Discordance between diagnosis tools for assessing eczema in infants: a challenge for intervention trials

Pittet, Laure; Messina, Nicole L; Gardiner, Kaya; Freyne, Bridget; Abruzzo, Veronica; Morrison, Clare; Vuillermin, Peter; Allen, Katrina J; Ponsonby, Anne-Louise; Robins-Browne, Roy; Shann, Frank; Flanagan, Katie L; Donath, Susan; Casalaz, Dan [**and 2 more**]

How to cite

PITTET, Laure et al. Discordance between diagnosis tools for assessing eczema in infants: a challenge for intervention trials. In: Dermatitis, 2022, vol. 33, n° 3, p. 207–214. doi: 10.1097/DER.0000000000000842

This publication URL: https://archive-ouverte.unige.ch/unige:172904

Publication DOI: 10.1097/DER.0000000000000842

Discordance between diagnosis tools for assessing eczema in infants: a challenge for intervention trials

ABSTRACT

Background: There is no standardised definition for infant eczema and various tools have been used across studies, precluding direct comparison.

Objective: To assess and compare the accuracy of diagnostic tools for infant eczema using the extensive data collected in MIS BAIR, an eczema prevention trial.

Methods: Eczema incidence was assessed by three questionnaire-based measures: modified UK diagnostic tool, parent-reported medically diagnosed eczema and parent-reported use of topical steroids. Agreement between the definitions was quantified using kappa coefficient. Eczema severity was assessed by 3-monthly POEM scores and a SCORAD clinical assessment at a 12-month visit. ClinicalTrial.gov: NCT01906853.

Results: Among the 538 participants fulfilling at least one of the three questionnaire-based eczema definitions, only 197 (37%) participants met all three definitions. Agreement between the definitions was poor with kappa coefficients ranging from -0.11 to 0.62. The most frequently reported symptoms were generally dry skin (483/538, 90%) and pruritus (400/538, 74%). The face (352/538, 65%) and the trunk (306/538, 57%) were more frequently affected than the creases (257/538, 48%). Participants fulfilling all three questionnaire-based definitions of eczema were more likely to have higher severity scores and earlier onset of symptoms.

Conclusions: There is poor agreement between currently available tools for assessing infant eczema.

Capsule summary

- Only 37% of 538 infants with eczema fulfilled all of three widely used definitions (modified UK diagnostic tool, medically diagnosed, and topical steroids use)
- Caution should be exercised in the interpretation of the results of eczema prevention trials in light of the imprecision of the currently available definitions.
- Improved diagnostic tools are needed for assessing infant eczema in trials.

Key words

Atopic dermatitis; patient-oriented eczema measure, POEM; prevention; SCORing Atopic Dermatitis scoring system, SCORAD; William's UK diagnostic criteria.

Word count

2484 words, 21 references, 3 tables, 4 figures, 1 appendix.

INTRODUCTION

An increasing number of clinical studies are assessing preventive interventions to address the rise in atopic diseases.[1] Various definitions are used for eczema, each of them having limitations, particularly in infants.[2, 3] Although the Hanifin and Rajka criteria are considered by many as the standard,[4] they are complex and inconvenient to use in non-hospital settings. Simpler tools, such as the UK diagnostic criteria, are commonly used in clinical studies.[2, 3] There is a need for a standardised case definition for eczema, particularly in infants, that are easy to use in clinical trials.

One aim of the *Melbourne Infant Study: BCG for allergy and infection reduction* (MIS BAIR) trial was to determine whether neonatal bacille Calmette-Guérin (BCG) vaccination reduces the prevalence of eczema in the first years of life.[5, 6] The primary outcome for the cumulative incidence of eczema at one year of age was defined using the modified version of the UK diagnostic tool,[7, 8] and various other definitions were used for secondary outcomes.[5] Here we used the data from the MIS BAIR trial to highlight the discordance between a range of easy-to-use eczema definitions, and their limitations. We also explored the severity of disease across the various outcome measures.

