

# Recommended methods to derive clinical inputs for proposals to PHARMAC

**Version 1B**

**20 July 2005**

## Table of contents

Recommended methods to derive clinical inputs for proposals to PHARMAC .....	1
Table of contents.....	2
Context of these recommendations .....	3
Status of these recommendations .....	3
Why produce these recommendations? .....	4
PHARMAC's own processes.....	4
So opportunistic use of data on hand is OK, but state how evidence is compiled.....	4
In summary:.....	5
Source and precedence elsewhere .....	5
Aim of treatment.....	7
Choice of comparator.....	8
Epidemiology, Public Health Significance and Estimated Utilisation .....	9
Defining and quantifying disease-severity groups .....	11
Need by Maori and Pacific peoples.....	12
Search strategy.....	14
Outcome measures .....	17
<i>Effectiveness</i> .....	17
<i>Withdrawals and continuations</i> .....	17
Criteria for including or excluding source data, and data to include .....	19
Quality of the evidence .....	20
1. Critical appraisal and Scoring for quality.....	20
2. Hierarchy of evidence.....	22
3. Relevance of the evidence.....	23
4. Applicability of the evidence .....	24
Data synthesis.....	26
Data abstraction .....	26
Effectiveness.....	26
Withdrawals and continuations in randomised controlled trials.....	26
Assessment of missing data and possible publication bias .....	27
Additional clinical opinion and clinical contacts .....	28
Overall summary of benefits and harms; balance sheets .....	30
Review process .....	31
Appendix: Calculating person-year weighted incidence rates, weighted rate ratios and relative risk reductions .....	32
References/endnotes.....	34

## Context of these recommendations

The following recommendations are made to help suppliers and others most effectively propose to PHARMAC new products (mainly pharmaceuticals) to be funded through the New Zealand Pharmaceutical Schedule.

The recommendations cover both clinical data submitted to PHARMAC's Pharmaceutical and Therapeutics Advisory Committee (PTAC) and clinical data submitted to PHARMAC as part of economic analyses.

Clinical input data examine the effectiveness of interventions and the impact the underlying disease/disability state being treated has on all-cause mortality and overall quality of life (QoL).

Underlying disease/disability states can be described by the incidence of specific events and prevalence of specific health states. Generic quality of life (QoL) scores in turn describe these health states. Data on both the effectiveness of a treatment and the impact on mortality and quality of life are necessary to calculate the benefits of that treatment, measured in terms of gains in quality-adjusted life years (QALYs).<sup>1</sup>

Rigorous clinical input data and assumptions are necessary both for:

- how PTAC and PHARMAC assess the effectiveness of pharmaceutical agents; and
- economic analyses either submitted to and/or developed by PHARMAC.

These inputs include uptake and continuation rates, used in expenditure forecasts/cost modelling.

Clinical (including epidemiological) inputs are critical to PHARMAC's decision making, in order to assess health need, effectiveness, total costs and cost-effectiveness. These aspects cover five of PHARMAC's eight decision criteria (see [PHARMAC's Operating Policies and Procedures](#) 2.2 decision criteria (a) health needs of all eligible people within New Zealand; (b) particular health needs of Maori and Pacific peoples; (c) clinical benefits and risks; (d) cost-effectiveness; and (e) budgetary impact).

Details of the broad principles used by PHARMAC for pharmacoeconomic evaluations are on-line (see [A Prescription for Pharmacoeconomic Analysis](#) Version 2, May 2007). These principles include the use of costs to the Pharmaceutical schedule, other health sector costs and direct patient costs when measuring effects on costs overall; measuring QALY gains; discounting both costs and QALY gains according to PHARMAC's current discount rate [8% at July 2005]; and uni/multivariate sensitivity analyses. Include these aspects in any economic analyses you submit to PHARMAC, along with the detail of clinical assumptions as described below.

## Status of these recommendations

The following recommendations are discretionary. They aim to expedite supplier's applications, not to be an unnecessary hurdle.

Both PTAC and PHARMAC consider all relevant evidence supplied to them.

However, PTAC and PHARMAC find it easier to expedite and consider proposals when, within reason, applicants comprehensively collate and systematically analyse good quality data, and then present these simply, clearly and in prescribed formats. The intention of these recommendations is to help suppliers and others to do this.

---

<sup>1</sup> where:

- life expectancy and QoL combine to give quality-adjusted life expectancy (QALE);
- excess mortality and loss in QoL combine to give loss in quality-adjusted years of life (QALYL = usual QALE minus QALE for that disease/disability state); and
- combining effectiveness with mortality and QoL measures gives gains in quality-adjusted life years (QALY gains) from applying an intervention to a target population.

## Why produce these recommendations?

Evidence-based policy-making demands that there is at least recognition of where input data and assumption come from, how they are identified, and the potential for bias or uncertainty. PHARMAC is committed to this approach.

Yet periodically PHARMAC receives applications that are difficult to sift through, giving varying levels of evidence and not stating how well that evidence was compiled.

Although probably unintentional, such disorganisation, padding and failing to include all relevant information creates difficulties and inefficiencies for PTAC and PHARMAC staff. Conversely, applications with strong comprehensively-derived outcomes and cost-effectiveness evidence seldom need padding – their data speak for themselves.

Due diligence requires PTAC and PHARMAC to fully read applications in order to adequately consider all relevant evidence. Poor clinical data not identified as such can cloud the real issues, and deflects from being able to concentrate on the core quality evidence. It also risks subliminal bias, through persistent repetition of less robust assertions. These factors again create difficulties and inefficiencies, resulting in less time for core decision-making.

