

CHAPTER 1

Methodology of evidence-based medicine

Jeremy Allgrove

When one admits that nothing is certain one must, I think, also admit that some things are much more nearly certain than others.

—Bertrand Russell. ‘Am I an Atheist or an Agnostic?’ 1947
British author, mathematician and philosopher (1872–1970)

Introduction

Over the past two decades evidence-based medicine has become increasingly important in determining the way in which medicine is practised. The medical profession has always had a reputation for questioning its own practices, as demonstrated by the number of scientific publications that have appeared since medical journals were invented. As a result, considerable advances in health care have been achieved.

Nevertheless, it is not always the case that ideas that have developed are necessarily correct, and dogmatic statements or assumptions that have been made have sometimes turned out to be false when re-examined more rigorously. Although it has been suggested that ‘it is curious, even shocking, that the adjective “evidence-based” is needed’ [1], it is nevertheless the purpose of evidence-based medicine to limit these false assumptions and incorrect dogma so that patients may be treated in the best possible way with the tools available.

What is evidence-based health care?

The Cochrane library [2] quotes three slightly different definitions of evidence-based health care:

- Evidence-based health care is the conscientious use of current best evidence in making decisions about the care of individual patients or the delivery of health services. Current best evidence is up-to-date information from relevant, valid research about the effects of different forms of health care, the potential for harm from exposure to particular agents, the accuracy of diagnostic tests and the predictive power of prognostic factors [3].
- Evidence-based clinical practice is an approach to decision-making in which the clinician uses the best evidence available, in consultation with the patient, to decide upon the option which suits that patient best [4].
- Evidence-based medicine is the conscientious, explicit and judicious use of current best evidence in making decisions about the care of individual patients. The practice

of evidence-based medicine means integrating individual clinical expertise with the best available external clinical evidence from systematic research [5].

All of these definitions are very similar but differ slightly in emphasis on such matters as patient involvement and reliance on diagnostic tests.

What constitutes proof?

Scientific proof has always depended on probabilities rather than absolute proof and is determined by observation and perception. Both of these are open to misinterpretation and can be refuted by other observations that may be made under different circumstances. Statistical analysis is frequently used to ‘verify’ observations and it has become usual practice to accept that a probability of something being true with 95% certainty ($p < 0.05$) means that observation is ‘true’. By definition, it also means that there is a 5% chance that it will not be true.

In contrast, there is a fundamental difference between a scientific proof and a mathematical proof [6, pp. 21–2]. In the latter, proof is absolute and remains so forever. If proof is not absolute, i.e. if a flaw can be found in the logic, then proof does not exist. A simple example of this is the proof of the well-known formula of Pythagoras:

$$a^2 + b^2 = c^2$$

where a , b and c are the values of the sides of a right-angled triangle, c being the hypotenuse. The proof of this theorem is straightforward [6, pp. 333–4] and it can be shown that the relationship is true under *all* circumstances. Thus, if the values of any two numbers are known, the third can always be calculated.

However, this relationship can be rewritten as:

$$a^x + b^x = c^x$$

where the value of x is any whole number greater than 2. The French mathematician Pierre de Fermat (1601–1665) postulated that there is *no* solution to this equation. This has become known as Fermat’s last theorem. He died having claimed that he had found a proof that there is no solution, but the proof was lost and the challenge to rediscover it became the most exciting in the field of mathematics for the next 329 years until finally solved by Andrew Wiles in 1994.

Fermat’s last theorem is fiendishly difficult to prove. Initial attempts resulted in proofs that the postulate is true for values of $x = 4$ and $x = 3$. The problem is that even if it is possible to show that for all values between, say, 3 and 1000 the postulate is also true, this does not prove the theorem, as there could still be values greater than 1000 that do satisfy the equation. This is shown by another conjecture, that of the Swiss mathematician Leonhard Euler, which states that there are also no solutions to the equation:

$$x^4 + y^4 + z^4 = w^4$$

Initial attempts to solve it proved fruitless and the lack of a counter-example was taken as proof of its truth until a solution* was eventually found in 1988 some two centuries after it was postulated [6]. Therefore, Euler’s postulate is absolutely not true in mathematical

* $2,682,440^4 + 15,365,639^4 + 18,796,760^4 = 20,615,673^4$

terms, although in scientific terms it had been taken to be so. Thus, to obtain an absolute proof, it is necessary to go back to first mathematical principles and demonstrate that the conditions apply to *all* numbers.

