

DR CARSTEN FLOHR (Orcid ID : 0000-0003-4884-6286)

DR JONATHAN I SILVERBERG (Orcid ID : 0000-0003-3686-7805)

DR KATRINA ABUABARA (Orcid ID : 0000-0002-7736-6946)

Article type : Systematic Review

## **Systematic review of atopic dermatitis disease definition in studies using routinely-collected health data**

**Short Title:** AD definitions in studies using routinely-collected health data

M. P. Dizon<sup>1</sup>, A. M. Yu<sup>2</sup>, R. K. Singh<sup>3</sup>, J. Wan<sup>4</sup>, Mary-Margaret Chren<sup>1</sup>, Carsten Flohr<sup>5</sup>, J. Silverberg<sup>6</sup>, D. J. Margolis<sup>7</sup>, S. M Langan<sup>\*8</sup>, K. Abuabara<sup>1</sup>

\*Co senior author with equal contribution

### **Affiliations:**

1. Program for Clinical Research, Department of Dermatology, University of California, San Francisco School of Medicine, San Francisco, CA, USA.
2. Faculty of Medicine, University of Ottawa, Ottawa, Ontario, Canada.
3. University of California-Los Angeles, David Geffen School of Medicine at UCLA, Los Angeles, CA, USA.
4. Department of Dermatology, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA, USA.
5. Guy's and St Thomas' NHS Foundation Trust, London, United Kingdom.
6. Departments of Dermatology, Preventive Medicine and Medical Social Sciences, Northwestern University Feinberg School of Medicine, Chicago, IL, USA.
7. Department of Biostatistics, Epidemiology and Informatics and the Department of Dermatology, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA, USA.
8. Department of Non-communicable Disease Epidemiology, Faculty of Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, London, United Kingdom.

### **Figures:**

Fig 1. Flow Diagram

### **Tables:**

Table 1. Characteristics of included studies

Table 2. Terminology

Table 3. Prevalence estimates

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/bjd.16340

This article is protected by copyright. All rights reserved.

### **Supplemental Tables/Figures:**

Table S1. Classification system

Table S2. Search strategy

Table S3. List of included studies

**Key words:** atopic dermatitis, eczema, coding, validation, electronic medical records, epidemiology, prevalence

### **Abbreviations used:**

Atopic dermatitis (AD)

Routinely collected data (RCD)

### **Corresponding author:**

Katrina Abuabara, MD

Department of Dermatology

2340 Sutter Street, Room N421

San Francisco, CA 94143-0808

Phone: 415-502-0433; Email: Katrina.abuabara@ucsf.edu

**Funding:** This study was supported in part by the University of California, San Francisco Resource Allocation Program for Trainees (RAPtr). KA is supported by grants from the Dermatology Foundation and National Center for Advancing Translational Sciences of the NIH under Award Number KL2TR001870. Research reported in this publication is solely the responsibility of the authors and do not necessarily represent the official views of the NIH. SML is supported by a Wellcome Senior Clinical fellowship in Science (205039/Z/16/Z). JIS is supported by grants from the AHRQ K12 HS023011 and Dermatology Foundation. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Conflicts of interest:** JW receives funding for research from the National Institutes of Health (NIH), the Dermatology Foundation, and the American Skin Association. DM receives funding for research from the NIH and Valeant, and he is a consultant for Sanofi and GlaxoSmithKline. KA receives funding for research from the NIH, the Dermatology Foundation, and the Robert Wood Johnson Foundation.

**What's already known about this topic?** Increasingly, studies are using routinely collected data to study atopic dermatitis, though it is unclear how AD patients are identified and whether methodological differences could impact study findings.

**What does this study add?** We performed a systematic review of methods for identifying AD patients in studies using routinely collected data, and found variability in methods associated with up to a 3-fold variation in prevalence estimates. We encourage validation of methods and offer suggestions for reporting to allow for meaningful interpretation and comparison of results.

## **ABSTRACT**

**Background:** Routinely collected electronic health data (RCD) obtained for administrative and clinical purposes are increasingly used to study atopic dermatitis (AD). Methods for identifying AD patients in RCD differ, and it is unknown how this might affect study results.

**Objectives:** We sought to evaluate how AD patients have been identified in studies using RCD, to determine whether these methods were validated, and to estimate how the method for identifying AD patients affected variability in prevalence estimates.

**Methods:** We systematically searched PubMed, EMBASE, and Web of Science for studies utilizing RCD that reported on AD as a primary outcome. Studies of localized AD and other types of dermatitis were excluded. The protocol for this review was registered in PROSPERO (CRD42016037968).

**Results:** 59 studies met eligibility criteria. Medical diagnosis codes for inclusion and exclusion, number of occasions of a code, type of provider associated with a code, and prescription data were used to identify AD patients. Only two studies described validation of their methods and no study reported on disease severity. Prevalence estimates ranged from

0.18%-38.33% (median 4.91%) and up to 3-fold variation in prevalence was introduced by differences in the method for identifying AD patients.