For these reasons, PTAC and PHARMAC require applicants to state where and how they derived need and cost-effectiveness data for their pharmaceutical agents, to try to derive systematically, and present clearly.

## PHARMAC's own processes

In context, PHARMAC recognises and itself undertakes four levels of economic analysis: very rapid, preliminary, indicative, and detailed.<sup>1</sup>

Type	Description
Detailed	Detailed and systematic identification and synthesis of effectiveness, natural history, QoL and costings data. Follows Prescription for Pharmacoeconomics policies. Follows policies of Recommended Methods to Derive Clinical Inputs for Proposals to PHARMAC. Reviewed internally (PTAC for clinical assumptions, PHARMAC) and externally. 3-6 months FTE input.
Indicative	Interim assessment using opportunistic data but more detailed than preliminary CUA. Follows Prescription for Pharmacoeconomics policies. Largely follows policies of Recommended Methods to Derive Clinical Inputs for Proposals to PHARMAC. Typically reviewed internally (PTAC for clinical assumptions, PHARMAC). 4-6 weeks FTE input. Includes remodelling of supplier analyses.
Preliminary	Rapid assessment using opportunistic data. 1-2 weeks FTE input.
Very rapid	Very rapid (first cut) assessment using opportunistic data, 1-2 days FTE input. Includes supplier analyses not yet evaluated by PHARMAC staff.

PHARMAC recognises that it works in a pragmatic public policy/purchasing environment, where analytical capacity is not limitless. So there will be inevitable trade-offs between precision and timeliness. The level (extent and depth) of analysis will vary according to: individual policy issues; the availability of analyst resources at the time; the defensibility of recommendations derived from the results; and the extent of information available for analysis.

## So opportunistic use of data on hand is acceptable, but state how evidence is compiled

PHARMAC's approach above means that comprehensive searches and analyses are not necessarily undertaken every time. However, it does mean that with each analysis that PHARMAC does at least describe the level of analysis and search strategy etc (if any). This is done using standard reporting templates.

In turn, PHARMAC will not necessarily take information submitted by applicants at face value - unless applicants too demonstrate rigorous processes. Applicants may source data opportunistically, but do need to state what they did. When presenting data, state the sources.

Opportunistic use of data on hand, without formal search strategies may be acceptable - provided it is acknowledged as such. Applicants can collate data systematically or opportunistically - but should at least state what they did and why.

#### In summary:

When presenting data to PHARMAC:

- focus on the most important elements;
- be succinct; and
- explain all sources and assumptions.

#### Source and precedence elsewhere

The following recommendations derive from standard guidelines internationally. These include those of:

- Australia's Pharmaceutical Benefits Scheme (PBS)<sup>2 3 4</sup> (<http://www.health.gov.au/pbs/pubs/pharmpac/gusubpac.htm>, <http://www.health.gov.au/pbs/pubs/guidelines/content.htm> and <http://www.health.gov.au/pbs/pubs/pharmpac/interim/index.htm>), as largely adapted by the Health Funding Authority's draft Application Guidelines (1999);
- the UK's NHS Centre for Reviews and Development (CRD)<sup>5</sup> (<http://www.york.ac.uk/inst/crd/report4.htm>) and NICE<sup>6</sup> ([http://www.nice.org.uk/pdf/TAP\\_Methods.pdf](http://www.nice.org.uk/pdf/TAP_Methods.pdf));
- the Canadian Coordinating Office for Health Technology Assessment (CCOHTA)<sup>7</sup> <http://www.ccohta.ca/>;
- the US FDA<sup>8</sup> (<http://www.fda.gov/cder/mapp.htm>)

Refer to these and other resources for more details, for instance:

- the Cochrane Handbook (<http://www.cochrane.org/cochrane/hbook.htm>); and
- the Etext on Health Technology Assessment (HTA) Information (<http://www.nlm.nih.gov/nichsr/ehta/>, with companion text TA:101 Introduction to Health Technology Assessment [http://www.nlm.nih.gov/nichsr/ta101/ta101\\_c1.htm](http://www.nlm.nih.gov/nichsr/ta101/ta101_c1.htm)).

PHARMAC acknowledges the extensive work undertaken by the PBS, CRD, CCOHTA and other organisations. This work both sets precedence for the following PHARMAC guidelines for the New Zealand setting, and aids their content and style.

## **Description of disease, and Indication for the new pharmaceutical treatment**

When proposing to PHARMAC for new pharmaceuticals to be listed on the New Zealand Pharmaceutical Schedule, applicants should:

- ⇒ Describe the precise clinical indication(s) for the proposed treatment.
- ⇒ For the disease/condition, provide a summary including:
  - A brief description of the disease/condition;
  - The disease/condition's natural history;
- ⇒ Include not only the disease or condition to be treated, but also other details such as:
  - the stage of the disease;
  - co-morbid conditions;
  - (where appropriate) whether therapy is first- or second line; and
  - whether it is to be used alone or with other therapies.
- ⇒ Cite the literature/data sources used for the disease summary.

## **Aim of treatment**

Applicants should:

⇒ Categorise the aims of treatment in terms of:

- primary prevention;
- secondary prevention;
- cure;
- preventing deterioration;
- symptom relief/maintenance;
- rehabilitation; and/or
- palliation

⇒ Also describe the desired goals/outcomes of treatment in terms of:

- preventing premature death;
- preventing long-term sequelae from non-fatal disease events; and/or
- improving immediate disease/disability symptoms.

## Choice of comparator

Applicants should:

- ⇒ Describe what the current available alternative treatment(s) are for this condition.
- ⇒ List all appropriate comparators for each approved indication of the drug. State whether comparators are pharmacologic comparators (similar mechanism of action) or clinical comparators (similar clinical outcome).
- ⇒ If available, include dose equivalencies and justify them. State whether dose equivalencies were derived from direct or indirect comparisons.