METHODS

The MIS BAIR trial

MIS BAIR is a randomised controlled trial of 1272 infants allocated at birth to receive BCG or no BCG vaccine and followed up for at least 12 months, in Victoria, Australia. Parents were asked to complete online questionnaires using REDCap[9] at 3, 6, 9 and 12 months of age, and to attend a 12-month clinic visit.[5, 6] Because BCG causes a visible scar, families were aware of the group allocation, but the statisticians and the research nurses doing the 12-month visit remained blinded.

Eczema measures

Tables 1 and 2 list the tools used to measure the incidence and severity of eczema, respectively, and highlight their respective strengths and limitations. Further details on questionnaires and outcome definitions can be found in the supplementary appendix. Briefly, the cumulative incidence of eczema was assessed using three different questionnaire-based measures: (A) a modified version of the UK diagnostic tool;[7, 8] (B) parent-reported medically diagnosed eczema; and (C) parent-reported use of topical steroids. A fourth measure was (D) clinical assessment by an intervention-blinded research nurse at the 12-month visit. In the MIS BAIR trial, the modified version of the UK diagnostic tool (A) was the primary outcome measure for eczema; the other definitions (B, C, and D) were reported as secondary outcomes, including a combination of A and B, reported as the 'extended definition of eczema'.[5]

Eczema severity was estimated by the Patient Oriented Eczema Measure (POEM) score,[10] included in each 3-monthly questionnaire, and the SCORing Atopic Dermatitis scoring system (SCORAD)[11] assessed by the research nurse at the 12-month visit.

Statistical analysis

Agreement between the diagnostic tools was quantified using kappa coefficient, categorised as previously defined.[12] The proportion of participants reporting eczema symptoms (e.g. pruritus, dry skin) was assessed in the 538 participants fulfilling at least one of the questionnaire-based eczema definitions (A, B, or C; hereafter defined as 'having eczema'), as well as within each group and subgroup of definitions. The age at which eczema symptoms occurred, was taken as the time point (either 3, 6, 9, or 12 months) at which the participant first reported a given symptom. The mean (and standard deviation, SD) time point was calculated within each group and subgroup, as well as the mean POEM and SCORAD scores. Stata v.16 (StataCorp, College Station, Texas) was used for analysis.

Ethics and registration

This trial was done in accordance with NHMRC National Statement on Ethical Conduct in Human Research (2007) and the Note for Guidance on Good Clinical Practice (CPMP/ICH-135/95) and was approved by the Mercy Health Human Research Ethics Committee (R12-28) and all participating hospitals. A parent of each participant gave written informed consent prior to inclusion. ClinicalTrial.gov: NCT01906853.

RESULTS

Baseline characteristics and follow-up

Characteristics of the 1272 infants in the trial are presented in Table 3. Most families reported a history of atopic disease (83%, 1049/1271), reflecting the trial's inherent selection bias in recruiting to an allergy prevention trial. Most participants completed all four 3-monthly questionnaires (86%, 1091/1272), and attended the 12-month visit (86%, 1094/1272) at a median age of 1.1 years old (IQR 1.1 to 1.2).

Eczema incidence in the first year of life

Overall, 370 of 1103 participants (34%) fulfilled the modified UK diagnostic criteria for eczema in the first year of life (data missing for 169 participants); an additional 290 participants reported 3 to 4 minor criteria but never reported pruritus (the major criterion). Parents of 337 of 1164 participants (29%) reported a medical diagnosis of eczema (data missing for 108 participants), and 377 of 991 participants (38%) reported using topical steroids (data missing for 281 participants).

Discordance between the three eczema definitions is illustrated in Figure 1, as well as the additional participants who, while having 3 to 4 minor criteria, did not fulfil the UK diagnostic criteria as they did not report pruritus. Among the 538 participants fulfilling at least one of the three questionnaire-based eczema definitions ('having eczema'), there were only 197 participants (37%) who met all three definitions. The agreement between definitions was poor with low kappa coefficient (Figure 2), except between parent-reported medical diagnosis and topical steroid use (kappa 0.62). The agreement between the point prevalence of active eczema as assessed by the research nurse at the 12-month visit and the questionnaire-based tools is shown in Figure S1 (supplementary appendix).

