Send Orders for Reprints to reprints@benthamscience.net

The Open Allergy Journal, **2014,** *7,* **1-9**

Open Access

Evidence-based Medicine and Clinical Study Designs: Examples of Applications for Allergy Research

Simon F. Thomsen*

Center for Medical Research Methodology, Department of Biomedical Sciences, University of Copenhagen, Blegdamsvej 3B, DK-2200 Copenhagen N, Denmark

Abstract: Evidence of the effect of clinical interventions in allergology, and in medicine as a whole, can be hierarchically grouped based on the research design producing the evidence. The most weight is given to systematic reviews and meta-analyses, and to randomised controlled trials. These trial designs are superior to non-randomised controlled trials and cohort studies, which in turn are superior to case-control studies. The least weight is given to case-studies and anecdotal evidence. Herein, the principles of evidence-based medicine and clinical study designs are reviewed in the context of examples from the allergology literature.

Keywords: Case-control study, controlled trial, cohort study, Evidence-based medicine, meta-analysis, randomised, study designs, systematic review.

EVIDENCE-BASED MEDICINE

Clinical decisions should be based on the best available evidence. But when is the available evidence sufficient to subject a patient to a certain treatment, or even to recommend that treatment to the patient population as a whole? It is ethically unacceptable, both to the patient and society, to begin a new treatment if the effect of that treatment is undocumented. Accordingly, it is necessary to first conduct trials involving human patients to show that the new treatment has a positive effect that exceeds what is gained by merely observing the natural course of the disease or by giving any customary treatment for that disease. Moreover, it is also necessary to show that the new treatment has minimal side effects that do not outweigh its benefits. Lastly, in clinical practice, the physician is obliged not only to consult the best available evidence but also to weigh this evidence against the individual patient's personal and social circumstances. These concepts constitute evidence-based medicine and protect the patient from being given ineffective or even harmful treatments [1]. Nonetheless, this *conservatism* of the medical research society tends to hamper the introduction of new drugs. Moreover, the number of candidate compounds generated is huge compared to the small percentage that eventually reaches the patient. Every potentially new treatment must be subjected to the rigorous rules that apply to evidence-based medicine. In allergology, as in other areas of medicine, many complementary practices are being used for symptom control or even cures, but most of these are insufficiently documented or have been shown to have no benefits compared with placebo or conventional treatment.

*Address correspondence to this author at the Center for Medical Research Methodology, Department of Biomedical Sciences, University of Copenhagen, Blegdamsvej 3B, DK-2200 Copenhagen N, Denmark;
Tel: +45 2613 9838; Fax: +45 3531 3113;
E-mail: simonfrancisthomsen@gmail.com

HIERARCHY OF EVIDENCE

Evidence of the effect of clinical interventions can be hierarchically grouped based on the research design producing the evidence [2] (Table **1**).

This implies that the conclusion of a scientific study has more weight if the study is conducted in a certain way [3]. The most weight is given to randomised controlled trials (RCTs). RCTs are superior to non-randomised controlled trials and cohort studies, which in turn are superior to case-control studies. The least weight is given to case-studies and anecdotal evidence or so-called expert opinions as these are based only on personal feeling/preference and not on an experiment designed specifically to test the effect of the treatment. If many studies, irrespective of the weight or size, show a positive effect of a treatment, this speaks in favour of recommending that treatment. However, evaluation of these studies must be conducted in a systematic and unbiased manner so that all available evidence is gathered. This ensures that not only evidence from 'positive' studies, for example favouring a certain drug over another, is evaluated. For this reason systematic reviews are ranked the highest in the hierarchy of evidence in medical research, and a systematic review based on many RCTs therefore ranks above single RCTs.

In allergology - as in other areas of medicine - basic laboratory and animal experimental research provide an important backdrop or supplement to clinical and population studies in humans. The evidence from basic research including animal studies forms a basis for translational research in humans but cannot serve as the sole source of recommendation for medical therapy in human patients.

STUDY DESIGNS

Systematic Reviews and Meta-Analyses

A systematic review is an exhaustive evaluation of all previous studies concerning a specific medical intervention

1874-8384/14

Table 1. Hierarchy of evidence in medical research.

