BMJ Open

Machine learning-derived phenotypic trajectories of asthma and allergy in children and adolescents: protocol for a systematic review

Daniil Lisik [iD],[1] Gregorio Paolo Milani [iD],[2,3] Michael Salisu [iD],[1] Saliha Selin Özuygur Ermis [iD],[1] Emma Goksör [iD],[4] Rani Basna,[1,5] Göran Wennergren [iD],[1,4] Hannu Kankaanranta [iD],[1,6] Bright I Nwaru [iD][1,7]

For numbered affiliations see end of article.

Correspondence to
Mr Daniil Lisik;
daniil.lisik@gmail.com

ABSTRACT

Introduction Development of asthma and allergies in childhood/adolescence commonly follows a sequential progression termed the 'atopic march'. Recent reports indicate, however, that these diseases are composed of multiple distinct phenotypes, with possibly differential trajectories. We aim to synthesise the current literature in the field of machine learning-based trajectory studies of asthma/allergies in children and adolescents, summarising the frequency, characteristics and associated risk factors and outcomes of identified trajectories and indicating potential directions for subsequent research in replicability, pathophysiology, risk stratification and personalised management. Furthermore, methodological approaches and quality will be critically appraised, highlighting trends, limitations and future perspectives.

Methods and analyses 10 databases (CAB Direct, CINAHL, Embase, Google Scholar, PsycInfo, PubMed, Scopus, Web of Science, WHO Global Index Medicus and WorldCat Dissertations and Theses) will be searched for observational studies (including conference abstracts and grey literature) from the last 10 years (2013–2023) without restriction by language. Screening, data extraction and assessment of quality and risk of bias (using a custom-developed tool) will be performed independently in pairs. The characteristics of the derived trajectories will be narratively synthesised, tabulated and visualised in figures. Risk factors and outcomes associated with the trajectories will be summarised and pooled estimates from comparable numerical data produced through random-effects meta-analysis. Methodological approaches will be narratively synthesised and presented in tabulated form and figure to visualise trends.

Ethics and dissemination Ethical approval is not warranted as no patient-level data will be used. The findings will be published in an international peer-reviewed journal.

PROSPERO registration number CRD42023441691.

INTRODUCTION

Asthma and allergic diseases, such as atopic dermatitis, allergic rhinitis and food allergy, are among the most common non-communicable paediatric diseases and constitute a substantial public health burden. Prevalence varies widely across regions, but globally, about 10% report having ever had asthma or eczema by the age of 13–14 years, while around 15% report of ever having had hay fever.[1] Food allergy, in turn, is reported by roughly 5% of children and adolescents.[2,3] Often, these diseases develop in a sequential progression, termed the 'atopic march', beginning with atopic dermatitis in infancy, followed by food allergy, asthma and allergic rhinitis.[4–6] However, recent studies have highlighted a substantial heterogeneity in the trajectories of allergic diseases, both in terms of composition, sequential order and timing.[5,7] It has furthermore been suggested that the observed progressions may not in fact be trajectories per se, but rather a manifestation of comorbidities occurring more often in certain individuals at certain ages.[8] Underlying risk factors have also been demonstrated to be differentially associated with different disease trajectories. For example, breastfeeding has been found to be protective against early transient wheezing,

STRENGTHS AND LIMITATIONS OF THIS STUDY

⇒ 10 databases, including of grey literature, will be searched using exhaustive queries with no limitation by language to encompass all relevant literature.

⇒ Study quality and risk of bias will be assessed thoroughly through a form based on an in-depth review and compilation of related guidelines, checklists and quality assessment tools.

⇒ Two reviewers will independently perform screening, data extraction and quality assessment, minimising the risk of systematic/non-systematic bias and error.

⇒ The explorative nature and data of the investigated literature will limit comparative analysis of computational methodology and characteristics/frequency of the derived trajectories.

but the association appears to be non-significant for early-persistent and intermediate/late-onset wheezing.[9]

Facilitated by the increase of longitudinal clinical data,[10] a substantial number of studies characterising trajectories of asthma and allergic diseases have been published, including those using machine learning models.[11–18] The historically dominant hypothesis-driven approach of disease characterisation has commonly been based on the clinical presentation of patients and is susceptible to bias,[19–21] while data-driven approaches, in contrast, have the potential to explore large datasets more effectively and identify novel latent patterns.[22] Phenotypic trajectories, by capturing dynamics across multiple time points, also enable deeper understanding of disease pathophysiology, optimisation of care, as well as development of prediction models.[10] Although systematic reviews summarising phenotype discoveries in individual diseases such as asthma (including limited findings on phenotypic trajectories)[21 23] and risk factors of phenotypic trajectories, for example, wheezing[24] have been published, the present work will be the first to focus on machine learning-derived phenotypic trajectories in children/adolescents and encompassing a broad spectrum of allergic diseases as well as asthma, thereby providing a comprehensive overview of how these diseases develop during the first 18 years of life.