Eczema symptoms and distribution

The symptoms reported from participants in the different groups and subgroups of eczema definitions are presented in Figure S2 (supplementary appendix). Pruritus (74%, 400/538) and generally dry skin (90%, 483/538) were reported by a high proportion of all participants with eczema. Creases were affected in 48% (257/538) of participants with eczema; 70% (137/197) of those fulfilling all three definitions. The face (65%, 352/538) and the trunk (57%, 306/538) were the most commonly reported affected areas in participants with eczema, affecting 84% (165/197) and 78% (153/197) of infants fulfilling all three definitions, respectively.

Age of onset

The mean age of reported eczema symptoms in the 3-monthly questionnaire across the different sub-groups are presented in Figure 3. Participants reporting symptoms at a younger age were more likely to fulfill all three questionnaire-based definitions of eczema. Participants who were medically diagnosed only or who were reporting steroid use only and met less than 3 minor UK criteria were more likely to have a later onset of eczema. Among the symptoms reported, involvement of the face was reported at a mean age of 4.9 months (SD 2.7), involvement of the trunk was reported at a mean age of 6.6 months (SD 3.0), involvement of the creases was reported at a mean age of 6.1 months (SD 2.8). Generally dry skin was reported at a mean age of 5.1 months (SD 2.7), whereas pruritus was reported at a mean age of 6.9 months (SD 3.0).

Severity scores

Eczema severity measured at the 12-month SCORAD assessment and by the 3-monthly POEM scores are shown in Figure 4. Participants fulfilling all three questionnaire-based definitions of eczema were more likely to have higher severity scores. Participants who only reported using steroids and/or having 3 to 4 minor UK diagnostic criteria without reporting pruritus and had not been medically diagnosed had low SCORAD scores.

DISCUSSION

In the absence of an agreed easy-to-use tool to reliably estimate the cumulative incidence of eczema in infancy, various definitions have been used.[2] Using the data from our large prevention trial, we found discordance between three different measures of infant eczema, with only approximately one third of participants with presumed eczema (as defined as those fulfilling at least one of the three questionnaire-based eczema definitions) meeting all three definitions (modified UK diagnostic criteria, parent-reported medically diagnosed, and topical steroids treatment).

The results of previous studies of the cumulative incidence of eczema have been prone to similar discordance.[13-15] In a cross-sectional study of infant allergy, among the 2129 infants who fulfilled at least one of the measures of eczema, only 20% fulfilled all three measures (parent-reported medically diagnosed eczema, parent-reported itchy rash treated with topical steroids, and objective assessment of eczema at 12 months of age).[13] In another RCT investigating the effect of neonatal BCG on eczema incidence, there were discrepancies between the number of participants with parent-suspected eczema, parent-reported medically diagnosed eczema, parent-reported use of topical steroids, and clinical assessment of eczema at a 13-month visit; no details were provided on the overlapping rates.[14] In a cohort study, among the 70 infants who fulfilled at least one of the eczema definitions at 12 months of age, only 18 (26%) were positive for all four measures (Hanifin and Rajka criteria, Schultz and Larsen criteria, UK diagnostic criteria, and objective assessment using study-defined criteria) with no clear pattern of agreement or disagreement between tools.[15] Interestingly, another 18 participants of the 70 (26%) were positive for the objective assessment only, and negative for all questionnaire-based tools.

Discordance in categorisation by definition tools might reflect a different ability to distinguish between atopic and non-atopic eczema; transient eczematous rashes occur in a high proportion of infants, but only a small proportion will have eczema after the age of 3 years.[16]

One of the major limitations of the UK diagnostic tool in the MIS BAIR trial is the constraints inherent in its use in infants. Having been developed for use in older children and adults, the score requires the subject to have itch (major criterion). This excludes the many infants with eczema who do not have significant itch or might not yet be developmentally capable to scratch themself. Also, the minor criteria are primarily based on flexural lesions which is not the typical distribution in infancy. Infants often have eczema predominantly on the trunk [not included anywhere in the criteria] and face [which is included in the minor criteria of the modified version].[16, 17] As parents are considered able to accurately report eczema in their infants,[18] two of the minor criteria are virtually the same. Finally, when recruiting to an eczema prevention trial, participants are more likely to have a family history of atopic disease, which results in most participants having at least one minor criterion when using the modified version.