**Conclusions:** This systematic review highlights the need for clear reporting of methods for identifying AD patients in routinely collected electronic health data to allow for meaningful interpretation and comparison of results.

## INTRODUCTION

Increasing data suggest atopic dermatitis (AD, also known as atopic eczema or eczema) is a systemic inflammatory disease affecting both children and adults.<sup>1</sup> There is an unmet need for additional research in large, representative populations with longitudinal follow up and data on comorbid conditions. “Routinely collected” electronic health data obtained for administrative and clinical purposes often meet these criteria and are increasingly being used to study the epidemiology, natural history, and association of AD with other diseases.<sup>2</sup> They could include data for clinical management (e.g., primary care databases), health system planning (e.g., health administrative data), documentation of clinical care (e.g., electronic health record data repositories), or epidemiological surveillance (e.g., cancer registries and public health reporting data). Because these data are not generated specifically for research purposes, they require careful validation to ensure accuracy and reproducibility.<sup>2</sup>

Unlike some conditions for which diagnosis may be based on diagnostic tests or lab values easily retrievable from medical records, AD diagnosis is typically based solely on clinical signs and symptoms, and physician assessment is considered the ‘gold standard’.<sup>3-5</sup> Moreover, AD is clinically heterogeneous, with variable morphology, severity, and clinical course; all of which can present challenges to accurately identifying AD patients in routinely-collected health data. It is possible that AD prevalence and severity estimates are influenced

by the method used to identify patients. Therefore, we aimed to provide an overview of AD disease definition in studies using routinely collected data. The primary objectives of this systematic review were to evaluate how AD patients have been identified and how disease severity was defined. We also aimed to determine whether these methods were validated (i.e. whether any information was included about the accuracy of methods for identifying AD), and when applicable, to estimate how AD disease definitions affected the variability in AD prevalence.

## METHODS

The review protocol was registered in PROSPERO, number CRD42016037968 (<http://www.crd.york.ac.uk/PROSPERO>). We followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) and Reporting of studies Conducted Using Observational Routinely-collected Data (RECORD) guidelines, which are an extension of the STROBE guidelines.<sup>2 6 7</sup>

### Types of studies

We included cohort, case-control, and cross-sectional studies using routinely-collected health data reporting on AD as a primary outcome. Studies that examined AD as a predictor of a separate outcome (e.g., cancer diagnosis) were not included. Routinely-collected health data were defined as data collected without specific *a priori* research questions and developed prior to utilization for research.<sup>2</sup> Data sources designed to investigate specific questions about AD or atopic diseases, such as the International Study of Asthma and Allergies in Childhood (ISAAC), birth cohorts of patients with AD, and registries of AD patients were excluded. Studies of localized AD, such as hand eczema, or other types of dermatitis such as contact dermatitis and seborrheic dermatitis were also excluded.

## **Outcome**

Our primary endpoint was the criteria used to identify patients with AD in each study. We also examined whether these criteria were validated (i.e whether any information was included about the accuracy of methods for identifying AD), how disease severity was defined, and the prevalence of AD.

## **Search strategy**

With the help of a professional librarian, we searched MEDLINE via PubMed, EMBASE, and Web of Science for studies indexed until April 10, 2016. Table S2 shows the detailed search strategy. Studies in any language were included. Because the focus of this review was to describe how AD has been defined in the mainstream published literature, we excluded literature reviews, abstracts, conference proceedings, unpublished studies, ongoing studies, and the grey literature (i.e., reports and research disseminated outside of commercial publishing). We cross-referenced review articles and reference lists to ensure completeness.

## **Selection and data extraction**

Three authors (MPD and one of AMY, RKS) performed the study selection independently and in duplicate. Titles and abstracts were screened for inclusion, followed by full text review if abstracts were insufficient to determine whether studies met inclusion and exclusion criteria. Discrepancies were resolved through discussion and consensus with additional authors (KA and SML). For each eligible study, we extracted information on the database used, country of study, study objective, patient demographics, features of algorithms used to identify cases of AD, and prevalence estimates. When possible, missing data, including specific diagnosis codes, were obtained by contacting the study author(s).

## **Data synthesis and analysis**

We described the characteristics of relevant studies. Features of algorithms used to identify cases of AD were tabulated for the included studies, and the median prevalence and variability (interquartile range and ratio of 75<sup>th</sup> to 25<sup>th</sup> percentile) were calculated by subgroup. We also reported the proportion of studies in which these algorithms have been validated and described the methods of validation.