*Notes:*

*The most realistic alternative might be a different regime, different drug/treatment, different modality, placebo, next-best alternative, or no treatment.*

*Different modalities might include non-pharmaceutical medical treatment, pharmaceutical treatment, surgery, primary health care, mental health services, preventive programmes, etc.*

## Epidemiology, Public Health Significance and Estimated Utilisation

Applicants should:

- ⇒ For each recommended indication, estimate for the first five years of listing (shown year-by-year) the number of people in New Zealand:
- suffering from the particular disease/condition(s);
  - likely to seek treatment for the disease/condition(s);
  - likely to be prescribed the pharmaceutical;
  - likely to be treated for the disease/condition, broken down by:
    - those that can be successfully treated by the pharmaceutical only;
    - those that can be treated by both the pharmaceutical and other pharmaceuticals that treat the same condition;
    - those that can be treated by only other pharmaceuticals; and
    - those that can be treated, completely or partially, by other therapies.
- ⇒ For the disease/condition, describe the burden of the disease/condition (impact of the disease/condition on the individual and the community), including:
- the incidence and prevalence of disease/condition in New Zealand;
  - estimates of the morbidity (e.g. annual no. of hospitalisations) and premature mortality of the condition in New Zealand (e.g. annual no. of deaths; no. of potential years of life lost before age 80 (PYLL80)<sup>ii</sup>);
  - where available, estimates of average disability-adjusted life years (DALYs) lost by an individual patient because of the disease(s)<sup>iii</sup>;
  - where available, population loss of disability-adjusted life years (DALY loss) for the disease(s), and the disease's % of population total DALY loss across all diseases (543,000 DALYL in 1996);<sup>9</sup>
  - where available, estimates of the extent (%) to which the disease(s) are modifiable by the proposed treatment.
- ⇒ Describe estimated uptake using the above data, i.e.
- $$\frac{\text{no. of people likely to be prescribed the pharmaceutical}}{\text{no. of people with the particular disease/condition(s)}}$$
- ⇒ Also estimate the change in the extent of use of other agents.
- ⇒ Cite the data sources used, and clearly and explicitly state, the bases of the assumptions made for the estimated numbers. Estimates and assumptions can be indicated in ranges.

---

<sup>ii</sup> PYLL80 = sum across all individuals of (80 years minus age at death for an individual)

<sup>iii</sup> where an individual's normal life expectancy minus their loss of disability-adjusted life years (DALY loss) from their disease/condition equals their disability-adjusted life expectancy (DALE)

Notes:

*Use an epidemiological approach to estimate the likely number of patients projected to be eligible under access scenarios. Base this on the incidence and/or prevalence of the condition to be treated, current patterns of and time trends in utilisation for the indication, and patterns seen in other markets with the introduction of the new agent(s).*

*Useful sources of mortality and morbidity data include cancer registry data, hospitalisation episode data<sup>10</sup>, mental health episode data<sup>11</sup>, and mortality data<sup>12</sup>, available from NZ Health Information Service (<http://www.nzhis.govt.nz>). Details on the range of data available can be obtained from the Ministry of Health's Public Health Metadata Database (see <http://www.moh.govt.nz/PHMetadata.nsf>).*

*Other prevalence and outcomes data may be located using relevant Medline searches (<http://www.ncbi.nlm.nih.gov/entrez/query.fcgi>) using search terms [the disease/indication], (epidemiology OR prevalence OR risk factors OR natural history OR prognosis OR outcomes OR survival), and specific outcomes relevant to the disease/indication. Local epidemiology/outcomes data can be located more specifically by including the terms (Australia OR Zealand) in the search.*

## Defining and quantifying disease-severity groups

Applicants should:

- ⇒ Define the range of disease symptoms and severity.
- ⇒ Then define and quantify a set of disease-severity subgroups within the proposed treatment population, where total benefit is likely to vary.  
  
Define groups according to combinations of [disease groups or conditions] by [symptom complexes] by [severities of disease] by [patient characteristics (age by sex by ± other factors)] by [co-morbidities] by [intractability].
- ⇒ Compile and then apply epidemiological data to estimate the distribution of disease states across the subgroups (i.e. the prevalence of each subgroup).

*Notes:*

*Benefit will vary between groups by age, gender and other patient-related factors; disease-related factors, and factors specifically affecting the effectiveness of treatment.*

*The degree of breakdown depends upon the complexity of the rationing and targeting decisions to be made. Some situations will require many subgroups, others just the overall group. Groups will be a manageable set of groups comprising combinations of:*

- *disease groups or conditions, with*
- *symptom complexes, with*
- *severities of disease, with*
- *patient characteristics (age x sex x ± other factors), with*
- *co-morbidities, with*
- *intractability.*

*Epidemiological data relevant to estimating the distribution of disease states across subgroups (ie prevalence of each subgroup) can be compiled from a variety of sources, including:*

- *data from formal epidemiological surveys and registries for NZ;*
- *survey and registry data from elsewhere, where there are no NZ data;*
- *observational cohort studies;*
- *control groups of controlled clinical trials (RCTs);*
- *national mortality and utilisation statistics; and*
- *at times where data are otherwise still insufficient, then commissioning surveys locally, including sampling of primary care settings.*

## Need by Maori and Pacific peoples

Applicants should:

⇒ Describe the following ethnic-specific data:

- The availability and quality of data able to indicate the extent of disparity in disease prevalence, incidence, notifications of new cases, public hospital hospitalisations, deaths, etc between Maori, Pacific peoples and non Maori/non Pacific people.
- Where data are available, the degree of impact on Maori and Pacific people relative to non Maori/non Pacific, using age-adjusted relative risks for notification / hospitalisation / mortality rates for disease for Maori or Pacific people versus non Maori/non Pacific (as reference group). In each instance describe:
  - the number of cases occurring in Maori with rate per 1000 age-standardised (further details below);
  - the number of cases occurring in Pacific people with rate per 1000 age-standardised;
  - the number of cases in non Maori/non Pacific with rate per 1000 age-standardised;
  - both risk differences for age-standardised rates and age-adjusted rate ratios with 95% confidence limits for [Maori vs non Maori/non Pacific]; and
  - both risk differences for age-standardised rates and age-adjusted rate ratios with 95% confidence limits for [Pacific peoples vs non Maori/non Pacific].
- Where data are available, also describe key age-specific differences between Maori or Pacific people and non Maori/non Pacific, i.e. both:
  - which particular age-sex groups for Maori have the greatest age-specific risk differences when compared with non Maori/non Pacific; and
  - which particular age-sex groups for Pacific peoples have the greatest age-specific risk differences when compared with non Maori/non Pacific.
- Comment on the extent to which notification / hospitalisation / mortality rates correlate with need for pharmaceuticals.

⇒ Use standard information sources (e.g. NZHIS mortality, hospitalisation and mental health data), supplemented where required by discussion with relevant experts and/or Medline searches<sup>13</sup> on [[disease]] \* [Maori or Pacific] \* [Zealand] etc. Sources include prevalence data, incidence/notifications data, hospitalisation data<sup>14</sup>, mental health episode data<sup>15</sup>, and mortality data<sup>16</sup> for relevant indicator(s), combined with New Zealand age-ethnic-specific populations<sup>17</sup>.

⇒ Where able, calculate age-standardised rates using the direct method against Segi's Standard World population. Calculate age-standardised rate ratios (relative risks) and 95% confidence limits using standard binomial techniques.

*Notes:*

*Effects on Maori and Pacific people's health is important to PHARMAC, and is considered carefully. Criterion (b) of [PHARMAC's Operating Policies and Procedures](#) states the particular health needs of Maori and Pacific peoples to be one of PHARMAC's decision criteria.*

*Note that use of incidence (notification) rates to proxy need for pharmaceuticals is valid only if survival times are similar between ethnic groups.*

*For hospitalisations and deaths since 1996, Maori population denominators comprise any Maori inclusive. For mental health data, Maori denominators comprise exclusive sole-defined. Use of sole-defined and similarly restricted denominators attempts to mitigate for under-counting of Maori by mental health statistics.*

*Age-standardised rate ratios summarise disparity for Maori over all ages when compared with non-Maori, mitigating bias due to differences between ethnic groups' age structures.*

*Hospitalisation rates do not necessarily correlate well with prevalence (here, need for pharmaceuticals), although they might proxy total burden of disease.*

*Historically there have been major problems with the accuracy and validity of the data routinely available within the New Zealand health sector. In general there are few timely data available about the types of patients, disease/disability rates and pharmaceutical indications conditions treated in primary care, particularly when subdividing ethnicity by disease/indications. Mortality data are less timely than hospitalisations data and give less precise estimates of risk (given smaller numbers of deaths). Maori death rates have in past decades been substantially under-represented due to ethnicity miscoding. Causes of death described on death certificates can be inaccurate. Routinely collected hospitalisations data are less accurate than mortality data, with historically over one-quarter of discharge diagnoses being incorrectly coded. Hospitals had systematically under-counted Maori admissions, hence numerator-denominator mismatches historically. Hospital admissions only indirectly measure need, being also affected by supply factors (regional etc variations in admitting practices relating to bed/service availability and clinical protocols), in turn affecting ethnic rates. Double counting of readmissions and of inter-hospital transfers as "new" admissions further biases the data.*

*Note that describing Maori or Pacific peoples' representation (i.e. what proportion of patients in the community with [disease] are Maori or Pacific peoples) has only partial relevance to PHARMAC's decision-making. Disease prevalence does not necessarily correlate with need for pharmaceuticals. In addition, deviations from expected-need for ethnic groups have little relevance if these ethnic groups are not accessing appropriate treatment. This is where accessing appropriate treatment comprises having a GP or similar primary care professional available and feeling able and inclined to see them and the disease is identified and treatment prescribed - all at probabilities equal to non-Maori. There is evidence with many aspects of healthcare that Maori access is lower than for non-Maori.*

## Search strategy

Applicants should:

- ⇒ Define the research question and the topics used by their literature search(s) to identify all relevant evidence.
- ⇒ Identify relevant clinical data and economic analyses using key features of Sheffield University's ScHARR Internet search protocol<sup>18</sup> at <http://www.shef.ac.uk/uni/academic/R-Z/scharr/ir/proto.html> and NZ Health Technology Assessment's COSI search protocol at <http://nzhta.chmeds.ac.nz/nzhtainfo/protocol.htm>.