The primary aim of this systematic review will be to summarise the childhood/adolescence trajectories of asthma and/or allergic disease that have been identified and their characteristics (including with the use of meta-analysis) and frequency. The secondary aim will be to summarise variables and computational approaches used to derive these trajectories, as well as to synthesise the risk factors and outcomes associated with the derived trajectories (including with the use of meta-analysis).

METHODS

This protocol has been outlined in accordance with the Preferred Reporting Items for Systematic Review and Meta-Analysis protocol (PRISMA-P)[25] guidelines (completed checklist can be found in online supplemental table 1). The final report will be written in accordance with the PRISMA[26] and the Meta-analysis Of Observational Studies in Epidemiology[27] reporting guidelines. In addition, the protocol has been prospectively registered in the international prospective register of systematic reviews (PROSPERO).

Eligibility criteria

The following studies will be considered for inclusion:

► *Study design*: primary longitudinal observational studies in which trajectory-defining data are available from at least two time points in the same subject, with at least 1 year from first to last time point.
► *Population*: children and adolescents (up to 18 years old (ie, trajectory-defining data/follow-up no later than until the age of 18 years)) from population-representative samples. In studies where trajectory-defining data extends beyond the age of 18 years,

but there is possibility to extract any useful trajectory characteristics or associated risk factors/outcomes up until the age of 18 years, the study in question will be eligible
► *Objective*: utilisation of machine learning approaches (any data-driven method in which investigated subjects are classified into subgroups/trajectories by an algorithm) to identify and characterise (either through self-report/parental report, clinical assessment/measurement/diagnosis or medical records (from registers)) trajectories (subtyping by temporal data) of asthma (including recurrent episodes of wheezing) and/or allergies (including atopic dermatitis, allergic rhinitis/conjunctivitis/rhinoconjunctivitis, atopic dermatitis and food allergy, as well as (indirect) measurements of allergy, such as allergic sensitisation).

There will be no restriction on sample size. Due to the large and increasing number of studies, particularly in recent years, and the fact that studies commonly employ methods built on previous advancements, we will restrict our searches to studies published in the last 10 years (from 1 January 2013 until the date of respective database search). This will also ensure that the findings reflect recent methodological trends. Studies of any publication status will be considered (relevant articles under embargo will be noted but not assessed further with data extraction, narrative synthesis, quality assessment and the like). Likewise, relevant conference abstracts and abstracts without a full text will be noted but not assessed further. Relevant letters to the editor will be included and synthesised as far as possible as full-length articles. There will be no restriction based on language. Non-English articles will be translated using Google Translate.[28] Reviews (including systematic reviews) will not be included, but relevant reviews will be screened for relevant literature. Finally, the reference lists of included studies will be screened for additional relevant literature.

Search strategy and data sources

CAB Direct (including CAB Abstracts and Global Health), CINAHL, Embase, Google Scholar, PubMed, Scopus, Web of Science (including KCI and SciELO) and WHO Global Index Medicus (including AIM (Africa), IMEMR (Eastern Mediterranean), IMSEAR ?(South-East Asia), LILACS (Americas) and WPRIM (Western Pacific)) will be searched using exhaustive queries to capture all relevant literature. Likewise, PsycInfo and WorldCat dissertations and theses will be searched for grey literature. Given the indexing nature of Google Scholar, only the first 300 hits will be retrieved.[29] The search queries were adapted to the syntax of each database. Likewise, the search queries were modified based on character limit and on the existence/nomenclature of subject headings, filters and the like. The search queries were developed through pilot searches on PubMed in September, 2023 (during which additional relevant keywords were identified and the search queries iteratively refined) and consist of three blocks ('Asthma and allergies', 'Subgrouping and

trajectory modelling techniques' and 'Age-related inclusion terms', each comprised of 'OR' Boolean operator-separated search terms) concatenated with the 'AND' Boolean operator. Where possible and the number of studies exceed 1000 (arbitrary threshold above which substantial benefit is given by limitation of records), a filter was added to exclude adult-only studies. Finally, search results were limited to those published in the last 10 years (from 1 January 2013 until the date of respective database search), where possible through an additional block in the search query. Details of the final search queries are presented in online supplemental table 2A–J.

De-duplication and screening

Records retrieved from the searches will be imported to EndNote V.21 (Clarivate Analytics, 2023) for semi-automated de-duplication, following a method proposed by Bramer et al.[30] The de-duplicated records will subsequently be screened by pairs of reviewers (DL and GM, DL and MS, and DL and SSÖE) working independently using the Rayyan (https://rayyan.ai) web platform. Screening will be performed in two steps. In the first step, screening will be based on title and abstract, while the second step will consist of full-text assessment. Both steps will be performed in a double-blind fashion, with each reviewer independently evaluating every record for eligibility. Exclusion of records will be done according to the following order: (1) no abstract and no full text; (2) non-original article (ie, duplicate); (3) wrong study design; (4) wrong objective (including the exclusive use of non-machine learning methods, such as by manually defining trajectories) and (5) wrong population.[31] Following completion, the screening decisions will be unblinded for the other reviewer. Disagreements will be resolved through discussion and arbitration by the principal investigator (PI, BIN), if necessary. In the first step, records that are clearly eligible and records for which there is uncertainty of eligibility will be included to the second step, and cause of exclusion will not be documented. In the second step, records that are eligible will get included in the final manuscript, and each exclusion will be documented and reported (including cause of exclusion) in the supplementary material of the final manuscript (structure shown in online supplemental table 3). A PRISMA flow diagram will be produced to illustrate the screening process in the final manuscript.