Type of Study	Level of Evidence	Recommendation
Systematic reviews and meta-analyses of RCTs	Ia	A
Randomised controlled trials	Ib	
Non-randomised controlled trials	IIa	B
Cohort studies	IIb	
Case-control studies	III	C
Case-reports and expert opinions	IV	D

This method of ranking puts scientific studies into four (I-IV) categories based on their reliability in design and attempts to minimise bias. Each category corresponds to a recommendation (A-D) for clinical interventions so that studies with the highest quality translate into the strongest recommendation.

or scientific question [4]. It often uses a meta-analysis to combine the results of all these studies and draws a general conclusion about the effect of an intervention. The systematic review prioritizes studies and includes only those with sound methodology, omitting studies with flaws or major drawbacks in their design. By combining the results from all available and methodologically sound studies, additional weight is put on the conclusion because the 'new' meta-analytic study has more statistical power to detect beneficial or harmful effects. The Cochrane Collaboration is an international organisation of researchers that conducts systematic reviews (Cochrane Reviews) published in the Cochrane Library [5]. Cochrane reviews are by many believed to be the most reliable source of evidence on which to base medical decisions. Many practices within allergology have been evaluated in Cochrane reviews. For example, Cochrane reviews have concluded that: based on 12 RCTs, 'probiotics are not an effective treatment for childhood eczema' [6]; and based on 49 RCTs, 'sublingual immunotherapy is effective for allergic rhinitis and has been proven to be a safe route of administration' [7]; and based on 12 trials, 'use of systemic corticosteroids within 1 hour of presentation to an emergency department significantly reduces the need for hospital admission in patients with acute asthma' [8]. Although systematic reviews and meta-analyses are considered to provide the strongest evidence, critics have argued that they have several drawbacks, for example, the inability to control confounding and the constant need for update when new evidence appears. In addition, they are limited by the fact that negative studies, i.e., studies that show no effect of an intervention, tend not to be published and therefore cannot be included in the analysis.

The Randomised Controlled Trial (RCT)

A randomised controlled trial is an experiment that randomly assigns a patient population to two or more groups that undergo an intervention [9]. In the classical randomised, double-blind, placebo-controlled trial the patients in one of the groups receive the treatment of interest (the 'new' treatment), while the patients in the comparison group receive the 'standard' treatment or no treatment at all, i.e., placebo treatment.

For example, in a multinational, randomised, double-blind, placebo-controlled study of 278 children with grass-pollen-related allergic rhinoconjunctivitis patients were randomised 1:1 to two groups: one group received once-daily sublingual immunotherapy before and during the pollen season with allergen extract of five grass pollens in a tablet formulation, and the other group received placebo [10]. The placebo tablet corresponded to the active treatment in terms of size, shape, and colour but contained no active components. The study showed that compared with the placebo group, the actively treated group showed a mean improvement of 28% in their symptom score during the pollen season. Furthermore, actively treated patients experienced only few and mild side effects and used less rescue medication during the pollen season compared with patients in the placebo group.

Furthermore, another randomised, double-blind, placebo-controlled, multicenter study of peanut sublingual immunotherapy in 40 patients, 12-37 years of age, showed that, after 44 weeks of treatment, 70% of the patients receiving peanut sublingual immunotherapy were responders compared with only 15% of the patients receiving placebo [11]. Furthermore, the active treatment proved to be safe in these patients.

Finally, in a randomised, double-blind, placebo-controlled study, 18 children, 5-10 years of age, with perennial allergic rhinoconjunctivitis and who were monosensitized to house dust mite (Dermatophagoides pteronyssinus and Dermatophagoides farinae), received either active sublingual immunotherapy or placebo [12]. Although the active treatment was well-tolerated and reduced nasal sensitivity and skin test reactivity it was not superior to placebo in reducing isolated rhinoconjunctivitis symptoms within 12 months of treatment.

RCT Designs

The simplest RCT design is the 'parallel group design' where two - usually identically sized - groups are followed over a certain period. One group is termed the study group or the intervention group and is given the treatment under study (treatment A); the other group is termed the control group or

the placebo group, if a placebo treatment is given (treatment B). The 'cross-over design' is a design where the order of treatment in the two groups is reversed after a certain time so that treatment A is given first in one group followed by treatment B, while treatment B is given first in the other group followed by treatment A. This design can be useful if the number of eligible patients is small. Possible problems with the cross-over design are the so-called *period effect* and the *carry-over effect*. The period effect is based on the fact that many chronic diseases tend to better cyclically regardless of treatment and therefore the study drug given first apparently has the greatest effect. The carry-over effect is seen when the effect of one of the treatments continues after the crossing-over, thereby diluting the difference between treatments. A third type of design is the 'factorial design' where, for example, four groups of patients are given treatment A, treatment B, treatment A + treatment B, or placebo treatment. Factorial designs can also be where different doses of a study drug are given to different groups of patients.