At a minimum, the search needs to be for randomised controlled trials, review articles, meta-analyses, guidelines, economic analyses, and reports by regulatory/funding authorities, identifiable in:

- the Cochrane Library (abstracts available at <http://www.update-software.com/ccweb/cochrane/revabstr/mainindex.htm>);
- the DARE, NHS EED, and HTA databases compiled by the NHS Centre for Reviews and Dissemination at <http://nhscrd.york.ac.uk/>;
- Medline (US National Library of Congress's PubMed on-line at <http://www.ncbi.nlm.nih.gov/entrez/query.fcgi>, which includes the Clinical Queries Using Research Methodology Filters option at <http://www4.ncbi.nlm.nih.gov/PubMed/clinical.html> (from Hayes *et al* 1994<sup>19</sup>) for prognostic data); and
- the TRIP database at <http://www.tripdatabase.com>.<sup>20</sup>

Also use:

- BMJ Clinical Evidence Online at <http://www.clinicalevidenceonline.org/>;
- electronic searches of website indexed versions of the British Medical Journal, The Lancet, New England Journal of Medicine and JAMA, using key words; and
- reference lists of sourced articles to obtain further papers not identified in the above searches;

- ⇒ Describe the search strategies used to retrieve the relevant clinical studies in the published and unpublished literature. At a minimum, state:
  - the medium used to conduct the search and by whom – if at all (CD-ROM; on-line internet; in-house searches; search by external agency; search by an external library);
  - the specific databases searched – if at all;
  - the date the search was conducted – if at all;
  - the time period searched – if at all;
  - the complete search strategies used and the key words/MeSH heading and relationships (sets and Boolean) used – if at all; and
  - any supplementary search strategies.

- ⇒ Include in search strategies, singly or in combination, the search terms:
  - [the disease/indication] AND (epidemiology OR prevalence OR risk factors);
  - [the disease/indication] AND (natural history OR prognosis OR outcomes OR survival OR [specific outcomes]);

- [the disease/indication] AND (meta-analysis OR systematic overview OR review) ± (randomised controlled trial or whatever terms are used to elicit RCTs);
- [the disease/indication] AND (cost-benefit analysis OR cost-utility analysis OR cost-effectiveness analysis OR economic analysis);
- [the disease/indication] AND quality of life; and
- [the disease/indication] AND (Australia OR New Zealand).

⇒ Provide a citation list of all identified studies.

⇒ State which studies were excluded and why.

*Notes:*

*Any assessment needs to consider available evidence on effectiveness and cost-effectiveness. PTAC and PHARMAC need to be confident that there are no outstanding empirical data likely to alter any recommendations.*

*Given the tradeoffs between comprehensiveness and timeliness, highly extensive searches similar to those carried out for Cochrane reviews are not necessarily expected. However, applicants should make reasonable attempts to identify and retrieve randomised controlled trials that are in English, any relevant safety data and any published economic analyses.*

*The key factors that can influence the outcome of a literature search are:*

- *the date the search is carried out;*
- *the time period that is covered by the search;*
- *the database(s) used, and how it is provided; and*
- *the search terms and search strategy that is used.*

*The search should be constructed to find clinical trials, reports of safety issues and any published cost effectiveness analyses of the new technology. Further information on how to conduct such a literature search can be found on-line in the Cochrane Handbook (section 5. Locating and selecting studies, and Appendix 5c. Optimal search strategy for RCTs) at <http://www.cochrane.org/cochrane/hbook.htm>, the Etext on Health Technology Assessment (HTA) Information at <http://www.nlm.nih.gov/nichsr/ehta/>, and NZ Health Technology Assessment's literature advisory service (see <http://nzhta.chmeds.ac.nz>).*

*Given health technology assessment at PHARMAC eventually involves economic appraisal relevant to the Australasian setting, beyond clinical appraisal, then key words for searching should include:*

- *the disease/indication;*
- *epidemiology, prevalence, natural history, prognosis;*
- *guidelines;*
- *meta-analysis, [systematic] overview, review;*
- *randomised controlled trial, outcomes, [specific outcomes];*
- *cost-benefit analysis, cost-utility analysis, cost-effectiveness analysis, economic analysis; and*

- *Australia or New Zealand.*

*Other sources of evidence apart from formal medical literature include:*

- *reviews by regulatory authorities (e.g. a report from a health technology agency);*
- *consensus reports from expert panels; and*
- *reviews by expert bodies such as specialist colleges/professional bodies.*

*While these may not be primary sources of data, they may help identify relevant trials or studies.*

*Experts in New Zealand able to provide advice and training in searching the relevant literature include (but are not necessarily limited to):*

- *Auckland University's EPIQ (Effective Practice, Informatics & Quality Improvement) <http://www.health.auckland.ac.nz/population-health/epidemiology-biostats/epiq/>, and*
- *the New Zealand Health Technology Assessment Clearing House for Health Outcomes and Health Technology Assessment (NZHTA) (<http://nzhta.chmeds.ac.nz/>).*

## Outcome measures

Applicants should:

- ⇒ List and describe the key health outcomes measured by the empirical data and analysed.
- ⇒ Explain any methods of measurement of these outcomes that are likely to be unfamiliar.
- ⇒ Emphasise clinically relevant and valid outcomes of highest importance for the health of patients with the disease state.
- ⇒ If relevant (particularly when clinically relevant/valid outcome data are not available), use intermediate or surrogate outcomes. When using intermediate/surrogate outcomes, describe the strength of evidence showing that extrapolating the specific intermediate outcomes to clinically relevant patient outcomes is valid.
- ⇒ Also include withdrawals due to adverse effects (WDAE), total adverse events (AE) and total serious adverse events (SAE) as outcomes.