Data extraction

Data extraction will be performed independently in a double-blind fashion by pairs of reviewers (DL and GM and DL and MS), using a Microsoft Excel (Microsoft Corp., 2023) data extraction form (online supplemental file), prospectively piloted and modified by DL, BIN and RB based on relevant articles identified during the PubMed pilot searches. Following completion, the extracted data will be unblinded for the other reviewer. Disagreements will be resolved through discussion and arbitration by the PI (BIN), if necessary. Two attempts will be made to contact the corresponding author in case relevant data are missing.

Data items

The following data items will be extracted from each included article:

General study information
- First author and year of publication.
- Country/countries in which the study was conducted.

Subject information
- Number of subjects (included in modelling, at baseline and at end of follow-up, where appropriate).
- Age of subjects (age span in which trajectories were identified).
- Source and characteristics of subjects (eg, if they were derived from a cohort (including cohort abbreviation and link to paper or website with information), if they were selected based on the presence of a condition etc).
- Percentage of recruited subjects that participated in the study at baseline.
- Percentage of drop-outs/withdrawals and summary of discussion regarding potential causes and impact of the missing data.

Trajectory-defining data and preprocessing
- Rationale/process for selection of trajectory-defining variables.
- Variables used to define trajectories (including source of data and mechanism of assessment, for example, self-report or clinical assessment).
- Preprocessing performed on such data (eg, imputation, scaling, categorisation, dimensionality reduction, etc, as well as methods for assessing/dealing with time variance, noise/variation in data, etc).
- Reproducibility measures taken (eg, publication of analysis code/data, transparent description of methods or the like).

Trajectory modelling
- Rationale/process for selection of trajectory modelling technique(s), including (hyper)parameters.
- Technique(s) (including (hyper)parameters) used.
- Methods for optimising models for the given task/data, avoiding overfitting, etc.
- Methods for selecting optimal technique/number of trajectories.
- Reproducibility measures taken (eg, publication of analysis code/data, transparent description of methods, or the like).

Evaluation/validation of trajectories and associated risk factors/outcomes
- External validation (if it was performed, and if so, short description of results).
- Evaluation of clinical, epidemiological or pathophysiological meaning/impact of derived trajectories.

► Associated risk factors investigated (ie, variables investigated as risk factors for subsequently being assigned to the trajectory; including rationale for selection of said variables and methods for assessing association).

► Associated outcomes investigated (ie, variables for which assignment to the trajectory was investigated as a risk factor; as above).

► For each trajectory:
  – The given name (e.g., 'late-onset eczema').
  – Percentage of the full study population.
  – Details/timing of characteristics (separated by static (eg, gestational age) and dynamic (eg, frequency of wheezing) characteristics).
  – Point estimate and 95% CI for each investigated risk factor.
  – Point estimate and 95% CI for each investigated outcome.

Quality assessment

As there is no well-established quality assessment tool specific to studies of (computational) trajectory analysis, and given the specific characteristics of eligible studies, a custom quality assessment tool has been prospectively developed by DL, BIN and RB. The tool is based on the structure and rating system of the Effective Public Health Practice Project (EPHPP)[32] tool (with some core sections/questions remaining). The sections on methodological aspects of the trajectory exploration ((1) preprocessing; (2) trajectory modelling and (3) evaluation and reporting of results) were based on: related systematic reviews by Bashir et al,[33] Meijs et al[34] and Stafford et al[35 36]; a narrative review on computational patient trajectory analyses by Allam et al[10]; guidelines for reporting machine learning analyses by Luo et al[37] and Stevens et al[38]; quality assessment guidelines for machine learning analyses by Kocak et al[39] and Faes et al[40]; and the Guidelines for Reporting on Latent Trajectory Studies checklist by Van de Schoot et al.[41] See online supplemental text for details on the theoretical background and reasoning for each section and item in the quality assessment tool. Each section ((a) selection bias; (b) data collection methods; (c) withdrawals and drop-outs; (d) preprocessing; (e) trajectory modelling; (f) associated risk factors and outcomes and (g) evaluation and reporting of results) will be rated in terms of quality as 'weak', 'moderate', 'strong' or 'not applicable'. An overall rating will also be given to each study based on the number of 'weak' section ratings, following the rating system of the EPHPP tool: 'weak' if ≥2 sections, 'moderate' if one section and 'strong' if no section was rated 'weak'. We acknowledge that the extensive restructuring of sections and items renders the interpretation of the quality assessment largely different from how the developers of EPHPP intended, including the fact that while the overall rating in the original EPHPP tool is based on six domains, our tool consists of seven domains; thus, statistical possibility of a weaker overall rating is increased.[42] The quality assessment tool (online supplemental file) was piloted and modified based on relevant articles identified during the PubMed pilot searches. Results of the quality assessment will be presented in a table (structure shown in online supplemental table 4).