Scientific proof is not so rigorous and only demands that there is a sufficient body of evidence to suggest very strongly that a fact is 'true'. Medicine is no different in this respect from other scientific disciplines and, particularly because one is dealing with a biological rather than a physical system, is particularly open to variations in response. The most rigorous method available to scientists, in the realm of medicine, for determining the effectiveness of a treatment is the double-blind, placebo-controlled trial, properly conducted under clearly defined conditions with sufficient numbers of patients and with removal of bias. Some treatments have fulfilled these criteria, although others that are regularly used have never been tested under such circumstances. There has, for instance, never been such a trial of the use of insulin in type 1 diabetes mellitus (T1DM). It would, of course, be totally unethical to conduct such a trial now and yet there is little or no doubt that insulin therapy is effective in treating T1DM. The statement 'insulin is an effective treatment of T1DM' is taken to be true. Evidence-based medicine depends upon scientific observation rather than mathematical proof and is always open to some degree of doubt, however small. It is therefore necessary to have some means of gauging how reliable a piece of evidence is in scientific terms.

Grading of evidence

Several methods of grading evidence have been used and different guideline development groups (GDGs) have used different methods of classifying evidence. The classification used by the Scottish Intercollegiate Guideline Network (SIGN) is the most detailed [7]. The 'levels of evidence' are then converted into 'grades of recommendation' (A–D). In addition, they list 'good practice points' (GPPs).

The National Institute for Clinical Excellence (NICE), an independent body set up by the UK Department of Health, uses a similar, though not quite so detailed, classification [8]. It gives grades A–D and GPPs, and also recommendations from NICE technology appraisals.

The American Diabetes Association (ADA) has the simplest classification. This does not describe a level of evidence which is then converted into a grade but assigns a grade directly to a study [9]. The classification is shown in Table 1.1.

All of these grading methods are similar but, since this book is not designed to be another guideline but rather to present the evidence, we have chosen to use the ADA classification which does not include any GPPs, etc. The new International Society for Pediatric and Adolescent Diabetes (ISPAD) guidelines also use the same gradings. Where relevant, gradings have been assigned to references within the text.

Guidelines

Since the beginning of the 1990s there has been a move away from professional consensus towards more rigorous scientific methods, such as systematic reviews and meta-analyses [10]. This has usually been done in the context of creating guidelines, although the quality of these guidelines has varied depending on how rigorously the methodology has been applied. In 2003, Burgers *et al.* published a study, on behalf of the Appraisal of Guidelines,

Table 1.1 ADA evidence grading system for clinical practice recommendations

Level	Description
A	<p>Clear evidence from well-conducted, generalisable, randomised controlled trials that are adequately powered, including:</p> <ul style="list-style-type: none"> • evidence from a well-conducted multicentre trial • evidence from a meta-analysis that incorporated quality ratings in the analysis • compelling non-experimental evidence, i.e. 'all-or-none' rule developed by Centre for Evidence-Based Medicine at Oxford <p>Supportive evidence from well-conducted randomised controlled trials that are adequately powered, including:</p> <ul style="list-style-type: none"> • evidence from a well-conducted trial at one or more institutions • evidence from a meta-analysis that incorporated quality ratings in the analysis
B	<p>Supportive evidence from well-conducted cohort studies:</p> <ul style="list-style-type: none"> • evidence from a well-conducted prospective cohort study or registry • evidence from a well-conducted meta-analysis of cohort studies • supportive evidence from a well-conducted case-control study
C	<p>Supportive evidence from poorly controlled or uncontrolled studies:</p> <ul style="list-style-type: none"> • evidence from randomised clinical trials with one or more major or three or more minor methodological flaws that could invalidate the results • evidence from observational studies with high potential for bias (such as case series with comparison to historical controls) • evidence from case series or case reports <p>Conflicting evidence with the weight of evidence supporting the recommendation</p>
E	Expert consensus or clinical experience

Note: There is no Grade **D**.

Research and Evaluation for Europe (AGREE) study group [11], in which they described the structures and working methods of 18 national GDGs from 13 different countries worldwide. These did not include guideline development by NICE since this organisation was formed only in 1999 and produced its first report in 2002. They concluded that '*principles of evidence-based medicine dominate current guideline programs*'. As a result, it can be concluded that most of the current guidelines that have been developed are reasonably well evidence based and well referenced.