*Outcomes may include (but are not restricted to):*

### **Effectiveness**

- *period-related dependent mortality rates;*
- *survival;*
- *life expectancy;*
- *health service utilisation, e.g. hospitalisation rates, mean length of stay;*
- *disability levels;*
- *quality of life;*
- *adverse event(AE) rates;*
- *serious adverse event (SAE) rates;*
- *disability-adjusted or quality adjusted life years; and*
- *disability-adjusted or quality adjusted life expectancy.*

### **Withdrawals and continuations**

- *percentage of patients withdrawing because of adverse event (WDAE);*
- *percentage of patients withdrawing for any reason (including all deaths);*
- *percentage withdrawals in relevant survivors (where relevant patients are those remaining after censoring for non-epilepsy- or non-treatment-related withdrawals); and*
- *percentage continuations in relevant survivors.*

*Notes:*

*Important clinically relevant and valid outcomes can encompass both benefit and harm, depending on whether the incidence is reduced or increased.*

*Examples of important clinically relevant and valid outcomes for cardiovascular disease include all-cause mortality, cardiovascular-related mortality, all-cause morbidity, non-fatal MI, etc.*

*Examples of intermediate or surrogate outcomes with less clinical relevance or less clear validation of clinical relevance for patients for cardiovascular disease include blood pressure or cholesterol levels.*

*Further material on clinically relevant/valid outcomes (patient-focussed outcomes, clinical outcomes) and surrogate/intermediate outcomes can be found later in the section 'Quality of Evidence: 3. Relevance of the evidence'.*

*Adverse events can be defined as follows:*

- *Adverse event (AE): Any untoward medical occurrence that may present itself during treatment or administration with a pharmaceutical product, and which may or may not have a causal relationship with the treatment. (see <http://www.fda.gov/cder/guidance/iche3.pdf>)*
- *Serious adverse event (SAE): An adverse event which results in death, is life threatening, requires inpatient hospitalisation or prolongation of existing hospitalization, creates persistent or significant disability/incapacity, or a congenital anomaly/birth defect. (see <http://www.fda.gov/cder/guidance/iche3.pdf>)*
- *Adverse drug reaction (ADR): Any noxious or unintended response to a medicinal product related to any dose may be considered to be an adverse drug reaction. Here there is at least a reasonable possibility that there is a causal relationship between the medicinal product and the adverse event. (see <http://www.ncehr-cnerh.org/english/gcp/>)*

*Note the distinction between AE, which encompass all adverse events occurring during the trial, and adverse drug reactions (ADR), which reflect only events thought to be related to drug exposure.*

## Criteria for including or excluding source data, and data to include

Applicants should:

- ⇒ State *a-priori* inclusion and exclusion criteria for selecting relevant studies – if at all.
- ⇒ Include all identified randomised controlled trials published as full articles in peer-reviewed journals in the English language that report (or give sufficient data to calculate) outcomes by intention-to-treat.

Notes:

*Inclusion criteria can include:*

- *precise treatment;*
- *precise indication for specific disease;*
- *specific kinds of patients;*
- *type of study (e.g., RCTs, other);*
- *type of publication (full articles, abstracts, conference proceedings);*
- *where published (e.g. Medline-indexed peer-reviewed journal);*
- *language (e.g. English only);*
- *type of analysis (e.g. reporting (or giving sufficient data to calculate) outcomes by intention-to-treat).*

*Similarly, exclusion criteria can include features such as:*

- *treatments apart from the precise treatment;*
- *indications for diseases apart from the precise indication and disease;*
- *non-controlled and/or non-randomised and/or non-blinded experimental studies and/or observational studies (including cohort, case control and cross-sectional studies);*
- *data not yet published as full articles in peer-reviewed English language journals indexed on Medline;*
- *data published as conference proceedings only;*
- *data on patients not including [kinds of patients];*
- *trials of patients with not [the specific disease/indication]; and*
- *trials where outcomes by intention-to-treat were neither reported nor calculable from raw data.*

## Quality of the evidence

Quality of evidence is assessed using four components:

1. Critical appraisal of individual studies and scoring for quality;
2. Hierarchy of evidence;
3. Relevance of the study design and outcome to the new technology being assessed; and
4. Applicability of the study results to individual patients and groups of patients.

### 1. Critical appraisal and Scoring for quality

Applicants should:

- ⇒ First sort relevant studies according to level of evidence and study type (described below in “2. Hierarchy of evidence”).
- ⇒ Ideally, then critically appraise each study, systematically considering the aspects of the study design that are most likely to be subject to bias.
- ⇒ Ideally, appraise evidence according to established checklists, e.g. the GATE checklists devised by Auckland University’s EPIQ (Effective Practice, Informatics & Quality Improvement) available at <http://www.health.auckland.ac.nz/comhealth/ElectronicGateInterV10.doc> (accessed 12 June 2003), or the evidence tables described on the NZ Guidelines Group website at [http://nzgg.org.nz/tools/med\\_literature.cfm](http://nzgg.org.nz/tools/med_literature.cfm).
- ⇒ For RCTs, ideally use PHARMAC’s following modification of the Jadad criteria<sup>21</sup> to summary score each trial (score 0-5):

Criterion (modified)	Source of bias (Cochrane Handbook taxonomy)	Scoring system
Randomisation	Selection bias / confounding, i.e. systematic differences in comparison groups	Adequate =1, Inadequate/nil = 0
Concealed allocation	Selection bias / confounding	Adequate =1, Unclear/not described = 0 Inadequate/nil = 0
Blinding of receipt	Performance bias, i.e. systematic differences in care provided apart from the intervention being evaluated; recipients	Adequate, described =0.5, Unclear/not described = 0.25 Inadequate/nil = 0
Blinding of provision	Performance bias; providers	Adequate, described =0.5, Unclear/not described = 0.25 Inadequate/nil = 0
Follow-up	Attrition bias, i.e. systematic differences in withdrawals from the trial, affecting outcome measurement	Participants adequately accounted for =1, Unclear/not described = 0 Inadequate/nil = 0
Blinding of assessment	Detection bias, i.e. systematic differences in outcome assessment; assessors	0 (presumably incorporated into Blinding of provision)

⇒ Also use the simple grading system described by the Cochrane Handbook to derive overall gradings (A-C) for each RCT:

<b>Risk of bias</b>	<b>Interpretation</b>	<b>Relationship to individual criteria</b>
A. Low	Plausible bias unlikely to seriously alter the results	All criteria <i>fully met</i>
B. Moderate	Plausible bias that raises some doubt about the results	One or more criteria partly met ( <i>where other criteria are fully met</i> )
C. High	Plausible bias that seriously weakens confidence in the results	One or more criteria not met

*Italicised words* supplement the text of the Cochrane Handbook.