Relying only on parent-report of medically diagnosed eczema seems insufficient. Parents may not seek medical attention, or may have differing thresholds for seeking attention, particularly atopic families who know how to manage eczema. Moreover, the accuracy of the data is dependent on the ability of individual healthcare workers to diagnose eczema, as well as on parental interpretation of the doctor's explanation. This measure captures all grades of severity of eczema and therefore does not differentiate severe from transient forms. It is likely to capture transient eczema that is not necessarily atopic dermatitis.

Parent-reported use of topical steroids is attractive as a possible indirect and approximate measure of eczema, as it should capture cases of clinically relevant eczema. It is also a reflection of health service use. As we have witnessed in the MIS BAIR trial, limitations of this measure include the possibility of parents confusing other topical treatments with steroids, that it includes self-prescribed steroids available over the counter, and that it includes steroids given for other indications (e.g. cradle cap or diaper rash). We partially corrected for this latter problem by assuming that diaper rash rather than eczema was the likely diagnosis in any infant treated solely with a mild steroid / antifungal combination. In some countries a further limitation of this possible measure is that eczema is mostly treated with emollients during infancy to avoid steroid use.

Clinical assessment of active lesions by study staff is a robust measure that can be based on objective criteria and is considered as the gold standard.[3, 19, 20] Training of assessors can ensure consistency. However, this measure is limited as it only captures participants with lesions on the day of the visit. It is also strongly influenced by the seasonal fluctuation of eczema symptoms and the quality of management, as even severe eczema can have limited manifestations on the visit day when well treated. This measure estimates a point prevalence of active eczema, and therefore the agreement rate with questionnaire-based tools should be interpreted cautiously as the latter measures cumulative incidence of eczema disease over a longer period.

Among the numerous diagnostic tools proposed to define eczema, the Hanifin and Rajka score is the oldest.[4] Designed by clinicians for clinicians, this score is not practical to use in clinical studies as it includes 27 criteria, some of which are cumbersome, require biological testing or expert examination (e.g. ophthalmologist). The score was found to be inconvenient in non-hospital settings, and therefore inspired the development of simpler tools. Among them, the UK diagnostic tool is probably the most widely used in eczema prevention trials;[2] the International Study of Asthma and Allergies in Childhood (ISAAC) tool is usually preferred in large epidemiological surveys, being based on 3 simple questions,[21] but it performs poorly on an individual level.[3] Finally, the recently designed questionnaire-based REACH diagnostic tool does not live up to its name (Reliable Estimation of Atopic dermatitis in Childhood) in that it is not suitable in infants as it does not consider the typical infant distribution of eczema lesions.[22]

In our setting, a combination of the modified UK diagnostic tool and parent-reported medically diagnosed eczema (called the extended definition of eczema) may be the most suitable as it captured the children with the highest SCORAD score at the 12-month visit. Although the use of the UK diagnostic tool in infants appears limited by pruritus being a major criterion, in our trial over 74% of the infants meeting at least one of the three questionnaire-based eczema definitions reported pruritus at least once, mostly after 6 months of age. Moreover, most non-itchy infants who fulfilled 3 to 4 minor criteria had overall low severity scores, suggesting only very mild or no eczema, and the agreement rate with the other eczema definition tools was poor.

Based on our findings, we propose that revised eczema diagnosis criteria for infants should: (i) better consider the typical distribution of eczema in infancy, namely the convexities, such as the trunk and the face; and (ii) include a minimum duration of dermatitis to exclude most transient erythematous self-resolving eruptions in infants.

We plan to re-evaluate the incidence of eczema and other atopic manifestations in MIS BAIR participants at 5 years of age to determine which of the tools best predicts which infants had mild transient eczema, and which have severe persisting eczema, since the main aim of a diagnostic tool would be to predict the latter.

