Title page

Title: Real-world evidence: Methods for assessing long term health and effectiveness of allergy immunotherapy

Authors: Benedikt Fritzsching, Marco Contoli, Celeste Porsbjerg, Sarah Buchs, Julie Rask Larsen, and Nick Freemantle

Author affiliation: Benedikt Fritzsching, MD, Paediatric Pulmonology and Allergy, Children's Doctor Service; Heidelberg and University of Heidelberg, Germany

Marco Contoli, PhD, Respiratory Section, Department of Translational Medicine, University of Ferrara, Italy

Celeste Porsbjerg, PhD, Department of Respiratory Medicine, Bispebjerg Hospital, University of Copenhagen, Denmark

Sarah Buchs, MSc, Global Market Access, ALK, Denmark

Julie Rask Larsen, PhD, Global Medical Affairs, ALK, Denmark

Nick Freemantle PhD, Institute of Clinical Trials and Methodology UCL, UK

Manuscript count

999 words
9 references
1 figure and 1 table

Real-world evidence: Methods for assessing long term health and effectiveness of allergy immunotherapy

Recently scientific societies and regulatory authorities have stressed the need for robust real-world evidence (RWE) in allergic respiratory disease.[1-4] Despite their importance, only a limited number of RWE studies (RWS) has been conducted in allergy immunotherapy (AIT) [4] (Figure 1).

The highest level of evidence was suggested if coexisting RCTs and observational studies provide consistent findings.[2,5] If appropriate methods and standardized protocols are applied in RWS there is the opportunity to transform real-world data into evidence of high clinical relevance consistent with EAACI´s 2021 position.[2] However, as the robustness of RWS can be hampered by several factors, this paper aims to identify the caveats.

Data sources and study designs in RWE

Selecting the appropriate data source is central for the quality of generated data. As heterogeneity of the study population is an intrinsic characteristic of RWS, large cohorts are required. RWE may be generated prospectively through primary data collection or retrospectively, using secondary data sources, e.g. registries, healthcare claims, or prescription databases, which have become more complete and comprehensive in recent years. Routinely collected healthcare data offer advantages in providing access to large representative samples of patients in routine clinical practice along with the potential for long-term follow-up. Recent publications discussed strengths and limitations of such data sources in AIT. [2,6] The following will focus on retrospective data sources and the assessment of effectiveness in RWS.

Generating high-quality RWE

Conducting a high quality and scientifically reliable RWS requires rigorous methodology largely mirroring what is done in RCTs, although without randomisation, which alone provides unbiased estimates (Table 1). Quality standards for reporting RCTs are well-described in the CONSORT statement.[7] Frameworks to assess the quality of RWS are described by the STROBE checklist[8] or the RELEVANT tool with key principles reflecting the requirements for RCTs.[4]

1. Pre-specification and transparency: Hypotheses and research questions must be pre-specified to avoid post-hoc 'fishing' for interesting outcomes.[4] Transparency about the study design and analysis before execution is key and requires pre-registration of the RWS in a public registry or the publication of the study design, and a prespecified statistical analysis plan.[4] Similar to an RCT where the population is selected according to eligibility criteria and the database is locked prior to unblinding the results, RWS should have pre-specified enrolment criteria and study cohorts can be locked prior to testing the protocol-defined hypotheses.

2. Ensure comparable groups to avoid confounding: To minimize the risk of bias and confounding, it is important to identify, control for, and address potential confounders in retrospective database studies.[4] Due to the absence of randomization, retrospective studies are subject to confounding by indication/ disease severity where exposure is associated with additional unmeasured risk. Rosenbaum and Rubin who derived the methods for propensity score analyses pointed out that, to be valid, exposure to the treatment of interest should be 'strongly ignorable' as a source of additional information on underlying risk. Hence, if

a treatment is routinely given to higher risk subjects that risk must be described separately and accounted for.[9] The risk for confounding is greater, the more dissimilar the compared groups are and ideally, matching designs like propensity score matching (PSM) or instrumental variable techniques are used.[4] Comparison of the study groups at baseline before and after matching is important.[4] If key confounding variables differ at baseline, there is a higher likelihood for substantial confounding in the results.[4] To assess validity and potential confounding, RWS should aim to replicate the findings of RCTs in similar populations where possible, before bridging into different populations and longer follow-up.