Notes:

Good guides to critical appraisal are available on-line, for example at the McMaster University EBM site at <http://www.cche.net/usersguides/main.asp>. Clinical Evidence (<http://www.clinicalevidenceonline.org/>) contains similar approaches to appraisal (see <http://www.clinicalevidenceonline.org/lpBinCE/lpext.dll?f=templates&fn=main-hit-h.htm&2.0>).

Experts in New Zealand able to provide advice and training in critical appraisal include (but are not necessarily limited to) Auckland University's Institute EPIQ (Effective Practice, Informatics & Quality Improvement) (<http://www.health.auckland.ac.nz/population-health/epidemiology-biostats/epiq/>) and the New Zealand Health Technology Assessment Clearing House for Health Outcomes and Health Technology Assessment (NZHTA) (<http://nzhta.chmeds.ac.nz/>).

The Cochrane Handbook<sup>22</sup> recommends simple approaches for assessing validity that can be fully reported (i.e. how each trial scored on each criterion).

Jadad criteria, frequently cited as a scoring system (e.g. Cochrane Handbook, Bandolier<sup>23</sup>, systematic review 1999 commissioned by the NHS R&D HTA Programme<sup>24</sup>), allocate points for (1) (adequate) randomisation; (2) (adequate) blinding; (3) description of withdrawals and dropouts; and adequacy of methods/description of (4) randomisation and (5) blinding. The maximum score is 5 points, and studies scoring below 3 points are usually regarded as being of low methodological quality.<sup>25</sup> The Cochrane Handbook notes there is empirical evidence suggesting that, on average, both inadequate concealment of allocation<sup>26 27 28</sup> and lack of double blinding<sup>29 30</sup> result in over-estimates of the effects of treatment.

The PHARMAC modifications of the Jadad criteria largely relate those criteria to the sources of bias listed by the Cochrane Handbook, which include adequacy of allocation concealment ('randomisation blinding') and Blinding of assessment (for detection bias). PHARMAC modifications also separate performance blinding between recipients and providers.

The criteria used by the above simple A-C overall grading system (fully met, partly met, not met) are those of the above PHARMAC modified criteria (modified from Jadad criteria and Cochrane Handbook), i.e. Randomisation, Concealed allocation, Blinding of receipt, Blinding of provision, Follow-up, and Blinding of assessment.

Although the final quality score is only a guide, if a study rates poorly (factors are either not mentioned or are absent in the report) then that study is likely to be subject to significant biases, hence interpret its results with caution.

The Cochrane Handbook however notes that none of the currently available scales for measuring validity of trials can be recommended without reservation, and urges caution. The handbook notes empirical evidence of a relationship between parameters thought to measure validity and actual study outcomes is limited, and which validity criteria are the most important determinants are not yet known.

## 2. Hierarchy of evidence

Applicants should:

- ⇒ Classify component data according to PHARMAC's below modification of the Revised Scottish Intercollegiate Guidelines Network (SIGN) grading system<sup>31</sup>:
- 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, or RCTs with a low risk of bias
  - 1- Meta analyses, systematic reviews, or RCTs with a high risk of bias
  - 2++ High quality systematic reviews of case-control or cohort or studies; High quality case-control or cohort studies with a very low risk of confounding or bias and a high probability that the relationship is causal
  - 2+ Well conducted case control or cohort studies with a low risk of confounding or bias 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 significant risk that the relationship is not causal
  - 3a Non-analytic uncontrolled observational studies
  - 3b Case reports
  - 4 Expert opinion and/or epidemiological/economic modelling in absence of empirical data
  - 5 Other
- ⇒ Describe both the level of evidence used and denominate by a judgement of the likely potential level of evidence available, i.e.

Level of evidence:	level [xx]
Potential level of evidence available:	level [yy]

*Notes:*

*The relative weight carried by different type of studies is known as the hierarchy of evidence. There are a number of layers to consider, including:*

- *type of information (meta-analyses, RCTs, cohort studies, case-control studies, descriptive series, anecdotal/expert opinion);*
- *quality of information (e.g. meta-analyses: systematic identification of all RCTs, including and exclusion criteria, methods of combining data, conclusions based on data, etc. RCTs: allocation concealment, randomisation, blinding, intention-to-treat, follow-up, etc. Cohort studies: concurrent versus historical controls, etc.);*
- *homogeneity and heterogeneity of RCTs within meta-analyses; and*
- *precision and relevance of effects (clinically relevant effects, non-overlap of confidence intervals etc.).*

*Evidence ranges from the best/ideal (multiple RCTs with homogeneity of effects, identified systematically with sound inclusion/exclusion criteria, all with clinically relevant effects without overlapping confidence intervals) to least robust. There are a number of taxonomies for hierarchies of evidence available.*

*The Revised SIGN grading system<sup>32</sup> arose from the AHCPR grading system<sup>33 34</sup> and the MERGE checklist (NSW Department of Health)<sup>35</sup>, and was recommended by the New Zealand Guidelines Group. PHARMAC modifications consider features in other key systems<sup>36 37 38 39 40 41 42</sup>, and include*