## **Risk of bias**

Systematic reviews often include an assessment of the risk of bias. This involves rating each included study on the methods used for selection of the study groups, comparability of groups, and ascertainment of exposure and outcome.<sup>6,8</sup> These categorizations were not applicable to the objective of our study (i.e. we focused exclusively on how an outcome, AD, was defined), so no risk of bias assessment was performed.

## **RESULTS**

### **Selection and characteristics of studies**

Our search identified 1354 studies. Title and abstract review identified 127 articles for full text review. Of these, 68 were excluded, and 59 met criteria for inclusion (Figure 1). Table 1 shows the characteristics of the included studies. The vast majority of studies (90%) included children; only 6 studies (10%) included only adults. Most studies (81%) included both male and female patients, and 61% were from North America or Europe, while the remainder came from Asia Pacific/Oceania. The included studies were published between 1994 and 2016, while data came from the years 1967-2014. Most studies (58%) were conducted using administrative databases (e.g. insurance databases, birth/death registries or

employment registries). Primary care databases, national patient registers, institutional electronic medical records (EMR), and hybrid databases compiling information from multiple sources were also used.

### **Algorithm features**

We identified four features of algorithms that were used to identify patients with AD: (1) diagnosis codes, (2) number of occasions of a code, (3) type of provider associated with a code, and (4) prescription data.

(1) *Diagnosis codes*: Multiple codes from multiple coding systems, including ICD9/10 and READ/OXMIS were used. We grouped studies into those that used codes specifically for AD, and those that used codes referring to a broader group of dermatitis-related disorders (including, for example contact dermatitis, and eczema not otherwise specified (Table S1). When we compared the terminology used in each published study (i.e. “atopic dermatitis” versus “eczema”) to the types of codes we found imperfect overlap (Table 2). Only one study directly incorporated codes for exclusionary diagnoses (“diaper or napkin rash” and “contact dermatitis and other eczema”) as part of their algorithm for AD.<sup>9</sup>

(2) *Number of occasions of a code*: A total of 10 studies used algorithms that required multiple instances of codes to identify patients as having AD. Two studies<sup>10 11</sup> used unique algorithms which identified cases if they had at least 1 inpatient claim or 2 outpatient claims associated with diagnosis codes for AD. Nine studies specified a minimum of two occasions associated with AD,<sup>12-20</sup> while one study required three medical visits coded for AD to identify patients.<sup>21</sup>

(3) *Type of provider associated with a code*: A total of 7 studies, all of which were conducted in Taiwan using the National Health Insurance Research Database, specified the type of provider required to enter diagnosis to define a case of AD. Three studies required a

diagnosis by a dermatologist,<sup>14 17 22</sup> three studies required a diagnosis by a dermatologist or a pediatrician,<sup>15 23 24</sup> and one study required a diagnosis by a “specialist.”<sup>18</sup>

(4) *Prescription data*: Five studies included medication prescriptions in the algorithms used to identify patients with AD. One study required a patient to have a diagnosis code and a recorded prescription of a treatment for AD (e.g., emollients, topical corticosteroids, and topical calcineurin inhibitors).<sup>25</sup> Two studies specified either a diagnosis code or by prescription records for calcineurin inhibitors or topical corticosteroids,<sup>26 27</sup> and one study used prescription codes for either topical corticosteroids or topical calcineurin inhibitors to identify patients with eczema.<sup>28</sup>

### **Validation Information**

Only two studies described methods to validate the algorithms used to identify cases of AD. The first study assessed the validity of ICD-9-CM codes by conducting a detailed chart review of randomly selected visits of 1,000 patients.<sup>29</sup> The ICD-9-CM diagnosis was confirmed in 93% of cases, however, the study included a range of skin diseases and was not limited to AD. Criteria used in chart review to confirm diagnoses were not described. The second study calculated the positive predictive value (PPV) of filled prescriptions of topical corticosteroids or immunosuppressants for identifying patients with an “umbrella diagnosis” of dermatitis or eczema based on diagnosis codes, and found PPVs 82% and 45%, respectively.<sup>30</sup> The study described specific clinical criteria used to validate asthma diagnoses in a subset of pediatric patients through review of medical records, however, coded diagnoses for dermatitis and eczema were not evaluated against clinical criteria.

## Severity Information

None of the studies included in the review reported on the severity of AD within the study population.

## Variation in prevalence estimates

Of the 59 included studies, 40 (68%) provided a prevalence estimate of AD, which ranged from 0.18% to 32.49%. Estimates varied by features used to identify AD patients, study characteristics, and the duration of time used to calculate the period prevalence. Of note, the variation in prevalence introduced by differences in the methods used to identify patients was similar in magnitude to the variation in prevalence introduced by study characteristics (e.g. prevalence in studies where the algorithm included prescription data was 16.9% vs. no prescription data was 4.5%; prevalence among studies including children only was 11% vs. adults and/or children was 4%, Table 3).