Research on the prevention of eczema has been hampered by inconsistency in outcome definitions,[2] and our data underline the need for the development of more reliable tools to diagnose infant eczema. Our data also show that the face and the trunk, which are rarely considered in the eczema definition tools available, are more commonly affected than the creases in this age group. A robust tool to diagnose infant eczema would enable preventive interventions to be more accurately assessed in trials, and to be compared with one another.

References

1. Asher MI, Montefort S, Bjorksten B, et al. Worldwide time trends in the prevalence of symptoms of asthma, allergic rhinoconjunctivitis, and eczema in childhood: ISAAC phases one and three repeat multicountry cross-sectional surveys. *Lancet*. 2006;368(9537):733-43.
2. Simpson EL, Keck LE, Chalmers JR, et al. How should an incident case of atopic dermatitis be defined? A systematic review of primary prevention studies. *J Allergy Clin Immunol*. 2012;130(1):137-44.
3. Lee SC. Various diagnostic criteria for atopic dermatitis (AD): A proposal of Reliable Estimation of Atopic Dermatitis in Childhood (REACH) criteria, a novel questionnaire-based diagnostic tool for AD. *J Dermatol*. 2016;43(4):376-84.
4. Hanifin J, Rajka G. Diagnostic features of atopic dermatitis. *Acta Derm Venereol (Stockh)*. 1980;92(suppl.):44-47.
5. Pittet LF, Messina NL, Gardiner K, et al. Prevention of infant eczema by neonatal bacille Calmette-Guérin vaccination: The MIS BAIR randomized controlled trial. *Allergy*. 2021;
6. Messina NL, Gardiner K, Donath S, et al. Study protocol for the Melbourne Infant Study: BCG for Allergy and Infection Reduction (MIS BAIR), a randomised controlled trial to determine the non-specific effects of neonatal BCG vaccination in a low-mortality setting. *BMJ Open*. 2019;9(12):e032844.
7. Williams HC, Burney PG, Hay RJ, et al. The U.K. working party's diagnostic criteria for atopic dermatitis. I. Derivation of a minimum set of discriminators for atopic dermatitis. *Br J Dermatol*. 1994;131(3):383-96.
8. Williams HC, Burney PG, Pembroke AC, et al. The U.K. working party's diagnostic criteria for atopic dermatitis. III. Independent hospital validation. *Br J Dermatol*. 1994;131(3):406-16.
9. Harris PA, Taylor R, Thielke R, et al. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform*. 2009;42(2):377-81.
10. Charman CR, Venn AJ, Williams HC. The patient-oriented eczema measure: development and initial validation of a new tool for measuring atopic eczema severity from the patients' perspective. *Arch Dermatol*. 2004;140(12):1513-9.
11. Severity scoring of atopic dermatitis: the SCORAD index. Consensus report of the european task force on atopic dermatitis. *Dermatology*. 1993;186(1):23-31.
12. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33(1):159-74.
13. Martin PE, Koplin JJ, Eckert JK, et al. The prevalence and socio-demographic risk factors of clinical eczema in infancy: a population-based observational study. *Clin Exp Allergy*. 2013;43(6):642-51.
14. Thostesen LM, Kjaergaard J, Pihl GT, et al. Neonatal BCG vaccination and atopic dermatitis before 13 months of age: A randomized clinical trial. *Allergy*. 2018;73(2):498-504.
15. Jøhnke H, Vach W, Norberg LA, et al. A comparison between criteria for diagnosing atopic eczema in infants. *Br J Dermatol*. 2005;153(2):352-8.
16. Halkjaer LB, Loland L, Buchvald FF, et al. Development of atopic dermatitis during the first 3 years of life: the Copenhagen prospective study on asthma in childhood cohort study in high-risk children. *Arch Dermatol*. 2006;142(5):561-6.
17. Barnetson RS, Rogers M. Childhood atopic eczema. *BMJ*. 2002;324(7350):1376-9.
18. Fleming S, Bodner C, Devereux G, et al. An application of the United Kingdom Working Party diagnostic criteria for atopic dermatitis in Scottish infants. *J Invest Dermatol*. 2001;117(6):1526-30.
19. Eichenfield LF, Hanifin JM, Luger TA, et al. Consensus conference on pediatric atopic dermatitis. *J Am Acad Dermatol*. 2003;49(6):1088-95.
20. Saeki H, Furue M, Furukawa F, et al. Guidelines for management of atopic dermatitis. *J Dermatol*. 2009;36(10):563-77.
21. Asher MI, Keil U, Anderson HR, et al. International Study of Asthma and Allergies in Childhood (ISAAC): rationale and methods. *Eur Respir J*. 1995;8(3):483-91.
22. Lee SC, Bae JM, Lee HJ, et al. Introduction of the Reliable Estimation of Atopic Dermatitis in ChildHood: novel, diagnostic criteria for childhood atopic dermatitis. *Allergy Asthma Immunol Res*. 2016;8(3):230-8.