Quality and risk of bias in each included study will be assessed independently in a double-blind fashion by the same pairs of reviewers that extracted data from said articles. Following completion, the ratings will be unblinded for the other reviewer. Disagreements will be resolved through discussion and arbitration by the PI (BIN), if necessary.

Data synthesis and statistical analysis

Extracted data items from each included study will be narratively synthesised and tabulated in a table of characteristics (structure shown in online supplemental table 5), except articles under embargo, conference abstracts and abstracts without a full text, which will only be noted/referenced in the manuscript and in a separate table (structure shown in online supplemental table 6). Line plots will be produced to illustrate: (a) the number of studies published across time; (b) the number of studies using each of the different trajectory modelling techniques across time and (c) the number of studies of low, moderate and high overall quality rating across time. Furthermore, a world map will be drawn, with each country coloured in a shade proportional to the number of studies from said country, to illustrate regional density of conducted research on the topic.

A table (structure shown in online supplemental table 7) will be produced to summarise trajectory-defining characteristics, associated risk factors/outcomes and the frequency at which distinct trajectories have been identified. Depending on the quantity and nature of the findings, additional tables may be produced to summarise, for example, disease-specific trajectories (or combinations thereof). Each section in the table(s) will be populated by one trajectory assessed to be distinct from the other trajectories described across the included studies and in which the ages of the subjects are comparable. The number of studies which have identified said trajectory (based on fraction/composition of identical or similar characteristics, as assessed by DL in agreement with BIN) will be presented. In the middle column, the trajectory characteristics will be described. Dynamic characteristics (eg, frequency of wheezing) will be plotted with one line representing the estimates of each study on the Y-axis (eg, percentage of subjects reporting wheezing) and age on the X-axis, or described narratively, depending on data form/availability. Static characteristics (eg, gestational age) will be presented as the percentage of subjects with said characteristic, together with the corresponding 95% CI, which will be calculated with the Wilson score interval method without continuity correction (suitable in case of small samples or proportions close to 0 or 1, which is expected in the present context).[43 44] The percentage with 95% CI from individual studies will be separated by a comma. In addition, the pooled percentage with 95% CI will be calculated and presented, where possible (details

in paragraph below). In the left column, risk factors (eg, maternal smoking during pregnancy) will be presented with the point estimate and 95% CI from each study separated by a comma, as well as the pooled point estimate and 95% CI, where possible (details in paragraph below). In the right column, outcomes (eg, asthma hospitalisation) will be shown, in a similar fashion as risk factors. The data in the left and right columns will be expressed as risk ratios (RRs) and converted to estimates of RR if needed (details in paragraph below). Characteristics, risk factors and outcomes will be color-coded according to the following domains (based on findings from the PubMed pilot searches as well as domain expertise among the authors; see online supplemental table 8) for more details):

1. Personal data (eg, sex and gestational age).
2. Atopy (eg, assessment through skin prick test).
3. Inflammation (eg, measures of blood neutrophils and eosinophils).
4. Food allergy (including family history, symptoms, diagnosis, healthcare use, medication and (indirect) measure of disease).
5. Atopic dermatitis (including family history, symptoms, diagnosis, healthcare use, medication and (indirect) measure of disease).
6. Allergic rhinitis, conjunctivitis and rhinoconjunctivitis (including family history, symptoms, diagnosis, healthcare use, medication and (indirect) measure of disease).
7. Asthma and wheezing (including family history, symptoms, diagnosis, healthcare use, medication and (indirect) measure of disease).
8. Behavioural and socioeconomic data (eg, absenteeism from school, day-care attendance, etc).
9. Environmental exposure (eg, maternal smoking during pregnancy, exposure to mould at home, diet types and food introduction timing, early childhood infection type/frequency, etc).
10. Comorbidity and related health measures (comorbidities and other health data not directly related to asthma or allergy, for example, body mass index (BMI), height, diabetes, etc).
11. Other (data not fitting elsewhere).

Given the heterogeneous and explorative nature of eligible studies and the aims of the present systematic review, we expect limited possibilities to conduct meta-analysis. Nevertheless, where numerical data on risk factors and outcomes associated with the derived trajectories are deemed comparable (in terms of study population, subject age, trajectory characteristics, control group and risk factor/outcome investigated, as assessed by DL in agreement with BIN), meta-analysis will be performed. Similarly, meta-analysis will be used to pool the percentages of static characteristics in those trajectories for which such data are deemed comparable (in terms of study population, trajectory modelling technique and nature of the specific data, as assessed by DL in agreement with BIN). As the eligible studies are expected to be

heterogeneous and estimate varying true effect sizes and percentages, the random-effects model is deemed most appropriate.[45 46]

