

General principles of epidemiology, as applied to studies of IgE-mediated food allergy

Anne Ferguson *

GI Laboratory, Department of Medicine, University of Edinburgh, Western General Hospital, Edinburgh EH4 2XU, UK

Abstract

Epidemiology is the study of the distribution of disease and of the factors which contribute to or determine the development of a disease. Descriptive studies measure the frequency of a disease whereas analytic studies are predominantly concerned with identifying determinants or causes. In any epidemiological study it is essential that the exact objectives are clearly understood in terms of definitions and diagnostic criteria for the condition under scrutiny. Rather than address all food-induced reactions and diseases, the issues raised by epidemiological studies of one main group of conditions with a common mechanism, IgE-mediated food allergy, are considered. Problems with population-based and hospital-based surveys include case definition, case ascertainment, calculations and presentation of results and small numbers with wide confidence limits. © 1997 Elsevier Science B.V.

Keywords: Epidemiology; Food allergy; Food intolerance

1. Introduction

Within the wide range of adverse reactions to foods and food ingredients it is possible to define several specific diseases, such as coeliac disease or lactase deficiency. There are also particular patterns of reaction which can be recognised and separated from the rest on the basis of the particular types of food involved (e.g. complex carbohydrates, precipitating colonic symptoms) or the disease mechanism implicated (e.g. an immunological hypersensitivity). Finally, there remains the puzzling problem of people who describe food intolerance, but do not fit into any of the above categories.

In order that the powerful tools of epidemiology can be applied to investigations of food-induced reactions and diseases, it is essential that the exact objectives of any study are clearly understood in terms of definitions and diagnostic criteria for the condition under scrutiny.

The availability of non-invasive hydrogen breath tests for lactose malabsorption means that accurate data on this form of food intolerance can fairly easily be acquired. There has been extensive research on the epidemiology of coeliac disease (as defined pathologically), now facilitated by the availability of highly predictive serological tests. The work on food additive intolerance, described by Dr Young (this issue), succeeded because resources were available for double blind, placebo controlled, food challenges. In this paper, we discuss some of the theoretical and practical issues which may arise in studies of the epidemiology of IgE-mediated food allergy. Even within this type of food sensitivity there is very considerable heterogeneity in clinical manifestations, and in the foods concerned. Manifestation of food allergy as disease will be the result of a combination of genetic and environmental factors, both host-related (atopy) and socio-geographic (foods available in that society, and the individual person or parent's selection of diet). The designs of epidemiological research studies must take account of these factors at an early stage.

* Tel.: +44 131 5371730; fax: +44 131 5371007.

2. Types of epidemiological studies

Epidemiology is the study of the distribution of disease and of the factors which contribute to or determine the development of a disease. Thus epidemiological studies are of two main types, descriptive and analytic. Descriptive studies measure the frequency of a disease in terms of time, place and person (in other words, answering the questions when?, where? and who?). Analytic studies (cohort or case-control) are predominantly concerned with identifying determinants or causes (Hennekens and Buring, 1987; Logan, 1992).

Descriptive studies have two purposes. Firstly, for the planning and provision of health services and as a contribution to decisions on prioritisation of efforts and resources. Secondly, descriptions of variation in disease frequency in different countries or areas, or variation over time, may give clues as to possible environmental factors and thus serve to generate hypotheses as to the aetiology of the condition. When descriptive studies have proved to be the source of fruitful hypotheses the differences in frequency involved have invariably been large, often 5-fold or greater. This point needs to be kept in mind; in descriptive work, enormous efforts devoted to identifying the last few percent of cases in a study are rarely justified, as the influence of the last few small changes in the overall data set are generally unimportant.

The usual measures of disease frequency are:

Incidence—the rate at which new events or conditions occur in a population, i.e.

$$\frac{\text{Number of new events in a specified period}}{\text{Number of persons exposed in this period}} \times 10n$$

Cumulative incidence—the proportion of a population who develop an event or condition in a specified time period

Prevalence—the proportion of a population who have an event or condition at a point in time or who have had an event or condition during a specified period (e.g. a year or lifetime).

Finding out the incidence or prevalence of a condition sounds deceptively simple. The essentials are an accurate count of the number of cases already identified or of new cases developing, and knowledge of the size of the population from which the cases arose. To estimate incidence, it is necessary to be able to date the onset.