Tables

Table 1: Eczema definition tools used in the MIS BAIR trial

	Description	Advantage	Limitations
A. Modified UK diagnostic criteria	Major criterion (mandatory): itchy skin reported at any 3mQ. Minor criteria (3 or more of the following): 1. History of involvement of the skin creases or cheeks, at any 3mQ 2. History of atopic disease (asthma, hay fever or eczema) in a first-degree relative 3. History of a general dry skin, at any 3mQ 4. Visible flexural eczema or eczema involving the head or limbs, at any 3mQ	• Well known • Widely used • Simple and convenient	• Less relevant for infant eczema: ○ often not accompanied by itch ○ flexural involvement less significant ○ involvement of non-flexural areas (such as torso) is common but not included in the criteria • Theoretically requires a clinic visit for the minor criterion 4, although parents are considered able to accurately report dermatitis in their infants[18] • Low specificity, limited ability in differentiating from other itchy skin rashes ○ higher false positivity rate in scabies-endemic areas
B. Parent-reported medically diagnosed	Positive answer to the following question in any of the 3mQ *"Has [child's-name] ever been diagnosed with eczema?"*	• Single question • Involves medical assessment	• Families may not seek medical attention • Depends on healthcare worker's experience • Susceptible to parental misunderstanding • Various severity, as detecting any eczematous lesion, not necessarily atopic dermatitis
C. Parent-reported use of topical steroid	Positive answer to the following question at any of the 3mQ *"Have you used any topical steroids on [child's-name]'s skin in the last 3 months?"* AND for the following question "*What is the name of the steroid cream you have used?*", the participants did not list only creams that do not contain steroids (e.g. emollient), or steroid-containing creams that are not typically used in isolation to treat eczema (e.g. steroid-containing cream for diaper rash)	• Simple question • Captures clinically significant lesions	• Can be confused with other topical agent • Topical steroid use could be for diaper rash or cradle cap • In some countries, eczema is mostly treated with emollient during infancy to avoid steroid usage
D. Researcher-diagnosed eczema	Participant diagnosed with eczema by the study nurse at the 12-month study visit, defined as a SCORAD score of 10 or more (mild to severe, see Table 2).	• Robust • Based on validated criteria • Use of same assessment can ensures consistent measure	• Only captures lesion present on the day of visit ○ may miss seasonal fluctuation of eczema symptoms • Depends on management of eczema, lesions may be absent if the disease is well treated • Captures any eczema lesion, not necessarily atopic dermatitis

3mQ: 3-monthly questionnaires. Please refer to supplementary appendix for details on questionnaire wordings and outcome definition.