For the risk factor and outcomes meta-analyses, random-effects robust variance estimation (RVE[47]; *robumeta*[48] R package) will be used, as it enables the inclusion of statistically dependent effect sizes (eg, based on the same control group, measurements at different time points and related measures of outcome) in the same model,[47] which is expected to constitute part of the eligible data.[24] Furthermore, the exact dependence structure does not need to be known when using the RVE method,[49] and assumptions, such as normal distribution of effect sizes and their estimates, are relaxed.[47] Pooled point estimates with 95% CI will be produced, using either the 'CORR' or 'HIER' model weighting scheme, depending on the type of statistical dependency in the included studies, 'CORR' being applicable if overall, the meta-analysis data stems from studies that report multiple estimates based on the same subjects, while non-independent data suitable for 'HIER' stems from different sets of subjects but share other influences, for example, being evaluated by the same group of researchers and/or using the same protocol/tools.[50] In case of 'CORR' (correlated effects), the default rho value (within-study effect size correlation) of 0.8 will be used.[51] Separate meta-analyses will be performed for each pair of risk factor/outcome and trajectory, if there are comparable numerical data from ≥2 separate studies.[52 53] Small sample correction for both the residuals and df, which increases performance in small samples of studies, will be used,[50] as we expect the number of studies in individual meta-analyses to be relatively low. Heterogeneity will be assessed through calculation of the: (a) proportion of between-study variance not due to random sampling error (I-squared; I^2)[54]; (b) between-study variance (Tau-squared; τ^2).[55–57] Forest plots will be created to present the meta-analysis results using the *forestploter*[58] R package. The pooled point estimates and corresponding 95% CI will also be displayed in online supplemental table 7. A p value of <0.01 instead of the default threshold of <0.05 will define statistical significance in meta-analyses with Satterthwaite df (df_{Sk})<4, as these have been reported to be prone to type I errors.[59 60] RR will be used as measure of effect due to intuitive interpretation.[61] Data expressed as incidence rate/risk ratio, prevalence ratio and relative risk ratio will be used without conversion, as these are calculated identically to RR.[62] Likewise, HR and OR data will be used without conversion as long as the outcome is <15% (at the end of follow-up). In case the outcome is more common (≥15%), estimates of RR will be calculated through the following formulae[63]:

- $RR \approx \sqrt{OR}$
- $RR \approx \dfrac{1 - 0.5\sqrt{HR}}{1 - 0.5\sqrt{\frac{1}{HR}}}$

For static characteristics, meta-analysis will be performed using a generalised linear mixed model (GLMM) with logit-transformed percentages from individual studies. GLMM was chosen due to the generally

lower risk of bias compared with two-step approaches and suitability for cases where the data contain small sample sizes or high proportions, which is expected in the present work.[64 65] The Wilson score interval method without continuity correction (suitable in case of small samples or proportions close to 0 or 1, which is expected in the present work)[43 44] will be used to produce the corresponding 95% CI for the percentage in individual studies. The *meta*[66] R package will be used for the meta-analyses of static characteristics.

Sensitivity analysis will be performed by repeating each meta-analysis in which ≥2 studies remain after excluding studies given an overall 'weak' rating. Publication bias will be assessed in case of ≥10 studies[67] in individual meta-analyses, using the *metafor*[68] R package through the means of[69]: (a) visual inspection of asymmetry in funnel plots; (b) statistical tests through Begg and Mazumdar correlation test[70] and Egger's regression test.[71] The trim-and-fill method[72] will be used to assess how many studies would be needed to normalise an asymmetric funnel plot. The code used to perform the above analyses will be written in R statistical software 4.2.3 (R Core Team, 2023) and together with underlying data made freely available at https://osf.io/ayf35/.

DISCUSSION

Promising research has been published in the field of trajectory exploration of allergic diseases and asthma, identifying novel and clinically meaningful subgroups. Our work—through the inclusion of a broad set of relevant diseases as well as an exhaustive search in ten databases without restriction by language—will provide a comprehensive overview of the current knowledge and methodological trends on this topic. While a restriction on publication date to the past 10 years will be implemented, the rapidly increasing body of research in this area, together with advancements in trajectory modelling techniques, will ensure a broad coverage of findings with focus on the latest methodological trends, building on previous literature and progress. Given the relative novelty and explorative nature of this area of research, interpretability will be limited due to the lack of well-established methodological principles on which assessment can be made regarding soundness of underlying computational approaches, reproducibility and clinical meaningfulness of the identified trajectories. Furthermore, the quality assessment form developed for the present work itself—although detailed and based on a broad set of guidelines, checklists and reviews—has not been externally validated, which warrants cautious interpretation of the rating results. Finally, as we anticipate low number of studies in most meta-analyses, the reliability of the pooled estimates may be relatively low. While some methods offer more reliable estimation in such scenarios, for example, Bayesian modelling,[73] the lack of strong priors in this field heavily limits our options. In summary, we believe this systematic review will provide value by summarising the central aspects of recent studies, highlighting repeatedly identified trajectories and their characteristics, as well as outlining methodological trends and limitations and perspectives for future work.