Studies that used only diagnosis codes specific to AD to identify patients found a lower median prevalence than studies that used more general dermatitis codes (4.3% vs 11.5%), and the amount of variability was similar (ratio of 75<sup>th</sup> to 25<sup>th</sup> percentile 3.1 vs 3.3). Studies that required patients to have specified diagnosis codes on multiple occasions found a lower median prevalence (4.3% vs 5.6%) and less variability (ratio of 75<sup>th</sup> to 25<sup>th</sup> percentile 1.5 vs 5.5) than studies that only required one instance of a code. The same was true of studies that required diagnosis by a specialist (4.5% vs 4.9%, ratio of 75<sup>th</sup> to 25<sup>th</sup> percentile 1.5 vs 4.6). Finally, studies that used prescription data to identify patients found a higher median prevalence (16.9%, vs 4.5%), possibly due to misclassification of patients receiving medications for other conditions.

Prevalence estimates also varied by region and age group with the median prevalence higher in studies that included only children. The duration of period prevalence ranged from

1-39 years; among the 9 estimates that calculated a one-year prevalence; the median prevalence was lower (2.8%) than in studies that calculated a multi-year period prevalence (8.9%).

## DISCUSSION

This review demonstrates variability in the way AD patients are identified in studies using routinely collected data. It also highlights a lack of standardization in terminology, validation studies, and information on disease severity, which are all crucial to allow for comparison of study results. These issues are not specific to AD, however, they are of particular importance in AD because it is a common condition and misclassification of even a relatively small percentage of patients could result in large absolute errors.

### Terminology

Much has been written about the inconsistent use of terminology in allergic disorders, and efforts are underway to improve the classification system and standardize terminology used in coding<sup>31-34</sup>. To ensure we captured all relevant studies, we used multiple terms in our search, including “atopic dermatitis”, “dermatitis”, “eczema” and other variants thereof (Table S2). Inconsistent use of terminology can be seen in Table 2, with some studies of “eczema” including only patients with AD-specific diagnosis codes and some studies of “atopic dermatitis” including patients with diagnosis codes for other types of inflammatory skin conditions such as contact dermatitis. We found that use of more broad diagnosis codes increased the median prevalence from 4.3% to 11.5%, Table 3). Such ‘lumping together’ of different disease entities could inflate AD prevalence estimates, though using more limited AD code sets might underestimate the true prevalence. For example, a recent validation study using EMR data found that 42% of patients with the nonspecific diagnosis code of

692.9 and no AD-specific code of 691.8 had a final diagnosis of AD after chart review.<sup>35</sup> If a study focuses on more than one type of dermatitis, authors should clearly delineate how each condition was defined, including which codes were used. Ideally, studies examining multiple types of dermatitis would report estimates separately by sub-groups to facilitate comparison with existing literature.

### **Validation studies**

This systematic review highlights the frequent use of non-validated algorithms to identify AD patients in routinely collected data; only 2/59 studies described any attempt to validate algorithms used. Validation research is a high priority to ensure patients are accurately identified and avoid misclassification bias,<sup>36</sup> and since completion of our search two new validation studies of AD using routinely collected data have been published.<sup>35 37</sup> Both highlighted the potential magnitude of misclassification bias, even when using physician-defined codes. For example, using a single code for “AD/eczema”, rather than one of 5 AD-related codes in a primary care database from the UK could result in a 50% reduction in prevalence estimates.<sup>38</sup>

The performance of coding algorithms for identifying AD patients is inherently context-specific. For example, the performance of an algorithm may depend on the baseline prevalence and the way in which diagnosis codes and pharmacy codes are entered in a given setting. Moreover, in any given context, the choice of coding algorithms may be related to the goals of the study, as there is often a trade-off between maximizing the number of true positives and reducing the number of false positives, and the value of optimizing sensitivity or specificity may depend on whether the study is aiming to identify all possible cases or to identify only those with definite disease. Therefore, each study should discuss evidence for the validity of the methods used to identify AD patients.<sup>2</sup> When possible, researchers also

may consider showing how changes in their definition of AD could affect their estimates. Additional research is necessary to understand the generalizability of coding algorithms and the extent to which these might be standardized across settings.

### **Severity data**

A secondary objective of this systematic review was to examine methods used to describe AD severity in studies using RCD, however, none of the studies meeting our inclusion criteria evaluated disease severity. Current approaches to measuring AD severity are often complex and not routinely documented in the medical record<sup>39</sup> or are not standardized,<sup>40</sup> and therefore are difficult to use in RCD studies. Treatment data and/or frequency of health care visits has been used to define severity in studies of psoriasis and asthma using RCD and may be applicable to AD:<sup>41-44</sup> such an approach, if applied to AD, should be carefully validated.