Randomisation

Correct randomisation is the most important feature of RCTs as it ensures that any difference in treatment response between the intervention group and the placebo group at the end of the study period can be ascribed to the intervention and not to other factors. Randomisation ensures a chance allotment of known and unknown confounding factors to each study group. This means that the different groups in the study are comparable and that the intervention is the only discriminating factor between groups. Correct randomisation is secured via computer-generated allocation of patients to each study group. Incorrect randomisation, which would introduce bias, includes, for example, allocation of patients by date of examination, month of birth, home address or even researchers' preferences.

Blinding

Blinding is the procedure that ensures the researcher and/or the patient (double/single blinded) are unaware of which treatment is given to which study subject, irrespective of whether the treatment is active or placebo. Triple blinding of a clinical trial is blinding of the analyst or the evaluator of the trial's data in addition to the blinding of the researcher and the patients. By blinding, all bias arising from the researcher's belief in the superiority of one of the interventions is eliminated and the trial meets so-called *clinical equipoise*, which implies a genuine uncertainty over which treatment will give the most benefit. A researcher who knows that a trial patient is receiving active treatment may be more inclined to offer the patient additional care to make the study drug seem superior. A patient who knows which treatment he or she is receiving may be more or less inclined to report symptoms or withdraw from the trial. In drug trials blinding is ensured with a placebo drug that resembles the study drug in appearance, smell, taste, etc. Sufficient blinding is not always achievable, for example, in trials involving surgical or invasive procedures; nevertheless, as much as possible should be done to ensure optimal blinding.

Selection of the Appropriate Endpoint

The optimal endpoint of an RCT is intuitively meaningful and should preferably be directly transferable to the daily clinical setting. This means that any given intervention should have a direct, positive influence on the patient's disease activity in terms of symptom relief or quality of life. An example of a clinically relevant endpoint is severity of symptoms, such as frequency of wheezing, shortness of breath, sneezing, intensity of itchiness, and ability to carry on with daily activities. Composite endpoints are composed of different clinical or paraclinical measures that jointly grasp a common entity, such as an asthma exacerbation.

Surrogate endpoints are biomarkers of disease activity that may or may not be directly related to the presence or the severity of the disease. Examples of surrogate endpoints are serum IgE, lung function, airway responsiveness, and eosinophil count. These endpoints are typically easier than the other endpoints to measure objectively and are sometimes more closely related to the action of a drug and may therefore be the first parameter that shows a change in relation to a specific treatment. However, it is important to remember that an improvement in a surrogate endpoint does not always reflect an equal improvement in disease activity. Results of randomised trials where surrogate endpoints are used as primary endpoints should therefore not be used as the only evidence on which a drug is introduced as treatment.

Termination of an RCT and Reporting of Findings

The termination of an RCT must be predetermined. If the researchers are unblinded, they may want to terminate the trial when one of the treatments shows, possibly by chance, to be superior to the other. Before initiation of a trial, researchers should estimate the number of patients needed to show a significant effect of the treatment because it is unethical to initiate a trial based on an estimation requiring more patients than can be allocated to the study.

A study finding can be statistically significant without being clinically significant and vice versa. For example, a large study may find that taking inhaled corticosteroids for mild intermittent asthma for three years compared with taking no treatment results in a statistically significantly higher lung function in the treatment group compared with that in the placebo group by the end of the study. However, if the difference in lung function is extremely small, it is not worth the effort (Table **2**).

Cohort Studies

A cohort study is an epidemiological study design that involves a large population of individuals, typically a random sample of subjects from the background population [13, 14]. Other types of cohort study include groups of patients or individuals with a certain characteristics, for example, all babies born to allergic mothers in a well-defined geographic area. A cohort study can be prospective or retrospective and include various measurements on the individual such as questionnaires, clinical tests, register data and paraclinical tests, such as blood samples for cell counts and DNA analyses.

Table 2. Evaluation of randomised controlled trials.

Checkpoint
Is the problem clinically relevant and well defined?
Were the treatment- and control-groups comparable at the beginning of the trial?
Did the groups receive the same care apart from the intervention?
Was the blinding sufficient?
Was the randomisation sufficient?
Are the inclusion and exclusion criteria relevant and understandable?
Were drop-outs of the trial sufficiently described?
Was the primary endpoint (clinically) relevant?
Was the follow-up time of adequate length?
Was the effect of the intervention convincing?
Were the safety and side-effects of the intervention evaluated?
Are the conclusions of the study in accordance with the results?
Was the effect attributable to the intervention?
Is the study biased?
Are the results relevant for the patient?