3. Pre-defined outcomes measured in a valid way and reported transparently: The objectives and primary outcome(s) must be defined and measured in a valid way. In retrospective studies, it is often necessary to use proxies for outcomes of effectiveness or disease severity, e.g. for AIT studies the prescription medications for AR and asthma, as well as confirmed diagnoses are used as proxies of disease severity in the lack of symptom scores.[4] RWS results should be presented for primary and secondary outcomes, as well as the results of sensitivity analyses.[4] RWS is that they often reflect the treatment of representative groups of subjects (avoiding the selection that occurs in randomized trials) and thus address an objective that may be considered a treatment policy strategy estimand. In other words has an objective which defines as irrelevant intercurrent clinical events which occur in the real world.[https://www.ema.europa.eu/en/documents/scientific-guideline/ich-e9-r1-addendum-estimands-sensitivity-analysis-clinical-trials-guideline-statistical-principles_en.pdf] Many randomised trials, through their participant selection processes, become more hypothetical in strategy as participants are not included who are likely to experience a range of realistic intercurrent events.

4. Results in perspective of existing research: Findings should be adequately discussed in line with available information and the clinical relevance.[4] Limitations including potential biases and confounding factors should be described following a discussion on how these may influence the results.[4]

5. Transparency on conflict of interest: Any conflict of interest must be transparently reported,[4] and measures should be taken to mitigate the conflict of interest, e.g. through involvement of third parties in the design, conduct, and analysis.[4]

Example: The Real-world effectiveness in allergy immunotherapy (REACT) study

The aim of the REACT study was to assess the effectiveness of AIT and provide high-quality RWE on how AIT works long-term and in real life (data on file). In this retrospective database study rigorous methodology was applied, comprising a pre-specified objective and processes described in the protocol, pre-registration at ClinicalTrials.gov, cohort-lock prior to outcomes analyses, and all analyses were conducted by an independent third party. PSM was used to ensure comparable groups and mitigate confounding. Subjects were matched based on many variables available in the database, including demographics, diagnosis codes for relevant comorbidities, prescriptions for AR and asthma medication, health resource utilization, and costs. The REACT study complements the evidence of existing RCTs and supports clinical decision-making on AIT for the treatment and sustained control of allergic rhinitis and asthma.

With an increasing amount of data, researchers have unique opportunities to generate valid and good quality RWE of the effectiveness of AIT with the application of rigorous and high scientific standards. While the focus herein was retrospective RWS, other RWE study types will likely benefit from using clinical RCT knowledge. Altogether, robust assessments of effectiveness of AIT in RWS complement the existing evidence of effect and safety of AIT in RCTs.