### **Strengths and limitations**

Strengths of this systematic review include a predefined and registered protocol and adherence to reporting standards.<sup>6</sup> We included 59 studies of varying study designs from a variety of settings to show all of the ways AD patients have been defined in the literature to date. We included only studies with AD as a primary outcome because we were primarily interested in whether differences in AD definitions would affect prevalence estimates, but were unable to synthesize these estimates using meta-analysis due to lack of standardized reporting of prevalence. Nonetheless, we include unadjusted median and interquartile prevalence ranges, which demonstrate variation in prevalence estimates.

## **Future research needs and recommendations for reporting**

Clinicians should be aware that estimates from studies using routinely collected data may vary depending on the algorithms used to identify patients, and should be wary of studies that do not provide data on the validity of these measures. The international Harmonizing Outcome Measures for Eczema (HOME) initiative was founded in response to the lack of standardization and validation of methods used to measure outcomes in randomized clinical trials, and the initiative has resulted in multiple publications suggesting standardized methods of measurement and reporting.<sup>45-49</sup> Similar international efforts are needed for questionnaire-based studies and studies of AD using RCD. In the meantime, we encourage authors to clearly report on their methods including: specific codes used to identify patients or exclude patients, whether there was a minimum number of codes or visits required, whether there were any restrictions on type of provider associated with the code or visit, and whether prescription data were used to identify patients. In addition, whenever possible, we encourage authors to report on the annual period prevalence of visits and/or prescriptions for AD by age to enable comparison across studies. All studies should include information on the validity of the algorithm used in their particular locale and practice setting.

## **ACKNOWLEDGEMENTS**

The authors would like to thank Elaine Allen for her statistical advice regarding study design and analysis, Evans Whitaker for helping us developing the search strategy, and Jochen Schmitt for his assistance in reviewing and extracting data from German-language articles.

## REFERENCES

1. Brunner PM, Silverberg JI, Guttman-Yassky E, et al. Increasing Comorbidities Suggest that Atopic Dermatitis Is a Systemic Disorder. *J Invest Dermatol* 2017;**137**(1):18-25.
2. Benchimol EI, Smeeth L, Guttman A, et al. [The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) statement]. *Zeitschrift fur Evidenz, Fortbildung und Qualitat im Gesundheitswesen* 2016;**115-116**:33-48.
3. Hanifin JM RG. Diagnostic features of atopic dermatitis. *Acta Derm Venereol* 1980;**92** (suppl):44-47.
4. 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.
5. Eichenfield LF, Tom WL, Chamlin SL, et al. Guidelines of care for the management of atopic dermatitis: section 1. Diagnosis and assessment of atopic dermatitis. *J Am Acad Dermatol* 2014;**70**(2):338-51.
6. Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. *BMJ* 2009;**339**:b2700.
7. Vandembroucke JP, von Elm E, Altman DG, et al. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration. *International journal of surgery* 2014;**12**(12):1500-24.
8. Wells GA SB, O'Connell D, Peterson J, Welsch V, Losos M, Tugwell P. The Newcastle-Ottawa Scale (NOS) for assessing the quality of non-randomized studies in meta-analyses. The Ottawa Hospital. . Secondary The Newcastle-Ottawa Scale (NOS) for assessing the quality of non-randomized studies in meta-analyses. The Ottawa Hospital. . [http://www.ohri.ca/programs/clinical\\_epidemiology/oxford.asp](http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp)
9. Sun H-L, Yeh C-J, Ku M-S, et al. Coexistence of allergic diseases: patterns and frequencies. *Allergy and Asthma Proceedings: The Official Journal of Regional and State Allergy Societies* 2012;**33**:e1-4.
10. Wang J-Y, Liu L-F, Chen C-Y, et al. Acetaminophen and/or antibiotic use in early life and the development of childhood allergic diseases. *International Journal of Epidemiology* 2013;**42**:1087-99.
11. Hua T-C, Hwang C-Y, Chen Y-J, et al. The natural course of early-onset atopic dermatitis in Taiwan: a population-based cohort study. *The British Journal of Dermatology* 2014;**170**:130-35.
12. Schmitt J, Schmitt NM, Kirch W, et al. Significance of atopic dermatitis in outpatient medical care. Analysis of health care data from Saxony. *Hautarzt* 2009;**60**:320-27.
13. Schmitt J. Medical care and health-economic relevance of atopic eczema. *Allergologie* 2010;**33**:279-88.
14. Chung S-D, Keller JJ, Lin H-C. Association of erectile dysfunction with atopic dermatitis: a population-based case-control study. *The Journal of Sexual Medicine* 2012;**9**:679-85.
15. Chen M-H, Su T-P, Chen Y-S, et al. Attention deficit hyperactivity disorder, tic disorder, and allergy: is there a link? A nationwide population-based study. *Journal of Child Psychology and Psychiatry, and Allied Disciplines* 2013;**54**:545-51.
16. Hwang C-Y, Hwang Y-Y, Chen Y-J, et al. Atopic diathesis in patients with Kawasaki disease. *The Journal of Pediatrics* 2013;**163**:811-15.
17. Tien K-J, Chou C-W, Lee S-Y, et al. Obstructive sleep apnea and the risk of atopic dermatitis: a population-based case control study. *PloS One* 2014;**9**:e89656.