The above list of checkpoints provides a basis for evaluating various aspects of a randomised controlled trial.

The Prospective Cohort Study

A prospective cohort study follows a group of individuals over a period of time and records new cases of a disease. The risk of the disease is measured in the context of different (environmental) exposures recorded before the onset of the disease, and risk factors for the disease of interest are identified. The influence of the identified risk factors is typically expressed as a relative risk in the exposed compared with non-exposed subjects. For example, in a prospective questionnaire study of 121,700 healthy adult women from the United States, the risk of asthma within a four-year period was increased by 63% in women who received acetaminophen for more than 14 days per month compared with non-users, indicating a role of acetaminophen in the development of adult-onset asthma [15].

The Retrospective Cohort Study

A retrospective cohort study - unlike the prospective cohort study - examines individuals after the outcome and exposures of interest have occurred. This study design therefore 'looks back' in time and tries to explain why certain subsets of individuals within the cohort became diseased.

Advantages and Disadvantages of Cohort Studies

A chief advantage of cohort studies is the ability to assess an exposure before the outcome of interest has occurred [16]. This secures temporality in the study and is crucial to minimise bias due to patients' recall. A second advantage is the large number of subjects that can be included in the study. This need for a large study population, however, is one of the weaknesses of cohort studies since it is time-consuming and expensive. A cross-sectional study is a type of cohort study that examines a study population on only one occasion. If representativeness of the cohort is achieved, this type of study design is useful for estimating the prevalence of certain common diseases, for example, asthma or hay fever or intermediate disease states, such as skin test positivity and airway hyperresponsiveness (Table **3**).

The Case-Control Study

The case-control study is - like a cohort study - an observational study design. It compares a group of patients (the cases) with one or more groups of matched individuals without the disease (the controls) in order to identify differences in exposures between cases and controls. For example, in a large case-control study from Sweden involving ~13,000 cases and ~800,000 control subjects the risk of asthma was 9% in children delivered by Caesarean section but only 7% in children from vaginal deliveries, thereby highlighting a possibly detrimental role of Caesarean section, or other factors associated with this type of delivery, in the aetiology of asthma [18].

The case-control study is often retrospective but may also be conducted prospectively. The key to conducting a good case-control study is the selection of appropriate control subjects to match the cases. Ideally, the control subjects should be 'healthy cases' thereby matching the cases in terms of different demographic factors. The appropriate control subjects differ between studies and are typically selected from a random background population, hospital records, blood donors or relatives to the cases. Conversely, selected cases should be homogeneous and reflect the population of patients as a whole.

Advantages and Disadvantages of Case-Control Studies

The main advantage of case-control studies is the ability to study rare outcomes and relate them to a variety of exposures. Another advantage lies in the relatively simple and achievable design. The main disadvantage is the high risk of

Table 3. Bradford Hill's criteria of causation.

Criterion	Explanation
Temporality	Exposure must precede the outcome
Strength	A strong association between exposure and outcome supports causality
Plausibility	Findings should be biologically plausible
Consistency	Findings should be reproducible in other settings such as in other age groups and countries and by different researchers
Gradient	Clear dose-response relationship between exposure and outcome supports causality
Specificity	Direct and simple association between exposure and outcome supports causality
Experiment	The association can possibly be experimentally induced or counteracted
Analogy	Alternative explanations should be ruled out
Coherence	The association follows existing theory

In 1965, Sir Austin Bradford Hill (British epidemiologist) formulated a group of minimal conditions necessary to provide adequate evidence of a causal relationship between an exposure and an outcome [17].

Table 4. Advantages and disadvantages of different study designs.

Type of Study	Advantages	Disadvantages
Systematic review	Evidence can be based on the highest quality studies	Often not possible to adjust for confounders
	High statistical power	Relies on published studies (and their quality)
	Exhaustive method	Needs update when new evidence emerges
Randomised controlled trial	Gold standard within medical research	Expensive and time-consuming
	Address causality (least biased)	Results applicable only within the specific frame
		Ethics: not all questions suitable for RCTs
Cohort study	If prospective, temporality is secured	Expensive if clinical data are collected
	Large sample size	Time-consuming
	Possible to control for confounders	
Case-control study	Economically feasible and fast	Retrospective (does not address causality)
	Requires few test subjects	Risk of recall-bias
	Possible to study rare outcomes	Non-randomised (lack of confounder-control)
	Provides guidance for later interventions	Estimates the relative risk, not the absolute

bias due to the retrospective nature of the study. For example, an adult with asthma who is asked whether her/his parents smoked when she/he was young could be more inclined to over-report such exposure because it is putatively causal.

