Benefits of best practice guidelines:

Evaluating and applying the evidence

Peter Didsbury MBChB Dip Obs FRNZCGP

Peter Didsbury is a general practitioner in Counties Manukau. He co-chaired the development of national guidelines on Congestive Heart Failure and Asthma in Adults. As Manager of Integration ProCare Health Limited he has worked on several integration projects in both the Counties Manukau and Auckland districts. Peter is chairman of the New Zealand Guidelines Group. In this article he highlights key points and exposes some myths surrounding evidence-based practice and the use of best practice guidelines.

Permissible evidence

Evidence-based practice (EBP) has been defined as being the conscientious, explicit, and judicious use of current best evidence in making decisions about the care of individual patients.

It is often thought that EBP is based solely on evidence from randomised controlled trials (RCTs), with the view that only 'gold standard' evidence is acceptable. If there are no RCTs on that aspect of clinical practice, then the practitioner and patient are considered left in some kind of decision-making void. However, EBP actually involves systematically identifying and appraising the best evidence. It does not set a prerequisite benchmark about the standard of evidence to be used.

Evidence does in fact come in many forms. At a high level of differentiation there is both quantitative and qualitative evidence. Quantitative research focuses on the frequency or the occurrence of events. It counts events and is useful for guiding decisions around screening, diagnosis, therapy and prognosis. Qualitative evidence answers questions about meaning and how people feel about and experience situations.

Quantitative evidence

Various types of study or experimental design provide quantitative evidence. These can be ranked in a hierarchy according to how reliably the findings of a study design will predict outcomes when applied to clinical practice. This relates to the risk of bias within the study. Bias refers to systematic errors in the design or execution of a study which may lead to an overestimation or underestimation of the 'true' effect of an intervention. Different types of study have different inherent risk of bias and the hierarchy of study types differs for different clinical questions.

Quantitative research focuses on the frequency or occurrence of events, i.e. it counts events and is useful for guiding decisions around screening, diagnosis, therapy and prognosis.

Qualitative evidence answers questions about meaning and how people feel about and experience situations.

Well-designed double blind RCTs can provide the most reliable evidence on questions of therapy or harm, while cross-sectional studies suit questions of diagnosis, and cohort studies suit questions about prognosis. Table 1 summarises different study designs and the questions they best answer.

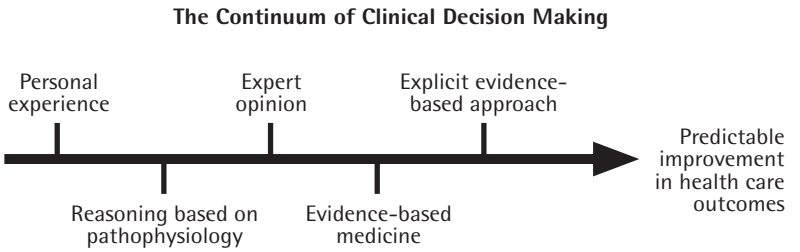
Qualitative evidence

Qualitative studies are designed to find out how people feel or experience certain situations. The main methods for collecting qualitative data are case studies, in-depth interviews, participant observation, and focus groups. The unit of analysis is generally a thought, concept or theme rather than an event. A qualitative study in South Wales,[1] for example, explored the views held by general practitioners, practice nurses, and patients about the role of guided self-management plans in asthma care. It found that neither health practitioners nor patients were enthusiastic about guided self-management plans

Table 1. Different clinical questions answered by different study designs

Clinical question	Most appropriate study design	Outcome measures
Diagnosis	Cross-sectional	Sensitivity, specificity Likelihood ratios
Prognosis	Cohort	Patient expected event rate
Harm	RCT, cohort or case control	Number needed to harm
Therapy	Systematic review, or RCT	Absolute risk reduction Number needed to treat

Figure 1. The continuum of clinical decision making



or certain about their usefulness. Quantitative studies, such as a Cochrane review, conversely, provided good evidence of the efficacy of guided self-management plans in improving health outcomes for patients with asthma.[2] Both sets of evidence are useful for the clinical care of patients with asthma.

Qualitative evidence then can, and indeed should, inform and assist practitioners in the 'art' of medicine. Ideally, the practitioner should combine the three dimensions of evidence, skill, and empathy in their clinical practice. Each is insufficient on their own and each is complementary. Quantitative evidence generally assists decisions of 'what to do' while qualitative evidence assists in understanding and relating to patients, informing the 'how to do it'.