ETHICS AND DISSEMINATION

Ethical approval is not warranted due to the exclusive use of publicly available aggregated data. The findings in this study will be published in a peer-reviewed journal and underlying data and analysis code made freely available in an online repository (https://osf.io/ayf35/).

Author affiliations
[1]Krefting Research Centre, Institute of Medicine, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden
[2]Department of Clinical Science and Community Health, University of Milan, Milan, Italy
[3]Pediatric Unit, Ospedale Maggiore Policlinico, Milano, Italy
[4]Department of Pediatrics, University of Gothenburg Sahlgrenska Academy, Gothenburg, Sweden
[5]Department of Clinical Sciences, Lund University, Lund, Sweden
[6]Tampere University Respiratory Research Group, Faculty of Medicine and Health Technology, Tampere University, Tampere, Finland
[7]Wallenberg Centre for Molecular and Translational Medicine, University of Gothenburg, Gothenburg, Sweden

ORCID iDs
Daniil Lisik http://orcid.org/0000-0002-0220-5961
Gregorio Paolo Milani http://orcid.org/0000-0003-3829-4254
Michael Salisu http://orcid.org/0009-0001-2100-9392
Saliha Selin Özuygur Ermis http://orcid.org/0000-0003-3507-773X
Emma Goksör http://orcid.org/0000-0001-9595-1877
Göran Wennergren http://orcid.org/0000-0002-7010-7191
Hannu Kankaanranta http://orcid.org/0000-0001-5258-0906
Bright I Nwaru http://orcid.org/0000-0002-2876-6089

REFERENCES

1 García-Marcos L, Asher MI, Pearce N, *et al*. The burden of asthma, hay fever and eczema in children in 25 countries: GAN Phase I study. *Eur Respir J* 2022;60:02866–2021.
2 Loh W, Tang MLK. The Epidemiology of Food Allergy in the Global Context. *Int J Environ Res Public Health* 2018;15.
3 Spolidoro GCI, Amera YT, Ali MM, *et al*. Frequency of food allergy in Europe: an updated systematic review and meta-analysis. *Allergy* 2023;78:351–68.
4 Hill DA, Spergel JM. The atopic march: critical evidence and clinical relevance. *Ann Allergy Asthma Immunol* 2018;120:131–7.
5 Tsuge M, Ikeda M, Matsumoto N, *et al*. Current Insights into Atopic March. *Children (Basel)* 2021;8:1067.
6 Spergel JM. From atopic dermatitis to asthma: the atopic march. *Ann Allergy Asthma Immunol* 2010;105:99–106.
7 Dharmage SC, Lowe AJ, Tang MLK. Revisiting the Atopic March Current Evidence. *Am J Respir Crit Care Med* 2022;206:925–6.
8 Maiello N, Giannetti A, Ricci G, *et al*. Atopic dermatitis and atopic march: which link? *Acta Biomed* 2021;92:e2021525.
9 Owora AH, Zhang Y. Childhood wheeze trajectory-specific risk factors: a systematic review and meta-analysis. *Pediatr Allergy Immunol* 2021;32:e13313:34–50:.
10 Allam A, Feuerriegel S, Rebhan M, *et al*. Analyzing Patient Trajectories With Artificial Intelligence. *J Med Internet Res* 2021;23:e29812.
11 Kim JH, Chang HS, Shin SW, *et al*. Lung Function Trajectory Types in Never-Smoking Adults With Asthma: clinical Features and Inflammatory Patterns. *Allergy Asthma Immunol Res* 2018;10:614–27.
12 Amat F, Saint-Pierre P, Bourrat E, *et al*. Early-onset atopic dermatitis in children: which are the phenotypes at risk of asthma? Results from the ORCA cohort. *PLoS ONE* 2015;10:e0131369.
13 Hu C, Nijsten T, van Meel ER, *et al*. Eczema phenotypes and risk of allergic and respiratory conditions in school age children. *Clin Transl Allergy* 2020;10:7.
14 Suaini NHA, Yap GC, Bui DPT, *et al*. Atopic dermatitis trajectories to age 8 years in the GUSTO cohort. *Clin Exp Allergy* 2021;51:1195–206.
15 Belgrave DCM, Granell R, Simpson A, *et al*. Developmental profiles of eczema, wheeze, and rhinitis: two population-based birth cohort studies. *PLoS Med* 2014;11:e1001748.
16 Bacharier LB, Beigelman A, Calatroni A, *et al*. Longitudinal Phenotypes of Respiratory Health in a High-Risk Urban Birth Cohort. *Am J Respir Crit Care Med* 2019;199:71–82.
17 Park SY, Jung HW, Lee JM, *et al*. Novel Trajectories for Identifying Asthma Phenotypes: a Longitudinal Study in Korean Asthma Cohort, COREA. *J Allergy Clin Immunol Pract* 2019;7:1850–7.
18 Kurukulaaratchy RJ, Zhang H, Patil V, *et al*. Identifying the heterogeneity of young adult rhinitis through cluster analysis in the Isle of Wight birth cohort. *J Allergy Clin Immunol* 2015;135:143–50.
19 Deliu M, Sperrin M, Belgrave D, *et al*. Identification of Asthma Subtypes Using Clustering Methodologies. *Pulm Ther* 2016;2:19–41.
20 Oksel C, Haider S, Fontanella S, *et al*. Classification of Pediatric Asthma: from Phenotype Discovery to Clinical Practice. *Front Pediatr* 2018;6:258.
21 Cunha F, Amaral R, Jacinto T, *et al*. A Systematic Review of Asthma Phenotypes Derived by Data-Driven Methods. *Diagn (Basel)* 2021;11:644.
22 Yu S, Liao KP, Shaw SY, *et al*. Toward high-throughput phenotyping: unbiased automated feature extraction and selection from knowledge sources. *J Am Med Inform Assoc* 2015;22:993–1000.