Figure 1

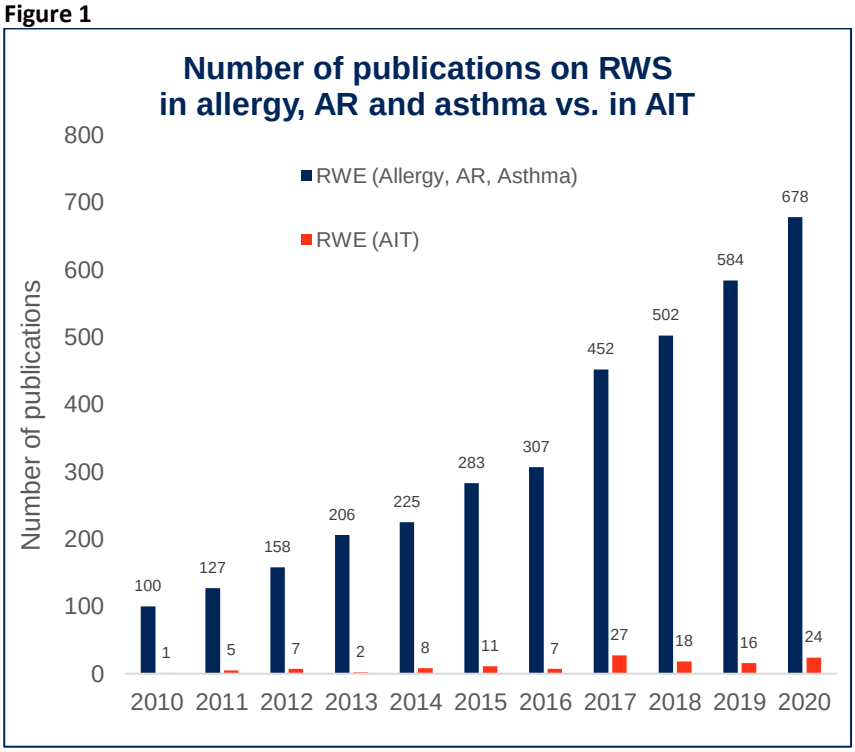

Figure 1 While the number of published real-world evidence studies have increased significantly the past 10 years within allergy, allergic rhinitis, and asthma, except for allergy immunotherapy, where real-world studies remain scarce. Search terms: (Real-world data OR Real-world evidence OR Registry) AND (asthma OR allergy OR allergic rhinitis) AND (allergy immunotherapy), with or without AIT included. Results were summarized by calendar year. AR, allergic rhinitis; AIT, allergy immunotherapy; RWS, real-world studies.

Table 1: Key methodology aspects for assessing the effects of allergy immunotherapy (AIT) in retrospective real world evidence studies of high quality by mirroring randomised controlled clinical trials.

	Randomised controlled clinical trials	*Retrospective real world evidence studies*
Transparency	Pre-register study protocol in a public registry or report study design in a publication	Pre-register study protocol in a public registry or report study design in a publication
	Report conflict of interest and follow ICMJE authorship recommendations	Report conflict of interest and follow ICMJE authorship recommendations
	Adhere to reporting standards, e.g. CONSORT	Adhere to reporting standards, e.g. STROBE
Minimise risk of bias and confounding	Pre-specify eligibility criteria, outcomes, and statistical analyses plan	Pre-specify eligibility criteria, outcomes, and statistical analyses plan
	Ensure comparable groups through randomisation	Ensure comparable groups through appropriate methods, e.g. propensity score matching
Analyses	Efficacy of AIT is assessed using daily diaries and often reported as TCS, DSS, DMS and RQLQ	Effectiveness of AIT can be assessed using proxies like prescriptions, diagnosis codes, HRU and cost
	Database lock before analyses and/or unblinding	Cohort lock before analyses
Interpretation	Identify key strengths and limitations and critically discuss results	Identify key strengths and limitations and critically discuss results in context of existing RCT evidence

ICMJE: International committee of Medical Journal Editors. CONSORT: CONsolidated Standards of Reporting Trials. STROBE: Strengthening the Reporting of Observational Studies in Epidemiology. TCS: Total Combined Score. DSS: Daily Symptoms Score. DMS: Daily Medication Score. RQLQ: Rhinitis Quality of Life Questionnaire. HRU: Health Resource Utiliszation. AR: Allergic Rhinitis

References

1. FDA. Submitting Documents Using Real-World Data and Real-World Evidence to FDA for Drugs and Biologics Guidance for Industry. May 2019. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/submitting-documents-using-real-world-data-and-real-world-evidence-fda-drugs-and-biologics-guidance.
2. Paoletti G, DiBona D, Chu DK, et al. Allergen immunotherapy: the growing role of observational and randomisedtrial "Real-World Evidence". *Allergy* 2021.
3. Roche N, Anzueto A, Bosnic Anticevich S, et al. The importance of real-life research in respiratory medicine: manifesto of the Respiratory Effectiveness Group: Endorsed by the International Primary Care Respiratory Group and the World Allergy Organization. *Eur Respir J* 2019; **54**(3).
4. Roche N, Campbell JD, Krishnan JA, et al. Quality standards in respiratory real-life effectiveness research: the REal Life EVidence AssessmeNt Tool (RELEVANT): report from the Respiratory Effectiveness Group-European Academy of Allergy and Clinical Immunology Task Force. *Clin Transl Allergy* 2019; **9**: 20.
5. Gershon AS, Jafarzadeh SR, Wilson KC, Walkey AJ. Clinical Knowledge from Observational Studies. Everything You Wanted to Know but Were Afraid to Ask. *Am J Respir Crit Care Med* 2018; **198**(7): 859-67.

6. Devillier P, Demoly P, Molimard M. Allergen immunotherapy: what is the added value of real-world evidence from retrospective claims database studies? *Expert Rev Respir Med* 2020; **14**(5): 445-52.
7. Schulz KADM, D; The Consort Group. CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. *The Lancet Regional Health - Europe* 2010.
8. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *PLoS Med* 2007; **4**(10): e296.
9. ROSENBAUM PR, RUBIN DB. The central role of the propensity score in observational studies for causal effects. *Biometrika* 1983; **70**(1): 41-55.

-