Evaluating the evidence

The volume of evidence is mounting at an ever-increasing rate. Medical knowledge is doubling every 15 years with some 23 000 journals publishing two million new articles every year. If a general practitioner wished to keep up with the 10 leading medical journals it would require review of 200 papers and 70 editorials per month. Pity the poor generalist who would need to read 19 papers per day, 365 days of the year!

The role of guidelines

Obviously it is impossible for practitioners to keep abreast of all the published evidence. They need reliable summaries of evidence, such as those provided by systematically developed, evidence-based clinical practice guidelines. Guidelines are a critical counter to information overload.

Clinical practice guidelines are systematic statements to assist practitioners' and consumers' decisions about appropriate health care for specific clinical circumstances.[3] Figure 1 outlines the continuum of clinical decision-making. There is nothing new in health practitioners seeking information to help with decision-making in their clinical practice. Practitioners consult text books, seek and read articles, or consult colleagues. Many guidelines, such as consensus-based guidelines, and particularly older guidelines, were developed by agreement of a nominated group of experts. Evidence-based guidelines in contrast are developed only after systematic search, retrieval and appraisal of the evidence. Evidence-based guidelines:

- clearly differentiate opinion from evidence;
- document the strength of the evidence that supports each recommendation.

Evidence-based guidelines not only make statements about which of two options (treatments) is better, but quantify the difference in outcome, including benefit and harms, between them.

Because guidelines provide a comprehensive summary of the available evidence, the practitioner is freed from the onus of having to search, review and assimilate all such evidence, a task akin to trying to do one's tax return in one's head. The process of guideline development also systematically minimises bias, offering the most reliable evidence. This means that the guidelines more reliably predict the outcome of a decision (intervention) when applied to a clinical setting.

Clinical practice guidelines are systematic statements to assist practitioners' and consumers' decisions about appropriate health care for specific clinical circumstances

Guidelines assist decision making

A common criticism of guidelines is that they foster 'cook book' medicine and do not take patients' individual circumstances or preferences into account. However, good guidelines are not 'decisions' but rather 'decision

aids'. They articulate summary statements that facilitate decision-making. Imperative statements, such as 'children with acute otitis media should be prescribed antibiotics' are to be found in clinical protocols (as opposed to guidelines) and may have a place in clinical situations where there is little scope or benefit from variation (e.g. emergency resuscitation).

Guidelines are not only for practitioners but should be aids to interaction between the patient and their health care provider. Guidelines should provide information to permit joint decision-making between parties to pursue outcomes that they both agree are desirable.

Guideline development and implementation

The New Zealand Guidelines Group (NZGG) promotes and uses a systematic process in guideline development (see Figure 2).[4] The process starts with identifying which clinical issues are best addressed by the guideline, then follows a process of formulating questions to pose to the literature, systematically collating and assessing the evidence, quantifying the benefits and harms, and then developing recommendations and algorithms. Once published, guidelines are disseminated, implemented, and evaluated.

Guideline evidence grading

Guidelines explicitly describe the strength of evidence supporting their recommendations. Guideline teams critically appraise papers identified by the literature search to assess their strengths and weaknesses, such as bias, and to attribute a level of evidence (see Figure 3) for each study. In NZGG publications, readers can see the level of evidence for each study annotated in the text next to the reference. The process for evaluating the strength of a study varies according to the study type (RCT versus cross-sectional design etc).
Once all the studies are appraised, the guideline team then considers the studies as a whole. They consider the volume and consistency of the evidence, and its applicability to and potential clinical impact in the New Zealand setting. They draw up recommendations based on their findings and allocate a grade of evidence for each recommendation (see Figure 4).

Users of the guideline then know how reliable (in predicting the outcome) each recommendation will be when used in clinical practice. If a rec-