23 Bosma AL, Ascott A, Iskandar R, *et al*. Classifying atopic dermatitis: a systematic review of phenotypes and associated characteristics. *J Eur Acad Dermatol Venereol* 2022;36:807–19.
24 Owora AH, Zhang Y. Childhood wheeze trajectory-specific risk factors: a systematic review and meta-analysis. *Pediatr Allergy Immunol* 2021;32:34–50.
25 Moher D, Shamseer L, Clarke M, *et al*. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Syst Rev* 2015;4:1.
26 Page MJ, McKenzie JE, Bossuyt PM, *et al*. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 2021;372:71.
27 Stroup DF, Berlin JA, Morton SC, *et al*. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. *JAMA* 2000;283:2008–12.
28 Jackson JL, Kuriyama A, Anton A, *et al*. The Accuracy of Google Translate for Abstracting Data From Non-English-Language Trials for Systematic Reviews. *Ann Intern Med* 2019;171:677–9.
29 Haddaway NR, Collins AM, Coughlin D, *et al*. The Role of Google Scholar in Evidence Reviews and Its Applicability to Grey Literature Searching. *PLoS ONE* 2015;10:e0138237.
30 Bramer WM, Giustini D, de Jonge GB, *et al*. De-duplication of database search results for systematic reviews in EndNote. *J Med Libr Assoc* 2016;104:240–3.
31 Pires GN, Arnardóttir ES, Islind AS, *et al*. Consumer sleep technology for the screening of obstructive sleep apnea and snoring: current status and a protocol for a systematic review and meta-analysis of diagnostic test accuracy. *J Sleep Res* 2023;32:e13819.
32 Armijo-Olivo S, Stiles CR, Hagen NA, *et al*. Assessment of study quality for systematic reviews: a comparison of the Cochrane Collaboration Risk of Bias Tool and the Effective Public Health Practice Project Quality Assessment Tool: methodological research. *J Eval Clin Pract* 2012;18:12–8.
33 Bashir MBA, Basna R, Zhang G-Q, *et al*. Computational phenotyping of obstructive airway diseases: protocol for a systematic review. *Syst Rev* 2022;11:216.
34 Meijs C, Handoko ML, Savarese G, *et al*. Discovering Distinct Phenotypical Clusters in Heart Failure Across the Ejection Fraction Spectrum: a Systematic Review. *Curr Heart Fail Rep* 2023;20:333–49.
35 Stafford IS, Gosink MM, Mossotto E, *et al*. A Systematic Review of Artificial Intelligence and Machine Learning Applications to Inflammatory Bowel Disease, with Practical Guidelines for Interpretation. *Inflamm Bowel Dis* 2022;28:1573–83.
36 Stafford IS, Kellermann M, Mossotto E, *et al*. A systematic review of the applications of artificial intelligence and machine learning in autoimmune diseases. *NPJ Digit Med* 2020;3:30.
37 Luo W, Phung D, Tran T, *et al*. Guidelines for Developing and Reporting Machine Learning Predictive Models in Biomedical Research: a Multidisciplinary View. *J Med Internet Res* 2016;18:e323.
38 Stevens LM, Mortazavi BJ, Deo RC, *et al*. Recommendations for Reporting Machine Learning Analyses in Clinical Research. *Circ Cardiovasc Qual Outcomes* 2020;13:e006556.
39 Kocak B, Kus EA, Kilickesmez O. How to read and review papers on machine learning and artificial intelligence in radiology: a survival guide to key methodological concepts. *Eur Radiol* 2021;31:1819–30.
40 Faes L, Liu X, Wagner SK, *et al*. A Clinician's Guide to Artificial Intelligence: how to Critically Appraise Machine Learning Studies. *Transl Vis Sci Technol* 2020;9:7.
41 van de Schoot R, Sijbrandij M, Winter SD, *et al*. The GRoLTS-Checklist: Guidelines for Reporting on Latent Trajectory Studies. *Struct Equ Modeling* 2017;24:451–67.
42 Igelström E, Campbell M, Craig P, *et al*. Cochrane's risk of bias tool for non-randomized studies (ROBINS-I) is frequently misapplied: a methodological systematic review. *J Clin Epidemiol* 2021;140:22–32.
43 Newcombe RG. Two-sided confidence intervals for the single proportion: comparison of seven methods. *Stat Med* 1998;17:857–72.
44 Wei L, Hutson AD. A comment on sample size calculations for binomial confidence intervals. *J Appl Stat* 2013;40:311–9.
45 Borenstein M, Hedges LV, Higgins JPT, *et al*. A basic introduction to fixed-effect and random-effects models for meta-analysis. *Res Synth Methods* 2010;1:97–111.
46 Higgins JPT, Thompson SG, Spiegelhalter DJ. A re-evaluation of random-effects meta-analysis. *J R Stat Soc Ser A Stat Soc* 2009;172:137–59.
47 Hedges LV, Tipton E, Johnson MC. Robust variance estimation in meta-regression with dependent effect size estimates. *Res Synth Methods* 2010;1:39–65.