- Accepted Article
18. Chen YT, Chen YJ, Hwang CY, et al. Comorbidity profiles in association with vitiligo: a nationwide population-based study in Taiwan. *Journal of the European Academy of Dermatology and Venereology: JEADV* 2015;**29**:1362-69.
  19. Hirsch AG, Yan XS, Sundaresan AS, et al. Five-year risk of incident disease following a diagnosis of chronic rhinosinusitis. *Allergy* 2015;**70**:1613-21.
  20. Wei C-C, Lin C-L, Shen T-C, et al. Occurrence of common allergic diseases in children with idiopathic nephrotic syndrome. *Journal of Epidemiology / Japan Epidemiological Association* 2015;**25**:370-77.
  21. Tsai J-D, Chang S-N, Mou C-H, et al. Association between atopic diseases and attention-deficit/hyperactivity disorder in childhood: a population-based case-control study. *Annals of Epidemiology* 2013;**23**:185-88.
  22. Chu S-Y, Chen Y-J, Tseng W-C, et al. Comorbidity profiles among patients with alopecia areata: the importance of onset age, a nationwide population-based study. *Journal of the American Academy of Dermatology* 2011;**65**:949-56.
  23. Chen M-H, Su T-P, Chen Y-S, et al. Comorbidity of Allergic and Autoimmune Diseases Among Patients With ADHD: A Nationwide Population-Based Study. *Journal of Attention Disorders* 2013.
  24. Woon PY, Chang WC, Liang C-C, et al. Increased risk of atopic dermatitis in preschool children with kawasaki disease: a population-based study in taiwan. *Evidence-Based Complementary and Alternative Medicine: eCAM* 2013;**2013**:605123.
  25. Schmitt J, Schmitt NM, Kirch W, et al. Early exposure to antibiotics and infections and the incidence of atopic eczema: A population-based cohort study. *Pediatric Allergy and Immunology* 2010;**21**:292-300.
  26. Henriksen L, Simonsen J, Haerskjold A, et al. Incidence rates of atopic dermatitis, asthma, and allergic rhinoconjunctivitis in Danish and Swedish children. *The Journal of Allergy and Clinical Immunology* 2015;**136**:360-66.e2.
  27. Egeberg A, Andersen YMF, Gislason G, et al. Neonatal risk factors of atopic dermatitis in Denmark - results from a nationwide register based study. *Pediatric Allergy and Immunology: Official Publication of the European Society of Pediatric Allergy and Immunology* 2016.
  28. Mulder B, Schuiling-Veninga CCM, Bos HJ, et al. Prenatal exposure to acid-suppressive drugs and the risk of allergic diseases in the offspring: a cohort study. *Clinical and Experimental Allergy: Journal of the British Society for Allergy and Clinical Immunology* 2014;**44**:261-69.
  29. Henderson MD, Abboud J, Cogan CM, et al. Skin-of-Color Epidemiology: A Report of the Most Common Skin Conditions by Race. *Pediatric Dermatology* 2012;**29**:584-89.
  30. Ortvqvist AK, Lundholm C, Wettermark B, et al. Validation of asthma and eczema in population-based Swedish drug and patient registers. *Pharmacoepidemiology and Drug Safety* 2013;**22**:850-60.
  31. Johansson SG, Bieber T, Dahl R, et al. Revised nomenclature for allergy for global use: Report of the Nomenclature Review Committee of the World Allergy Organization, October 2003. *J Allergy Clin Immunol* 2004;**113**(5):832-6.
  32. Demoly P, Tanno LK, Akdis CA, et al. Global classification and coding of hypersensitivity diseases - An EAACI - WAO survey, strategic paper and review. *Allergy* 2014;**69**(5):559-70.
  33. Tanno LK, Calderon MA, Goldberg BJ, et al. Categorization of allergic disorders in the new World Health Organization International Classification of Diseases. *Clin Transl Allergy* 2014;**4**:42.