Figure 2. Process of guideline development

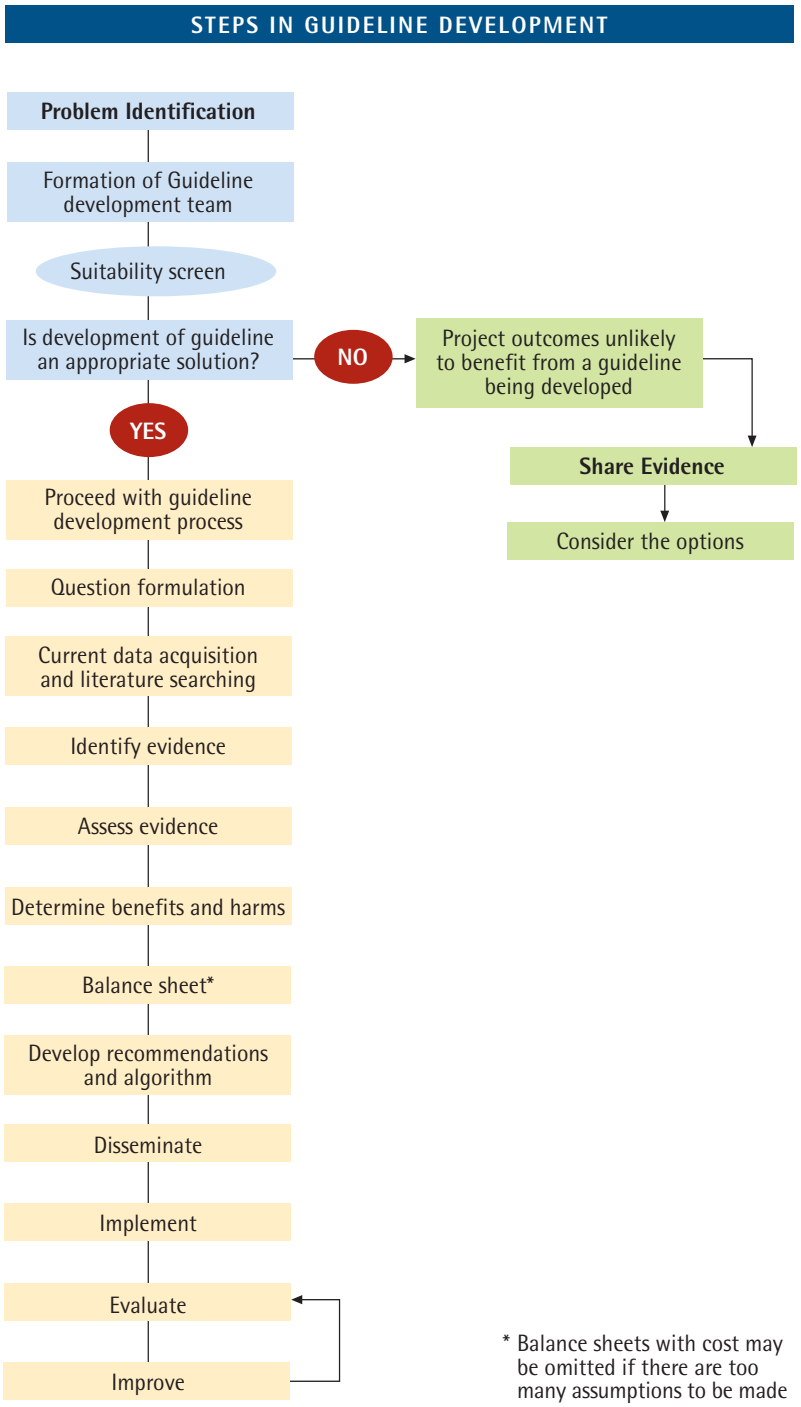


Figure 3. Levels of evidence

LEVELS OF EVIDENCE	
1++	High quality meta-analyses, systematic reviews of RCTs, or RCTs with a very low risk of bias.
1+	Well conducted meta-analyses, systematic reviews of RCTs, or RCTs with a low risk of bias.
1-	Well conducted meta-analyses, systematic reviews of RCTs, or RCTs with a high risk of bias.
2++	High quality systematic reviews of case-control or cohort studies with a very low risk of confounding, bias, or chance and a high probability that the relationship is causal.
2+	Well conducted case control or cohort studies with a low risk of confounding, bias, or chance and a moderate probability that the relationship is causal.
2-	Case control or cohort studies with a high risk of confounding or bias and a signifi cant risk that the relationship is not causal.
3	Non-analytic studies, eg, case reports, case series.
4	Expert opinion.

Figure 4. Grades of recommendation

GRADES OF RECOMMENDATION	
A	At least one meta-analysis, systematic review, or RCT rated as 1++, and directly applicable to the target population; **OR** A systematic review of RCTs or a body of evidence consisting principally of studies rated as 1+, directly applicable to the target population, and demonstrating overall consistency of results.
B	A body of evidence including studies rated as 2++, directly applicable to the target population, and demonstrating overall consistency of results; **OR** Extrapolated evidence from studies rated as 1++ or 1+.
C	A body of evidence including studies rated as 2 +, directly applicable to the target population and demonstrating overall consistency of results; **OR** Extrapolated evidence from studies rated as 2++.
D	Evidence level 3 or 4; **OR** Extrapolated evidence from studies rated as 2+.

