Big-Time Sensitization Rates in Young Germans: Big Numbers – Big Risks – Big Confusion?

Jörg Kleine-Tebbe

Allergy and Asthma Center Westend, Berlin, Germany

Welcome to well-powered data from wealthy Germany: almost 13,000 children and adolescents have been screened for serum IgE to 20 common inhalants and foods [1]. The nonselective design, a truly representative cohort, its huge sample size and appropriate statistics generated robust results. Sensitization rates based on allergen-specific serum IgE ≥ 0.35 kU_A/l to one or more allergen sources range from 29% in the youngest group (i.e. age 3–6 years) to 46.5% in the oldest group (14–17 years), this latter group being composed of 42% atopic girls and 51% atopic boys.

Extrapolate this into the future beyond adolescence – Yes, these are big numbers! What are they good for?

(1) They can provide trustworthy population-based information for everybody: doctors, scientists, the media, the general population, health-care stakeholders and politicians.

(2) Allergen manufacturers, diagnostic companies and industry will also be keen to dig into this data pool.

(3) Epidemiologists wanting to compare these statistics will look for similar cohorts or age groups and comment on these truly ‘epidemic’ results from Middle Europe.

However, extrapolations of any kind are not without some risks. There are major geographical differences due to variations in climate, natural allergen exposure (pollen, outdoor molds and food availability) and lifestyle factors

(e.g. the extent of indoor/outdoor living, pets and eating habits). The number of underlying variables is constantly growing, and this hampers comparisons between epidemiological studies on different regions.

Even more important, employing diagnostic extracts as tools for detecting allergic sensitizations (by skin prick test or serum IgE) carries inherent problems and can create misleading results:

(1) The peanut sensitization rate of 10.6% in the whole German cohort (age 3–17 years) is mainly driven by prevalent Bet v 1-related cross sensitizations to peanut allergen Ara h 8. This is presumably far beyond the much lower sensitization rate to ‘classical’, stable peanut allergens like Ara h 1, 2, 3 and 6 being associated with severe systemic symptoms. The same applies to soy (sensitization rate 6.3%) due to its Bet v 1 homologue Gly m 4 and the well-defined Bet v 1-cross-reactive foods like raw carrots, apples and potatoes with impressive sensitization rates of 9.7, 9.2 and 8.4%, respectively.

(2) Rates of high sensitization to rye (21.2%), wheat (9.9%) and rice (7.2%) are potentially misleading and are probably due to cross-reactive grass pollen allergens (e.g. Timothy grass pollen 22.7%), including cross-reactive carbohydrate determinants.

(3) Pollen sensitization rates (e.g. Timothy grass 22.7%, birch 14.1% and mugwort 10.9%) can be disputed for as

long as we do not separate the rates of primary sensitization (to genuine pollen major allergens) from those of secondary cross reactions (with markers like profilins, polcalcins and cross-reactive carbohydrate determinants), as this tends to obscure extract-based diagnostic methods to a certain extent.

(4) Rates of sensitization to animal hair (e.g. dog 9.7%, cat 8.1% and horse 4.4%) are also prone to overestimation due to newly identified cross-reactive single allergens in animals with fur.

Many of these extract-inherent problems could be solved by measuring IgE to single allergens, unfolding the fingerprint of each individual sensitization on a molecular basis. This approach, coined 'component-resolved diagnostics' long ago, has great potential for fuelling future epidemiological studies. At present, microarray techniques (i.e. ImmunoCAP ISAC, ThermoFisher) allow the detection of specific IgE to more than 110 single allergens from 50 sources with minute amounts of blood. Species-specific sensitizations could then be separated from cross-reactive phenomena. Summarized data on allergen sources could still be estimated by simply adding the IgE values of nonrelated single allergens from the same source.

Epidemiology studies based on molecular allergology would serve as powerful tools if combined with (regional) data on allergen exposure. Such maps, piloted with great enthusiasm in the diverse pollen regions of Spain [2], would offer novel insights: formerly extract-based sensitization rates could be resolved into truly species-specific or cross-reactive sensitization maps. Risky sensitizations to stable (food) allergens could be separated from sensitizations to rather labile proteins. Geographically defined differences in molecular sensitization rates could finally guide allergists to provide properly adapted diagnostic panels for improving allergen selection for subsequent immunotherapies [3].

When considering such maps on molecules combined with the solid KiGGs (German Health Interview and Examination Survey for Children and Adolescents) approach, what a wealth of data could be generated for Europe, the USA and even on a global scale! The resulting datasets would, however, require novel biometric and statistical approaches for the proper digestion of big numbers. Should we really detect sensitization rates in such great detail and consider them as being only clinically relevant in case of corresponding symptoms, if we feel not ready to synthesize and interpret their heterogeneity and complexity?

Genetic studies on allergy have been somewhat disappointing and epigenetic 'whole genome' approaches will

again demonstrate the risk of XXL data mining without having ways to recompose detailed mega-findings and make sense of highly scattered, partly redundant, mostly interdependent data with network-like relations. Stated simply: 'Why test it, if we do not know how to deal with it?'. Food scientists being confronted with thresholds in food detection assays that were too low – measuring everything almost everywhere – called this 'paralysis by analysis' (Steve Taylor, Lincoln, Nebr., USA).

Addressing complex problems by searching for even more details (thus adding more complexity) carries the risk of confusion... and, not unlikely, huge data cemeteries. Is this part of a general 'crisis in science'? (Donald W. MacGlashan Jr., Baltimore, Md., USA; personal communication, thank you, Don, for sharing!). The dilemma is: the more powerful our tools become for creating, handling, storing and crunching massive datasets with, for example, molecular, genetic or signaling information, the more we run the risk of being overwhelmed and no longer able to achieve clarity. Let us try and gain a new perspective. How could we avoid this rapidly increasing complexity in the biomedical sciences becoming a feature in molecular allergy and immunology? How could we improve our methods of analysis and recompose such huge detailed datasets in order to benefit from the information they offer? Novel, even unconventional ideas, like crowd-sampling via Adriano Mari's Allergome platform (www.allergome.org) [4] or prize-based contests [5], have already been put forward.

With this in mind, do enjoy the results and approaches presented by Roma Schmitz and her colleagues [1] and 'get your kicks from studying KiGGs'.

References

- 1 Schmitz R, Ellert U, Kalcklösch M, Dahm S, Thamm M: Patterns of sensitization to inhalant and food allergens – findings from the German Health Interview and Examination Survey for Children and Adolescents. *Int Arch Allergy Immunol* 2013;162:263–270.
- 2 Barber D, de la Torre F, Feo F, Florido F, Guardia P, Moreno C, Quiralte J, Lombardero M, Villalba M, Salcedo G, Rodríguez R: Understanding patient sensitization profiles in complex pollen areas: a molecular epidemiological study. *Allergy* 2008;63:1550–1558.
- 3 Sastre J, Landivar ME, Ruiz-García M, Andregnette-Rosigno MV, Mahillo I: How molecular diagnosis can change allergen-specific immunotherapy prescription in a complex pollen area. *Allergy* 2012;67:709–711.
- 4 Mari A, Rasi C, Palazzo P, Scala E: Allergen databases: current status and perspectives. *Curr Allergy Asthma Rep* 2009;9:376–383.
- 5 Lakhani KR, Boudreau KJ, Loh PR, Backstrom L, Baldwin C, Lonstein E, Lydon M, MacCormack A, Arnaout RA, Guinan EC: Prize-based contests can provide solutions to computational biology problems. *Nat Biotechnol* 2013;31:108–111.