- Accepted Article
34. Kantor R, Thyssen JP, Paller AS, et al. Atopic dermatitis, atopic eczema, or eczema? A systematic review, meta-analysis, and recommendation for uniform use of 'atopic dermatitis'. *Allergy* 2016;**71**(10):1480-5.
  35. Hsu DY, Dalal P, Sable KA, et al. Validation of international classification of disease ninth revision codes for atopic dermatitis. *Allergy* 2016.
  36. De Coster C, Quan H, Finlayson A, et al. Identifying priorities in methodological research using ICD-9-CM and ICD-10 administrative data: report from an international consortium. *BMC Health Serv Res* 2006;**6**:77.
  37. 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.
  38. Abuabara K, Hoffstad O, Troxel AB, et al. Patterns and predictors of atopic dermatitis disease control past childhood: An observational cohort study. *J Allergy Clin Immunol* 2017.
  39. Rehal B, Armstrong A. Health Outcome Measures in Atopic Dermatitis: A Systematic Review of Trends in Disease Severity and Quality-of-Life Instruments 1985–2010. *PLOS ONE* 2011;**6**(4):e17520.
  40. Futamura M, Leshem YA, Thomas KS, et al. A systematic review of Investigator Global Assessment (IGA) in atopic dermatitis (AD) trials: Many options, no standards. *Journal of the American Academy of Dermatology* 2016;**74**:288-94.
  41. Fuhlbrigge AL, Carey VJ, Finkelstein JA, et al. Validity of the HEDIS criteria to identify children with persistent asthma and sustained high utilization. *The American Journal of Managed Care* 2005;**11**:325-30.
  42. Bennett AV, Lozano P, Richardson LP, et al. Identifying high-risk asthma with utilization data: a revised HEDIS definition. *The American Journal of Managed Care* 2008;**14**:450-56.
  43. Michael Schatz MD, Robert S. Zeiger MD, Su-Jau T. Yang P, et al. Persistent Asthma Defined Using HEDIS Versus Survey Criteria. *American Journal of Managed Care* 2010;**16**.
  44. Seminara NM, Abuabara K, Shin DB, et al. Validity of The Health Improvement Network (THIN) for the study of psoriasis. *Br J Dermatol* 2011;**164**(3):602-9.
  45. Schmitt J, Williams H, Group HD. Harmonising Outcome Measures for Eczema (HOME). Report from the First International Consensus Meeting (HOME 1), 24 July 2010, Munich, Germany. *The British Journal of Dermatology* 2010;**163**:1166-68.
  46. Schmitt J, Spuls P, Boers M, et al. Towards global consensus on outcome measures for atopic eczema research: results of the HOME II meeting. *Allergy* 2012;**67**:1111-17.
  47. Chalmers JR, Schmitt J, Apfelbacher C, et al. Report from the third international consensus meeting to harmonise core outcome measures for atopic eczema/dermatitis clinical trials (HOME). *British Journal of Dermatology* 2014;**171**:1318-25.
  48. Schmitt J, Apfelbacher C, Spuls PI, et al. The Harmonizing Outcome Measures for Eczema (HOME) roadmap: a methodological framework to develop core sets of outcome measurements in dermatology. *The Journal of Investigative Dermatology* 2015;**135**:24-30.
  49. Chalmers JR, Simpson E, Apfelbacher CJ, et al. Report from the fourth international consensus meeting to harmonize core outcome measures for atopic eczema/dermatitis clinical trials (HOME initiative). *The British Journal of Dermatology* 2016;**175**:69-79.

**Table 1. Characteristics of the included studies**

|                                                                  | N  | %   |
|------------------------------------------------------------------|----|-----|
| Age groups included in study sample                              |    |     |
| Children                                                         | 32 | 54% |
| Adults                                                           | 6  | 10% |
| Both                                                             | 21 | 36% |
| Sex of patients included in study sample                         |    |     |
| Male only                                                        | 4  | 7%  |
| Female only                                                      | 0  | 0%  |
| Both male and female                                             | 48 | 81% |
| Not reported                                                     | 7  | 12% |
| Countries represented in studies                                 |    |     |
| Taiwan                                                           | 19 |     |
| South Korea                                                      | 2  |     |
| USA                                                              | 7  |     |
| Canada                                                           | 1  |     |
| Sweden                                                           | 5  |     |
| Denmark                                                          | 3  |     |
| Norway                                                           | 1  |     |
| Finland                                                          | 1  |     |
| UK                                                               | 11 |     |
| The Netherlands                                                  | 4  |     |
| Germany                                                          | 4  |     |
| Australia                                                        | 2  |     |
| Type of Routinely Collected Data*                                |    |     |
| National patient register                                        | 3  | 5%  |
| Administrative databases<br>(insurance, birth/death, employment) | 34 | 58% |
| Primary care databases                                           | 15 | 25% |
| Institution-specific electronic<br>medical record                | 3  | 5%  |
| Hybrid*                                                          | 4  | 7%  |
| Provided an estimate of AD prevalence                            |    |     |
| Yes                                                              | 40 | 68% |
| No                                                               | 19 | 32% |

**Notes:** \*Hybrid datasets included patient data from sources spanning multiple categories (i.e., both primary care databases and administrative databases). National patient registers include hospital records in countries with government-funded universal healthcare. They are not specific to insurance claims, prescriptions, nor are they limited to primary care.