48 Fisher Z, Tipton E, Zhipeng H. robumeta: robust variance meta-regression. 2017.

49 Pustejovsky JE, Tipton E. Meta-analysis with Robust Variance Estimation: expanding the Range of Working Models. *Prev Sci* 2022;23:425–38.

50 Tipton E. Small sample adjustments for robust variance estimation with meta-regression. *Psychol Methods* 2015;20:375–93.

51 Lisik D, Ermis SSÖ, Ioannidou A, *et al*. Is sibship composition a risk factor for childhood asthma? Systematic review and meta-analysis. *World J Pediatr* 2023;19:1127–38.

52 Ahn E, Kang H. Introduction to systematic review and meta-analysis. *Korean J Anesthesiol* 2018;71:103–12.

53 Deeks JJ, Higgins JPT, Altman DG, *et al*. Analysing data and undertaking meta-analyses. *Cochrane Handb Syst Rev Interv* 2019;2019:241–84.

54 Higgins JPT, Thompson SG, Deeks JJ, *et al*. Measuring inconsistency in meta-analyses. *BMJ* 2003;327:557–60.

55 von Hippel PT. The heterogeneity statistic I2 can be biased in small meta-analyses. *BMC Med Res Methodol* 2015;15:35.

56 Rücker G, Schwarzer G, Carpenter JR, *et al*. Undue reliance on I(2) in assessing heterogeneity may mislead. *BMC Med Res Methodol* 2008;8:79.

57 Higgins JPT, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stat Med* 2002;21:1539–58.

58 Dayimu A. forestploter: create flexible forest plot. 2022.

59 Fisher Z, Tipton E. robumeta: an R-package for robust variance estimation in meta-analysis. 2015.

60 Tanner-Smith EE, Tipton E, Polanin JR. Handling Complex Meta-analytic Data Structures Using Robust Variance Estimates: a Tutorial in R. *J Dev Life Course Criminol* 2016;2:85–112.

61 George A, Stead TS, Ganti L. What's the Risk: differentiating Risk Ratios, Odds Ratios, and Hazard Ratios? *Cureus* 2020;12:e10047.

62 Martinez BAF, Leotti VB, Silva G de SE, *et al*. Odds Ratio or Prevalence Ratio? An Overview of Reported Statistical Methods and Appropriateness of Interpretations in Cross-sectional Studies with Dichotomous Outcomes in Veterinary Medicine. *Front Vet Sci* 2017;4:193.

63 VanderWeele TJ, Ding P. Sensitivity Analysis in Observational Research: Introducing the E-Value. *Ann Intern Med* 2017;167:268–74.

64 Lin L, Chu H. Meta-analysis of Proportions Using Generalized Linear Mixed Models. *Epidemiology (Sunnyvale)* 2020;31:713–7.

65 Lin L, Xu C, Chu H. Empirical Comparisons of 12 Meta-analysis Methods for Synthesizing Proportions of Binary Outcomes. *J Gen Intern Med* 2022;37:308–17.

66 Balduzzi S, Rücker G, Schwarzer G. How to perform a meta-analysis with R: a practical tutorial. *Evid Based Ment Health* 2019;22:153–60.

67 Dalton JE, Bolen SD, Mascha EJ. Publication Bias: the Elephant in the Review. *Anesth Analg* 2016;123:812–3.

68 Viechtbauer W. Conducting meta-analyses in R with the metafor package. *J Stat Softw* 2010;36:1–48.

69 Lin L, Chu H, Murad MH, *et al*. Empirical Comparison of Publication Bias Tests in Meta-Analysis. *J Gen Intern Med* 2018;33:1260–7.

70 Begg CB, Mazumdar M. Operating characteristics of a rank correlation test for publication bias. *Biometrics* 1994;50:1088–101.

71 Egger M, Davey Smith G, Schneider M, *et al*. Bias in meta-analysis detected by a simple, graphical test. *BMJ* 1997;315:629–34.

72 Duval S, Tweedie R. Trim and fill: a simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics* 2000;56:455–63.

73 Reis DJ, Kaizer AM, Kinney AR, *et al*. A practical guide to random-effects Bayesian meta-analyses with application to the psychological trauma and suicide literature. *Psychol Trauma* 2023;15:121–30.