappa (95% CI)		
	Parental Reports	**External Physician**	**Combined Reports**
Agreement with Clinical Assessments	0.19 (0.16, 0.22)	0.39 (0.34, 0.43)	0.41 (0.36, 0.47)

APPENDIX 7A

Risk ratio for each AD measure in predicting allergic outcomes at age 5 y relative to the criteria-based Clinical Assessments including all children (N = 3014)—Tabular format of Figure 3

	Risk ratios for each outcome (95% CI)			
Outcomes at 5 y	**Parental Report**	**External Physician**	**Combined Report**	**Criteria-based Clinical Assessment**
Allergic sensitisation	1.81 (1.51, 2.16)	2.04 (1.69, 2.47)	2.42 (1.96, 2.97)	2.92 (2.40, 3.55)
Asthma	2.07 (1.52, 2.83)	1.99 (1.41, 2.79)	2.61 (1.80, 3.72)	2.63 (1.85, 3.74)
Allergic rhinitis	2.47 (1.85, 3.31)	2.90 (2.15, 3.92)	2.99 (2.14, 4.11)	3.66 (2.67, 4.99)
Food allergy	4.18 (2.98, 5.88)	6.32 (4.51, 8.87)	7.76 (5.50, 10.97)	7.22 (5.13, 10.15)

APPENDIX 7B

Risk ratio for each AD measure in predicting allergic outcomes at age 5 y relative to the criteria-based Clinical Assessments, including only children with complete data for all questionnaires (N = 2221)

	Risk ratios for each outcome (95% CI)			
Outcomes at 5 y[a]	**Parental Report**	**External Physician**	**Combined Report**	**Criteria-based Clinical Assessment**
Allergic sensitisation	1.76 (1.42, 2.17)	1.73 (1.39, 2.15)	2.07 (1.65, 2.60)	2.77 (2.18, 3.52)
Number of events per AD measure (%)	195/790 (24.7)	125/460 (27.2)	107/332 (32.2)	91/213 (42.7)
Allergic rhinitis	1.90 (1.34, 2.69)	2.13 (1.50, 3.02)	2.33 (1.62, 3.36)	3.84 (2.67, 5.54)
Number of events per AD measure (%)	75/818 (9.2)	53/478 (11.1)	43/342 (12.6)	43/223 (19.3)
Food allergy	3.57 (2.27, 5.62)	6.18 (4.02, 9.52)	6.52 (4.32, 9.84)	6.19 (4.11, 9.33)
Number of events per AD measure (%)	67/820 (8.2)	62/478 (13.0)	54/343 (15.7)	41/221 (18.6)
Asthma	1.72 (1.17, 2.52)	1.68 (1.13, 2.51)	1.71 (1.11, 2.64)	2.55 (1.64, 3.98)
Number of events per AD measure (%)	58/821 (7.1)	37/480 (7.7)	28/344 (8.1)	26/224 (11.6)

[a] Number of available AD measures for each outcome may not be consistent due to missing data at the 5 year outcomes