GOOD PRACTICE POINT	
✓	Recommended best practice based on the clinical experience of the guideline development group

Figure 5. An example of recommendations appearing in the guideline: The Diagnosis and Treatment of Adult Asthma

INHALED CORTICOSTEROIDS (ICS)	
A	Treatment with inhaled corticosteroids is recommended in those who have daily symptoms of asthma or patients requiring SABAs daily.
B	Most adults with asthma should be initiated on treatment with low dose inhaled corticosteroids (beclomethasone dipropionate equivalent 400 µg/ day).
A	Fluticasone propionate is at least twice as potent as beclomethasone dipropionate.
A	Once-daily treatment with budesonide is as effective as twice-daily in mild asthma, once control has been achieved.
A	ICS have a relatively fl at dose response curve. Little additional benefi t is gained from doses above 500 µg/ day of fluticasone propionate or 800 µg/ day of beclomethasone dipropionate/budesonide.
A	There is evidence of an increased risk of cataracts, reduced bone mineral density, glaucoma and bruising of the skin with long-term treatment with high dose ICS (eg, more than 800 µg/ day beclomethasone dipropionate).
A	Early treatment with ICS in people with persistent symptoms and impaired lung function leads to better lung function in the medium term, and may help prevent the development of irreversible airflow limitation.
✓	High doses of ICS should be avoided where possible for adults with asthma who have pre-existing conditions or vulnerability to conditions such as osteoporosis or cataracts.
✓	The guideline team recommends that LABAs should always be considered in individuals who continue to experience symptoms despite taking moderate (800 µg BDP/day) doses of ICS as this is at the top end of the dose response curve of ICS and higher doses are associated with increased risk of adverse effects (opinion).

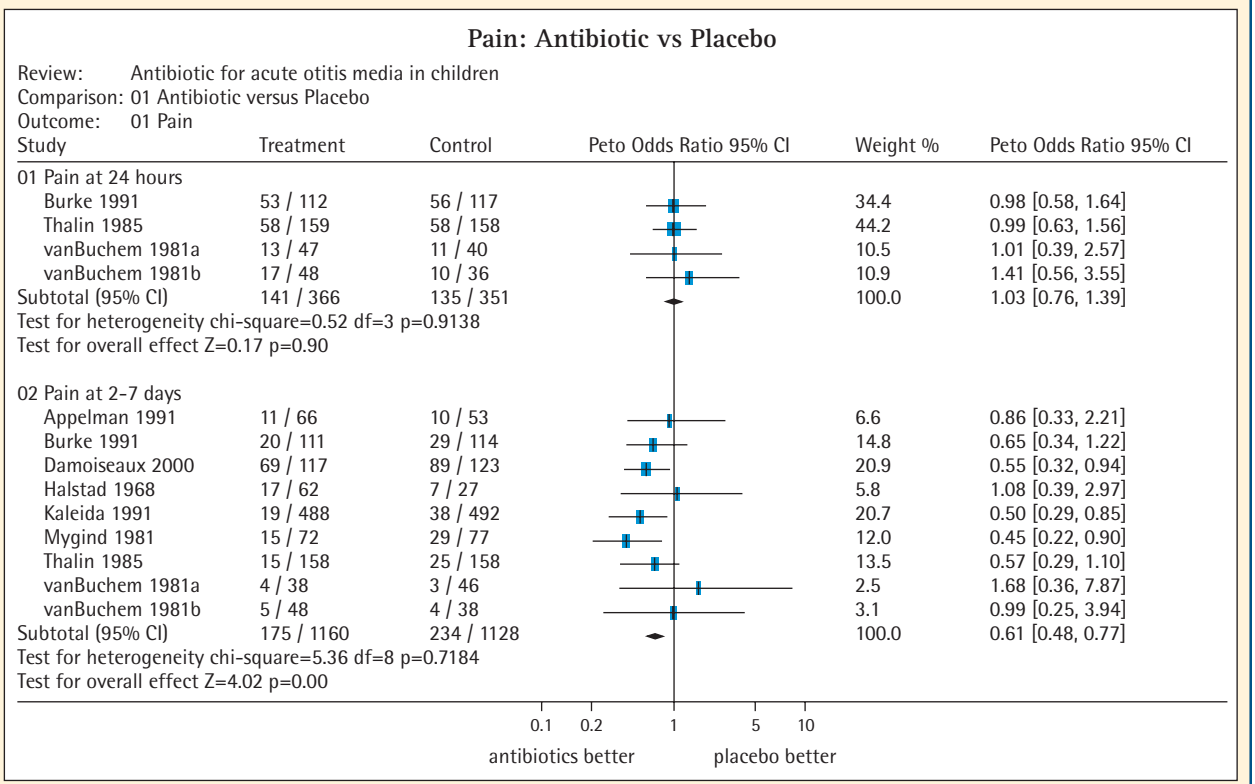
Table 2. Measures used to describe size of effect