However, this is not always the case. For instance, the Consensus Guidelines for the Management of Type 1 Diabetes in Children and Adolescents published by ISPAD in 2000 contained no references. It raises the question of how truly evidence-based they were and how much they depended on the views and opinions of the guideline development team. Having said that, they have proved invaluable as a resource. The situation is due to be rectified with the publication of the new ISPAD Clinical Practice Consensus Guidelines 2006/2007, which are heavily referenced. The first two chapters were published in 2006 (**E**) [12, 13], with the rest due to be published in 2007.

Bertrand Russell is quoted as saying [14], '*The fact that an opinion has been widely held is no evidence whatever that it is not utterly absurd; indeed in view of the silliness of the majority of mankind, a widespread belief is more likely to be foolish than sensible*'. Although he was referring to marriage, he could as easily have been referring to clinical guidelines. That is not to say that guidelines should not be followed, but it must be understood that, whilst

they are usually well researched, there are often aspects of the guidelines that are based solely on the personal opinions of those drawing them up with little or no hard evidence to support them and there may be individual circumstances where they do not necessarily apply.

There may also be a tendency, in some instances, for recommendations to be 'transferred' from one guideline to another by default. Let us examine, as an example, the statement made in all of the major national and international guidelines for the treatment of diabetic ketoacidosis (DKA) in children that the dose of insulin should be '0.1 unit per kilogram body weight per hour' (E) [8, 15–18]. The British Society for Paediatric Endocrinology and Diabetes (BSPED) guidelines (E) [16] state that '*Modifications (to their previous guideline) have been made in the light of the guidelines produced by the International Society for Pediatric and Adolescent Diabetes (2000) and the recent ESPE/LWPES consensus statement on diabetic ketoacidosis in children and adolescents*', and the NICE guidelines (E) [8] say that '*The current guidelines take account of recently published consensus statements developed by the European Society for Paediatric Endocrinology and the Lawson Wilkins Pediatric Endocrine Society. The guidelines highlight the need for further research to investigate the effectiveness of different concentrations of rehydration fluid, the rate of rehydration and the concentration of insulin infusion in the management of diabetic ketoacidosis*'. The implication of these two statements is that they are merely following previous recommendations and have not re-examined the evidence.

Despite claims to the contrary [15], the evidence for the stated dose of insulin is weak. The Lawson Wilkins Pediatric Endocrine Society/British Society of Paediatric Endocrinology and Diabetes (LWPES/BSPED) guidelines state that '*Physiologic studies indicate that IV insulin at a dose of 0.1 unit/kg per hour, which achieves steady state plasma insulin levels of ~100 to 200 μ U/mL within 60 minutes, is effective*'. However, as stated by Edge and Spinks in Chapter 4 of this book, '*there is a body of opinion that a dose of 0.05 units/kg/hour is sufficient to reverse the metabolic abnormalities and overcome any insulin resistance whilst reducing the blood glucose at a steadier rate*', and many units in the UK ignore the national and international guidelines and routinely use this lower dose.

The statement, which is given an A grading, is based on a study conducted in six adults with established diabetes who were rendered ketotic by the administration of two doses of dexamethasone and cessation of insulin in the 24 hours prior to the study [19]. They were then given insulin infusions at varying rates (0.01, 0.1 and 1 U/kg/h) in random order. Steady-state levels of insulin were measured and the rates of fall of glucose and ketones, as measured by β -hydroxybutyric acid and acetoacetate, observed with the different doses. The principal conclusions were as follows:

- 1 An infusion rate of 0.1 U/kg/h achieves a steady-state insulin concentration between 100 and 200 μ U/mL (an increase between 90 and 112 μ U/mL over baseline).
- 2 Logarithmic increases in infusion rates resulted in logarithmic increases in insulin concentration.
- 3 The effect of insulin on reducing ketones was maximal at 0.1 U/kg/h but the effect on reducing blood glucose had no such plateau effect; i.e. the rate of fall of blood glucose continues to increase with larger doses of insulin.

Unfortunately, an infusion rate of 0.05 U/kg/h was not tested but it can be deduced from the above that this lower rate of infusion would be likely to result in a steady-state concentration of insulin of ~55 μ U/mL, which may well be sufficient to switch off ketogenesis (the principal aim of insulin therapy in the treatment of DKA) whilst reducing

the rate of fall of blood glucose. This is supported by another study, also conducted in adults [20], and also quoted in the LWPES/BSPED guidelines, in which patients with newly diagnosed diabetes were admitted with DKA and treated with insulin at a rate of 1 mU/kg/min ($\equiv 0.06$ U/kg/h). This resulted in a steady fall in blood glucose at an acceptable rate of 3.3 mmol/L/h and correction of the acidosis.