Table 2: Eczema severity assessment tools used in the MIS BAIR trial

	Description	Advantage	Limitations
POEM score	The POEM score is calculated using data from the 3mQ *"The following questions relate to how much of [child's name]'s last week has been affected by itchy skin or eczema. Please answer the following questions in relation to the last week* • *How many days has [child's-name]'s skin been a. itchy, b. bleeding, c. weeping or oozing clear fluid, d. cracked, e. flacking off, f. felt dry or rough because of eczema/dry itchy skin?* • *How many nights has [child's-name]'s sleep been disturbed because of eczema/dry itchy skin?* Scoring for each of the 7 criteria listed above: No days = 0, 1-2 days = 1, 3-4 days = 2, 5-6 days = 3, every day = 4 Interpretation: 0-2 = clear, 3-7 = mild, 8-16 = moderate eczema, 17-24 = severe, 25-28 = very severe	• Widely used questionnaire-based tool	• Question on pruritus not reliable for young infants (unable to scratch themselves) • Question on sleep disturbance can be misinterpreted by parent as asking for "overall sleep disturbance" that is frequent in infancy • Depends on management of eczema, score can be low if the disease is well treated • Only captures the previous week's severity • Relatively complex scoring system
SCORAD	Sub-score A. Area: Percentage of the whole body affected Sub-score B. Intensity: On a representative area, an intensity score is given to each of the following signs (none = 0; mild = 1; moderate = 2; severe = 3): a. erythema (redness), b. oedema/papulation (swelling), c. oozing/crusting, d. excoriation (scratching), e. lichenification (skin thickening), f. xerosis (dry skin, assessed in another area where there is no inflammation). Sub-score C. Subjective symptoms: two questions in relation to itching and sleeping in the last week of the participant's life. Each question was given a value from unaffected (0) to worst possible (10). Scoring: $\frac{A}{5} + \frac{7*B}{2} + C$ Interpretation: 0 to 9.9 = clear, 10.0 to 28.9 = mild, 29.0 to 48.9 = moderate, 49.0 to 103 = severe	• Widely used score	• Requires clinic visit and trained staff • Depends on management of eczema, score can be low if the disease is well treated • Only captures the severity on the day of the clinic visit • erythema does not occur with all skin types (e.g. dark skin) • Sub-score C is highly subjective ○ Question on sleep disturbance can be misinterpreted by parent as asking for 'overall sleep disturbance' that is frequent in infancy • Relatively complex scoring system

3mQ: 3-monthly questionnaires; POEM: Patient Oriented Eczema Measure score[10]; SCORAD: SCORing Atopic Dermatitis scoring system.[11]
Please refer to supplementary appendix for details on questionnaire wordings and outcome definition.

Table 3: Participant characteristics

	N*	N (%) or mean (SD)#
Sex, female	1272	630 (50%)
Birth weight, kg	1272	3.4 (0.5)
Gestational age at birth, week	1272	39.3 (1.4)
Vaginal delivery	1272	812 (64%)
Family history of eczema	1270	514 (40%)
Family history of hay fever	1270	836 (66%)
Family history of asthma	1269	615 (48%)
Family history of any atopic disease†	1271	1049 (83%)
Family history of any allergy or atopic disease	1262	1070 (85%)
Family history of any non-food allergy‡	1231	306 (25%)

* Number of participants with data available.

Categorical variables are reported as number (%), continuous variables are reported as mean (standard deviation).

† Any of eczema, hay fever, asthma.

‡ Does not include hay fever.

Figure legends

Figure 1: Distribution of participants fulfilling eczema definitions.

Figure 2: Agreement between eczema definitions

Agreement between eczema diagnostic tools is quantified using kappa coefficients, categorised as previously defined.[12] "3-4 minor UK criteria non-itchy": participant fulfilling 3 to 4 minor criteria for the UK diagnostic tool but non-itchy (not fulfilling the major criterion).

Figure 3: Mean age of onset of eczema symptom in the different subgroups

Heatmap of mean time of first report of any eczema symptom in a 3-monthly questionnaire in subgroups of eczema definitions, presented in the same Venn diagram of Figure 1. Numbers are the mean time of first report within the sub-group, in months (m). Details available in Figure S2 (supplementary appendix).

Figure 4: Mean SCORAD scores in the different subgroups

Heatmap of SCORAD score at the 12-month visit in subgroups of eczema definitions, presented in the same Venn diagram of Figure 1. Numbers are the mean score within the sub-group. Details available in Figure S3 (supplementary appendix).