Case-Report and Expert Opinions/Anecdotal Evidence

Case-reports constitute the weakest type of evidence on which to base clinical decisions. Case-reports are detailed descriptions of a beneficial (or harmful) treatment of a single patient or a small group of patients with, for example, a type of drug that is normally not used for treating that particular condition. Case-reports can form the basis for future randomised studies.

Expert opinions usually built on one (senior) physician's personal belief of superiority of one treatment over another. This belief can be scientifically rigorous and important for the medical community as a whole but it can also be rooted in several anecdotal or idiosyncratic principles or be biased by conflicts of interest.

Other Study Designs

A range of other study designs are encountered in the medical literature, for example, studies that validate diagnostic tests or scoring systems [19]; studies that test the variation between different examiners' assessment of a clinical problem or reading of an objective test; or studies concerned with experimental procedures, such as biochemical and molecular characterisation of biological processes. For an overview of advantages and disadvantages of various study designs, please see Table **4**.

CONFLICT OF INTEREST

The author confirms that this article content has no conflict of interest.

ACKNOWLEDGEMENTS

Declared none.

REFERENCES

[1] Elphick HE, Smyth RL. Research: the principles of evidence-based medicine. Curr Paediatr 2004; 14: 525-31.

[2] Balshem H, Helfand M, Schünemann HJ, *et al.* GRADE guidelines: 3. Rating the quality of evidence. J Clin Epidemiol 2011; 64: 401-6.

[3] Atkins D, Best D, Briss PA, *et al.* Grading quality of evidence and strength of recommendations. BMJ 2004; 328: 1490.

[4] Liberati A, Altman DG, Tetzlaff J, *et al.* The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. J Clin Epidemiol 2009; 62: e1-34.

[5] The Cochrane Collaboration (www.cochrane.org).

[6] Boyle RJ, Bath-Hextall FJ, Leonardi-Bee J, Murrell DF, Tang ML. Probiotics for treating eczema. Cochrane Database Syst Rev 2008; 4: CD006135

[7] Radulovic S, Calderon MA, Wilson D, Durham S. Sublingual immunotherapy for allergic rhinitis. Cochrane Database Syst Rev 2010; 12: CD002893.

[8] Rowe BH, Spooner C, Ducharme FM, Bretzlaff JA, Bota GW. Early emergency department treatment of acute asthma with systemic corticosteroids. Cochrane Database Syst Rev 2001; 1: CD002178.

[9] Moher D, Hopewell S, Schulz KF, *et al.* CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. Int J Surg 2012; 10: 28-55.

[10] Wahn U, Tabar A, Kuna P, *et al.* Efficacy and safety of 5-grass-pollen sublingual immunotherapy tablets in pediatric allergic rhinoconjunctivitis. J Allergy Clin Immunol 2009; 123: 160-.e3.

[11] Fleischer DM, Burks AW, Vickery BP, *et al.* Sublingual immunotherapy for peanut allergy: a randomized, double-blind, placebo-controlled multicenter trial. J Allergy Clin Immunol 2013; 131: 119-27.

[12] Aydogan M, Eifan AO, Keles S, *et al.* Sublingual immunotherapy in children with allergic rhinoconjunctivitis mono-sensitized to house-dust-mites: a double-blind-placebo-controlled randomised trial. Respir Med 2013; 107: 1322-9.

[13] 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.

[14] Von Elm E, Altman DG, Egger M, *et al.* The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. J Clin Epidemiol 2008; 61: 344-9.

[15] Barr RG, Wentowski CC, Curhan GC, *et al.* Prospective study of acetaminophen use and newly diagnosed asthma among women. Am J Respir Crit Care Med 2004; 169: 836-41.

[16] Henneskens CH, Buring JE. Edited by Mayrent SL. Epidemiology in medicine. Little, Brown and Company: Boston/Toronto, 1987.

[17] Hill AB. The environment and disease: association or causation? Proc R Soc Med 1965; 58: 295-300.

[18] Håkansson S, Källén K. Caesarean section increases the risk of hospital care in childhood for asthma and gastroenteritis. Clin Exp Allergy 2003; 33: 757-64.

[19] Rennie D. Improving reports of studies of diagnostic tests: the STARD initiative. JAMA 2003; 289: 89-90.

Received: February 28, 2014 Revised: July 07, 2014 Accepted: July 15, 2014