3. Particular features of epidemiology in relation to food allergy

As the actual clinical reaction to a food depends on dietary exposure, whereas the primary aetiological fac-

tor is accepted to be the existence of immunological sensitisation, onset of food allergy will be difficult to date, and so calculating incidence is of dubious validity. Most studies have therefore assessed frequency in terms of prevalence or cumulative incidence. Since clinical (although maybe not immunological) recovery often occurs, particularly with food allergy in infancy, cumulative incidence over a lifetime, even a short life, will be higher than the prevalence within a particular age-defined population.

Population surveys are the standard method of measuring prevalence and estimates of cumulative incidence can usually be derived as well. If the aim is only to establish the existence of atopy or of IgE sensitisation to one or several foods, suitable blood tests, or skin prick testing can be incorporated within population surveys.

For many diseases an alternative to surveys is to use the information available from routinely collected data such as hospital admission statistics, requests for special foods, invalidity or mortality statistics. As with most chronic diseases, mortality statistics are generally unhelpful. It is particularly regrettable that anaphylaxis—whether induced by food, drugs, insect bite or other allergen—is not acceptable for coding in death certificates. Published reports indicate that the vast majority of recorded deaths from food-induced anaphylaxis have been in patients who also have severe asthma, and prior to the proper recognition of fatal food allergy, such deaths were probably attributed to asthma.

If patients with food allergy are admitted to hospital for investigation or treatment, then hospital discharge statistics are another possible source of data. However, usually there will be two possible ways to record the diagnosis, the aetiopathological diagnosis of food allergy, and the organ-related diagnosis such as chronic diarrhoea or eczema. Until safe and simpler diagnostic criteria for food allergic manifestations are defined and internationally agreed, it is most unlikely that information of real value will be obtainable from hospital statistics. Case ascertainment currently requires experienced review of recorded clinical data, and, often, a decision re diagnosis based on the balance of probabilities. A further problem is that in most countries, hospital statistics count the number of discharges, not individual patients, and so allowance needs to be made for re-admission rates.

An interesting contrast can be drawn between studies in food allergy and respiratory allergy. Even though the official, agreed diagnostic criteria and definition of asthma amount to several pages, it has been perfectly possible for the definition to be greatly simplified for epidemiological work, e.g. 'wheezing more than three times last year with symptom-free intervals between the episodes'. Perhaps compromises could be reached between diagnostic security and operational realism, in

defining limited forms of expression of food allergy, such as eczema or chronic diarrhoea in infants, for comparative population studies. It has to be borne in mind that false positives and false negatives will balance one another out in a well planned, large study. However this, again, depends on all participants applying exactly the same effort and detail in patient assessment, and thus needs a well resourced central organisation, training of participants in questionnaire use and examination methods, and audit.

4. Problems with population-based and hospital-based surveys

Apart from a few retrospective and prospective population surveys (some of which are described elsewhere in this issue) studies of food allergy frequency have essentially involved a retrospective search of various hospital sources and related the number of cases identified to the hospital catchment area. In both of these types of study, problems can arise in four areas: (1) case definition; (2) case ascertainment; (3) calculations and presentation of results; (4) and small numbers with wide confidence limits.

4.1. Problems of case definition

Any epidemiological study requires a clearly stated definition of what constitutes a case. For example, studies of coeliac disease have a considerable advantage in the disease being defined in histopathological terms (including the response to diet). A definition based on histopathology has obvious advantages in terms of objectivity and independent verification even allowing for problems of sampling and biopsy interpretation. The early descriptions of cows' milk protein intolerance, and of food-sensitive colitis, also reported characteristic food-associated pathological changes in addition to immunological abnormalities and clinical responses to food challenge. However these methods, suitable for research on pathogenesis and diagnostic criteria, are impractical (indeed are not used) in general clinical work, on which any large epidemiological study must rely.