**Table 2. Comparison of manuscript terminology and medical codes**

| Terms used in manuscript | Total | AD codes only* |      | AD plus other dermatitis codes* |      |
|--------------------------|-------|----------------|------|---------------------------------|------|
|                          |       | N              | %    | N                               | %    |
| Atopic dermatitis        | 31    | 26             | 84%  | 5                               | 16%  |
| Atopic dermatitis/eczema | 3     | 0              | 0%   | 3                               | 100% |
| Atopic eczema            | 5     | 5              | 100% | 0                               | 0%   |
| Dermatitis/eczema        | 3     | 0              | 0%   | 3                               | 100% |
| Eczema                   | 17    | 4              | 24%  | 13                              | 76%  |

\* Full listing of AD and AD plus other codes included in Table S1

**Table 3. Prevalence estimates by subgroup**

|                                                                 | # of studies | # of prevalence estimates | Median Prevalence | IQR (%)     | Ratio of 75 <sup>th</sup> to 25 <sup>th</sup> Percentile |
|-----------------------------------------------------------------|--------------|---------------------------|-------------------|-------------|----------------------------------------------------------|
| <b>Overall</b>                                                  | 59           | 40                        | 4.91%             | 2.64-11.51  | 4.36                                                     |
| <b>Features of algorithms used to identify AD patients</b>      |              |                           |                   |             |                                                          |
| Diagnosis code category                                         |              |                           |                   |             |                                                          |
| Limited to AD                                                   | 35           | 26                        | 4.30%             | 2.14-6.70   | 3.13                                                     |
| AD plus others                                                  | 23           | 13                        | 11.53%            | 9.20-30.23  | 3.29                                                     |
| Algorithm included number of visits/codes                       |              |                           |                   |             |                                                          |
| Required multiple occasions                                     | 8            | 8                         | 4.27%             | 3.85-5.61   | 1.46                                                     |
| No                                                              | 50           | 31                        | 5.60%             | 2.24-12.37  | 5.52                                                     |
| Algorithm specified the type of provider associated with a code |              |                           |                   |             |                                                          |
| Yes                                                             | 5            | 5                         | 4.53%             | 4.39-6.70   | 1.53                                                     |
| No                                                              | 53           | 34                        | 4.91%             | 2.50-11.50  | 4.60                                                     |
| Algorithm included prescription data                            |              |                           |                   |             |                                                          |
| Yes                                                             | 5            | 3                         | 16.93%            | 12.37-32.49 | 2.63                                                     |
| No                                                              | 54           | 37                        | 4.53%             | 2.50-10.94  | 4.38                                                     |
| <b>Study characteristics</b>                                    |              |                           |                   |             |                                                          |
| Age of patient population                                       |              |                           |                   |             |                                                          |
| Children only                                                   | 32           | 19                        | 10.94%            | 4.69-19.00  | 4.05                                                     |
| Adults +/- children                                             | 27           | 21                        | 3.80%             | 2.14-5.60   | 2.62                                                     |
| Continent                                                       |              |                           |                   |             |                                                          |
| Europe                                                          | 28           | 21                        | 10.35%            | 3.80-16.93  | 4.46                                                     |
| North America                                                   | 8            | 4                         | 6.05%             | 2.85-9.05   | 3.18                                                     |
| Asia                                                            | 21           | 15                        | 4.39%             | 2.21-6.70   | 3.03                                                     |
| Type of routinely collected data                                |              |                           |                   |             |                                                          |
| National patient register                                       | 3            | 1                         | 12.37%            |             |                                                          |
| Administrative database                                         | 34           | 26                        | 4.27%             | 2.24-6.70   | 2.99                                                     |
| Primary care database                                           | 15           | 11                        | 13.24%            | 2.79-31.40  | 11.25                                                    |
| Hybrid                                                          | 4            | 2                         | 6.81%             | 4.42-9.20   | 2.08                                                     |

| Terminology                  |    |    |        |            |      |
|------------------------------|----|----|--------|------------|------|
| Atopic dermatitis            | 31 | 21 | 4.69%  | 2.24-10.94 | 4.88 |
| Atopic dermatitis/eczema     | 3  | 2  | 2.85%  | 2.50-3.20  | 1.28 |
| Atopic eczema                | 5  | 4  | 3.85%  | 2.25-7.13  | 3.17 |
| Eczema                       | 17 | 13 | 11.50% | 4.42-30.23 | 6.84 |
| Prevalence calculation       |    |    |        |            |      |
| Time period                  |    |    |        |            |      |
| One-year period prevalence   | 9  | 9  | 2.79%  | 2.24-3.43  | 1.53 |
| Multi-year period prevalence | 29 | 29 | 8.90%  | 4.15-13.21 | 3.18 |

Notes: IQR Interquartile Range