In some units it is considered important to control the rate of fall of blood glucose with the use of systems that involve the use of solutions of different strengths of dextrose, used at different rates depending upon circumstances, a situation that arguably increases the risk of error. Even so, in one such study [21], which was conducted in children, the recommended dose of 0.1 U/kg/h was used and the blood glucose fell initially, when no glucose was being infused, by approximately 33 mmol/L in the first 5 hours (6.6 mmol/h), a rate which is now regarded as being too rapid. Although there is little evidence to support it, a maximum of 5 mmol/L/h is recommended by the ISPAD guidelines (E) [17].

It is therefore clear that the evidence for the recommended dose of insulin is weak and has never been properly tested in children. It is possible that this dose *is* correct (although it may be different at different ages) but, as stated in the NICE guidelines, '*further research to investigate the effectiveness of different concentrations of . . . insulin infusion in the management of diabetic ketoacidosis*' is required (see above). Evidence-based medicine should ultimately be able to provide an answer.

Guidelines are widely quoted throughout this book and in many instances, the recommendations are clearly evidence based and have a high degree of validity. Nevertheless, in view of the fact that they are all consensus documents, they are always given an E grading. Whilst there is clearly a hierarchy of validity between A and C, an E grading does not necessarily mean that this is the lowest level since consensus documents do often contain systematic reviews or meta-analyses, which, under other circumstances, might be rated A. Having said that, some C-graded articles, particularly those that are case reports, may still carry quite a lot of weight if they contain, for instance, convincing genetic data.

Sources of data

Electronic databases, such as MEDLINE, have proved enormously helpful in searching for relevant studies. Not only do they make the searches much faster than previously, but they are inevitably more thorough. We have made use of all the available databases including:

- Allied & Complementary Medicine – 1985 to date
- British Nursing Index – 1994 to date
- CINAHL (R) – 1982 to date
- DH-DATA – 1983 to date
- EMBASE – 1974 to date
- King's Fund – 1979 to date
- MEDLINE – 1950 to date
- PsycINFO – 1806 to date.

These have all been available either via KA24, the National Health Service (NHS) portal available to NHS employees (accessible via <http://www.hilo.nhs.uk/> to registered personnel) [22], or via PUBMED, a service of the National Library of Medicine and the National Institutes of Health (accessible via <http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?CMD=Pager&DB=pubmed>).

In addition, the relevant Cochrane databases have been examined. These are a series of systematic reviews based on available publications and are also available via <http://www.hilo.nhs.uk/> [22]. (This requires no special permissions.) Cochrane describes a systematic review as follows:

- To help identify which forms of health-care work, which do not and which are even harmful, Results from similar randomised trials need to be brought together. Trials need to be assessed and those that are good enough can be combined to produce both a more statistically reliable result and one that can be more easily applied in other settings. This combination of trials needs to be done in as reliable a way as possible. It needs to be systematic. A systematic review uses a predefined, explicit methodology. The methods used include steps to minimise bias in all parts of the process: identifying relevant studies, selecting them for inclusion and collecting and combining their data. Studies should be sought regardless of their results.
- A systematic review does not need to contain a statistical synthesis of the results from the included studies. This might be impossible if the designs of the studies are too different for an averaging of their results to be meaningful or if the outcomes measured are not sufficiently similar. If the results of the individual studies are combined to produce an overall statistic, this is usually called a meta-analysis. A meta-analysis can also be done without a systematic review, simply by combining the results from more than one trial. However, although such a meta-analysis will have greater mathematical precision than an analysis of any one of the component trials, it will be subject to any biases that arise from the study-selection process and may produce a mathematically precise, but clinically misleading, result.

The Cochrane databases deal mainly with adult practice and have little relevance to paediatrics. There is only one systematic review relating directly to children listed on their website [23]. Nevertheless, the principles of systematic reviews and meta-analyses are important and apply equally to children as to adults.

Summary and conclusions

Evidence-based medicine is becoming increasingly important in determining how best patients should be treated. There is an element of cost-effectiveness built into the system but this is not the principal aim of the process. Unfortunately, in paediatric practice, there is a certain paucity of studies in many areas and it has been necessary to rely on studies in adults which are then extrapolated into paediatrics. Whilst this is valid in some areas, it may not be so in others and one has to retain a certain degree of scepticism in doing so. The aim of this book is to present the data that are available in the hope that they will shed some light on why paediatricians treat their patients as they do and to highlight some of the areas where knowledge is lacking and which require further research.