Acronym	Measure	Definition	Formula
CER	Control Event Rate	Rate of events in the control group	
EER	Experimental Event Rate	Rate of events in the experimental group	
RR	Relative Risk	The ratio of the risk of an event or outcome occurring in a population exposed to an effect compared to the population not exposed	EER/CER
ARR	Absolute Risk Reduction	The arithmetic difference in event rates between the experimental group and control group	CER-EER
RRR	Relative Risk Reduction	The proportional reduction in rates of events between the experimental and control groups	RRR=CER-EER/CER
NNT	Number Needed to Treat	The number of patients that have to be treated to achieve an outcome	NNT=1/ARR

Otitis Media and Measures of Effect

The Cochrane review *'Antibiotics for acute otitis media in children'*[5] contains a meta analysis of studies conducted on the effect of antibiotics on reducing pain. Here is the forest plot for the meta-analysis.

Figure 6. Pain: Antibiotic v Placebo – Forrest Plot



2288 children were randomised to antibiotic therapy (1160) or placebo (1128).
In the experimental group, 175 children had pain after 2-5 days.
EER=175/1160=15%.
In the control group, 234 children had pain.
CER=234/1128=21%.

The Relative Risk Reduction is CER-EER/CER =21-15/21 or 29%.

The Absolute Risk Reduction is CER-EER=21-15=6%. In other words, 6% of children gained benefit in terms of relief of pain from having antibiotics.

The NNT is 1/ARR or 17. In other words, 17 children need to be treated with antibiotics to avoid one painful ear after 2–5 days.

CER	21%
EER	15%
RRR	21%
ARR	6%
NNT	17

The example demonstrates that the various measures of effect can have quite varying values.

Interestingly, the NNH (Numbers Needed to Harm) for antibiotics is also 17, i.e. one child in 17 treated with antibiotics experienced diarrhoea or vomiting.

A useful guideline then, rather than making a recommendation such as 'children with otitis media should be prescribed antibiotics', presents this summary of evidence. It is then available to assist the parent and their general practitioner in their decision whether to use antibiotics or not. Presumably the decision will be based on the relative value placed on the potential harms of diarrhoea and vomiting versus the potential benefits of pain relief.

ommendation is based on A grade evidence, there is a good likelihood that applying the recommendation will result in the same outcomes as seen in the clinical trials. If based on D grade evidence, there is greater uncertainty. Figure 5 provides an example of recommendations from the guideline: *The diagnosis and treatment of adult asthma*.[6]

Size of effect

Users of guidelines also want to know what the size of the effect is for the recommended intervention. Strength of evidence and size of effect are totally independent. There can be A grade evidence that an intervention is not effective.

ARR, RRR, NNTs et al.

There are various ways of expressing size of effect. NZGG guidelines tend to prefer the measures *Numbers Needed to Treat* or *Numbers Needed to Harm* (NNT/NNH) as these give the best indication of effect. Table 2 summarises the measures used to describe size of effect. The use of different measures is applied in the example in the Box.

The New Zealand Guidelines Group

NZGG provides regular courses on the development and use of guidelines. The NZGG website www.nzgg.org.nz contains a library of New Zealand guidelines and links to other trusted evidence-based sites. It also has resources for those who wish to develop or appraise guidelines.

References

1. Qualitative study on the views of health professionals and patients on guided self management plans for asthma. BMJ 2000; 321:1507–150.
2. Gibson P et al. Limited patient education programs for adults with asthma. The Cochrane library; 4, 2001.
3. Institute of Medicine. Clinical Practice Guidelines. Directions for a new program. Washington DC National Academy Press; 1990.
4. Handbook for the preparation of explicit evidence based clinical practice guideline. www.nzgg.org.nz
5. Glasziou PP, Del Mar CB, Sanders SL, Hayem M. Antibiotics for acute otitis media in children (Cochrane Review). In: The Cochrane Library, Issue 1, 2003. Oxford: Update Software.
6. New Zealand Guidelines Group. The diagnosis and treatment of adult asthma. New Zealand Guidelines Group; Sept 2002.