Thus it is important that epidemiological studies use case definitions that are realistic and reflect local clinical practice, and that it is accepted that there are likely to be some cases which do not fully meet whatever criteria are being used. Rather than excluding such cases it is advisable to collect what data are available and report the numbers as possible cases. Presenting results in this way also facilitates comparisons with other studies. For example in the UK there are several studies of peanut allergy in progress. The diagnostic criteria could require one or more of the

following: (i) clear history, consistent with an allergic reaction within a predetermined time after eating peanut; (ii) positive skin test (using test materials of a proven standard) and/or serum IgE antibodies to one or more peanut protein antigens; (iii) positive challenge test (open, single-blind, double-blind; with or without placebo controls); (iv) history of severe anaphylaxis produced by peanut; (v) atopic disease (e.g. eczema) and positive skin test or antibodies to peanut antigens; (vi) family history of atopy; (vii) personal history of other atopic conditions; and (viii) other relevant factors.

4.2. Problems of case ascertainment

As well as choosing an appropriate case definition, it is important that in any report of epidemiological research, the measures to be taken to search for potential cases are specified. Epidemiologists describe the process of searching various sources for cases as case ascertainment or case finding.

The problem for epidemiological studies is in distinguishing a higher frequency of disease in one country, social class, or other subgroup of the population, from a higher detection rate of 'mild' disease, perhaps because of the presence of an active research group, or because of high public awareness of food-related diseases in general.

One way of assessing differences in case ascertainment is to collect data on the mode of presentation of the cases. It is useful to collect data on symptoms and severity of the presenting illness as well as age and clinical features at diagnosis. Substantial increases in cumulative incidence in the presence of similar modes of presentation are likely to be real. Unfortunately, deciding whether differences in case ascertainment account for minor differences in incidence in place or time can be very difficult.

4.3. Pitfalls in the calculation and presentation of results

Having identified the cases to be included, it is necessary to relate them to the appropriate population. For a childhood disease such as milk allergy, in which the majority of cases present before age 3, it makes sense to relate the cases born in a defined period to the total number of births in that period. Adjusting the number of births for losses due to infant mortality, which mainly occurs in the first months, is a refinement which is scarcely necessary in developed countries, but which could be vitally important in the studies which should now be being carried out in the developing world, in populations where chronic diarrhoea remains a common cause of death in infancy.

Clearly, the total births need to be related to a defined geographic or catchment area and hospital-detected cases excluded if referred from beyond this area. All cases resident within the area at diagnosis should be included. Clinicians have an understandable but misplaced reluctance to omit any of their cases and may tend to overlook the need to exclude cases that do not meet the residency requirement.

Failure to allow for the varying periods of observation of different birth cohorts can cause problems. For example a direct comparison of the cumulative prevalence of a disease in children born in 1975 with that of children born in 1985 would fail to allow for the 1975 cohort being 10 years older. The pitfall is easily avoided by restricting comparisons to an age range that is common to both cohorts, such as age 0–5 years. While this does mean ignoring the contribution of cases diagnosed after age 5, in studies of food allergy the numbers involved are usually small.

Problems can also arise if mean age at diagnosis is calculated. In almost all series of childhood food allergy the distribution of age at diagnosis is extremely skewed, with the majority presenting in the first year or two of life. Thus relatively small increases in the numbers of older children diagnosed will produce a disproportionate increase in the mean age at diagnosis. Perhaps the most satisfactory approach to presenting data is to plot cumulative incidence by age (on the *X* axis). The resulting curve then demonstrates both the cumulative incidence at particular ages and the pattern with regard to age at diagnosis.

4.4. Problems of small numbers and wide confidence limits

An unavoidable limitation of most studies is the relatively small number of cases eventually identified.

5. Conclusion

Although this paper has concentrated on the problems posed by epidemiological studies, the importance of overcoming these problems should be obvious, and this continues to require interdisciplinary input into epidemiological, as well as clinical and laboratory, research in this area. Noteworthy is the relative lack of case control studies in food allergy, techniques which, as we have emphasised, can be extremely powerful research tools, and can, of course, be carried out without many of the problems incumbent in descriptive population research.

Acknowledgements

I am most grateful to Professor RFA Logan for his advice and comments on this paper.

References

- Logan, R.F.A., 1992. Problems and pitfalls in epidemiological studies of coeliac disease. In: Auricchio, S., Visakorpi, J.K. (Eds.), *Common Food Intolerances I: Epidemiology of Coeliac Disease*. Karger, Basel, pp. 14.
- Hennekens, C.H., Buring, J.E., 1987. *Epidemiology in Medicine*, chap. 2: Design Strategies. Little Brown, Boston, MA, pp. 16.