References

- 1 Dickersin K, Straus SE, Bero LA. Increasing, not dictating, choice. *BMJ* 2007; **334**(Supplement Medical Milestones): s10.
- 2 The Cochrane Library. *Evidence for healthcare decision-making*. Available at: <http://www3.interscience.wiley.com/cgi-bin/mrwhome/106568753/WhatAreSystematicReviews.html>.
- 3 National Institute of Public Health. *First Annual Nordic Workshop on how to critically appraise and use evidence in decisions about healthcare*. Oslo, Norway, 1996.

Chapter 1

- 4 Muir-Gray JA. *Evidence-Based Healthcare: How to Make Health Policy and Management Decisions*. Churchill Livingstone, London, 1997.
- 5 Sackett DL, Rosenberg WM, Gray JA *et al*. Evidence based medicine: what it is and what it isn't. *BMJ* 1996; **312**: 71–2.
- 6 Singh S. *Fermat's Last Theorem*. Clays Ltd., St Ives, UK, 1997.
- 7 Scottish Intercollegiate Guideline Network. *SIGN 50: A Guideline Developers' Handbook Section 6: Forming Guideline Recommendations*. Available at: <http://www.sign.ac.uk/guidelines/fulltext/50/section6.html>
- 8 NICE. *Type 1 Diabetes in Children and Young People: Full Guideline 2004*. Available at: <http://www.nice.org.uk/page.aspx?o=CG015childfullguideline>
- 9 American Diabetes Association. Summary of revisions for the 2007 clinical practice recommendations. *Diabetes Care* 2007; **30**(suppl 1): S3.
- 10 Grimshaw J, Russell I. Achieving health gain through clinical guidelines. I: developing scientifically valid guidelines. *Qual Health Care* 1993; **2**: 243–8.
- 11 Burgers JS, Grol R, Klazinga NS *et al*. Towards evidence-based clinical practice: an international survey of 18 clinical guideline programs. *Int J Qual Health Care* 2003; **15**: 31–45.
- 12 Craig ME, Hattersley A, Donaghue K. ISPAD Clinical Practice Consensus Guidelines 2006–2007: definition, epidemiology and classification. *Pediatr Diabetes* 2006; **7**: 343–51.
- 13 Hattersley A, Bruining J, Shield J *et al*. ISPAD Clinical Practice Consensus Guidelines 2006–2007: the diagnosis and management of monogenic diabetes in children. *Pediatr Diabetes* 2006; **7**: 352–60.
- 14 Russell B. *Marriage and Morals*. George Allen & Unwin, London, 1929.
- 15 Dunger DB, Sperling MA, Acerini CL *et al*. European Society for Paediatric Endocrinology/Lawson Wilkins Pediatric Endocrine Society consensus statement on diabetic ketoacidosis in children and adolescents. *Pediatrics* 2004; **113**: e133–40.
- 16 Edge JA. *BSPED Recommended DKA Guidelines*. Available at: <http://www.bsped.org.uk/professional/guidelines/docs/BSPEDDKAApr04.pdf>
- 17 Wolfsdorf J, Craig ME, Daneman D *et al*. ISPAD Clinical Practice Consensus Guidelines 2006–2007: Diabetic Ketoacidosis. *Pediatr Diabetes* 2007; **8**: 28–43.
- 18 NHMRC. *The Australian Clinical Practice Guidelines on the Management of Type 1 Diabetes in Children and Adolescents 2005*. Available at: http://www.chw.edu.au/prof/services/endocrinology/apeg/apeg_handbook_final.pdf
- 19 Schade DS, Eaton RP. Dose response to insulin in man: differential effects on glucose and ketone body regulation. *J Clin Endocrinol Metab* 1977; **44**: 1038–53.
- 20 Luzi L, Barrett EJ, Groop LC *et al*. Metabolic effects of low-dose insulin therapy on glucose metabolism in diabetic ketoacidosis. *Diabetes* 1988; **37**: 1470–7.
- 21 Grimberg A, Cerri RW, Satin-Smith M *et al*. The 'two bag system' for variable intravenous dextrose and fluid administration: benefits in diabetic ketoacidosis management. *J Pediatr* 1999; **134**: 376–8.
- 22 HILO (Health Information for London Online). KA24: Knowledge Access 24 hours. KA24. Available at: <http://www.hilo.nhs.uk/>
- 23 Clar C, Waugh N, Thomas S. Routine hospital admission versus out-patient or home care in children at diagnosis of type 1 diabetes mellitus. *Cochrane Database Syst Rev* 2003; Issue 3: Art no CD004099.