*pseudo-randomised controlled trials and modelling and differentiate between observational studies and case reports.*

*PHARMAC gives the greatest weight to good-quality randomised controlled trials, but sometimes these are unavailable. In such circumstances, consider other levels of evidence. If good-quality randomised controlled trials are available, but are considered inappropriate for use as the basis of the application, then state why they were excluded.*

*Denominating the strength of evidence helps address some of the potential for disparity for funding data-rich items at the expense of outcomes that are more difficult to measure and where data are less available.*

- *For example, there may be one case-control study available for a particular intervention, when realistically a series of RCTs should be or is available (maybe showing negative results), as occurs with most pharmaceuticals. This would mean, say, ranking the evidence not as “level 2 [Well conducted case control or cohort studies with a low risk of confounding or bias and a moderate probability that the relationship is causal]”, but rather as “level 2 [... case control ... studies ...], out of possible 1++ [High quality ... RCTs with a very low risk of bias]”. This would alert decision-makers as to possible poor quality of evidence, relative to what potentially should be available.*

<i>Level of evidence:</i>	<i>level [2-]</i>
<i>Potential level of evidence available:</i>	<i>level [1++]</i>

- *Conversely, where evidence was difficult to obtain reliably because of methodological constraints (as might occur, say, with assessing the impact of some public health programmes), then this indicates that although the evidence is less than ideal, this is not through lack of trying - e.g. “level 3a out of possible 3a”, not just “level 3a”.*

<i>Level of evidence:</i>	<i>level [3a]</i>
<i>Potential level of evidence available:</i>	<i>level [3a]</i>

### 3. Relevance of the evidence

Applicants should:

⇒ For each study, assess its relevance according to the following relevance grades:

- |           |   |
|-----------|---|
| Grade I   | Evidence of effect of long term <i>patient focused outcomes</i> , taking into account benefits and harms, especially quality of life and survival |
| Grade II  | Evidence of effect on short to medium term <i>clinical outcomes</i> , taking into account benefits and harms                                      |
| Grade III | Evidence of effect confined to <i>intermediate or surrogate outcomes</i>  |

Notes:

*'Relevance' of the evidence describes how appropriate the study design and outcomes are to the new technology. This relates to the nature of the outcomes used and the applicability of the research results to other treatments, patients or clinical settings.*

*In assessing the evidence, PHARMAC gives greatest weight to studies that measure patient-focussed outcomes, then those with clinical outcomes.*

#### **Patient focused outcomes**

*Patient focused outcomes are any health outcomes that are meaningful to the patient. They often include items that are not readily amenable to objective measurement or quantification. They may include measures of clinical effect, adverse effects, tolerability and change in quality of life. These effects may need to be individualized for each treatment.*

Types of patient focused outcomes can include:

- all cause mortality and/or survival;
- cause-specific mortality;
- change in morbidity, subdivided according to type (e.g. hospital admission, nursing home requirements);
- side effects of treatment, including adverse reactions to drug therapies, and co-interventions that may be necessitated by the primary treatment; and
- disease specific outcome, including disease-specific quality of life measures.

### **Clinical outcomes**

*Clinical outcomes tend to relate to the disease being studied. In many cases the outcomes chosen are those appearing to clinicians to be of primary concern and that can be measured (e.g. survival in cancer, vertebral fracture in osteoporosis, peptic ulcer healing and relapse rates, walking distance in angina). However, this approach does not always necessarily capture all the relevant outcomes from the patient's or society's perspective. Such clinical outcomes may not reflect the patient's or other's quality of life.*

### **Surrogate/intermediate outcomes**

*'Surrogate/intermediate' outcomes are essentially biological markers. Commonly a physiological variable (e.g. serum LDL-cholesterol concentration, blood pressure), a surrogate/intermediate outcome has a statistical association with clinical outcome of interest (e.g. bone mineral density with fracture, CD4 cell concentrations with progression of HIV). There will also be a biological and pathophysiological basis for believing that the surrogate/intermediate outcome is a major determinant of the clinical outcome in the disease being studied (e.g. glycosylated haemoglobin (HbA1c) and diabetes complications).*

*A surrogate/intermediate outcome should possess all of the above features, but few do. If using a study with surrogate/intermediate outcomes as the basis for an application, then present evidence linking the surrogate/intermediate outcome with the clinical outcome of interest. Evidence relying on surrogate/intermediate outcomes alone will be unlikely to be sufficient base judgement about the true benefits of a technology.*

## **4. Applicability of the evidence**

Applicants should:

- ⇒ For each study, assess how applicable it is to the New Zealand health sector and pharmaceutical funding and delivery environment.

Notes:

*One of the main questions for applicability of clinical evidence, relevant to PHARMAC's considerations for listing pharmaceuticals on the Pharmaceutical Schedule or changing their access criteria, is "Can the results be applied in other settings and to other patients?"*

- *Are there any known biological factors that may alter the effect of the intervention?*
- *What effects does the timing of the intervention have?*
- *What effects do variations in the nature and severity of the disease have?"*

*In a wider setting (i.e. funding of non-pharmaceutical technology), the question of applicability also includes “ Can the intervention be reproduced in the proposed setting?*

- *Is the effectiveness of the intervention operator dependant?*
- *Is it a complex intervention with many components?*
- *Are the components of the intervention available in New Zealand?*
- *Is any infrastructure required/available, such as monitoring the intervention with regular blood tests?”*

*However, such questions may not be relevant to PHARMAC’s considerations for listing pharmaceuticals on the Pharmaceutical Schedule or changing their access criteria.*

## **Data synthesis**

Applicants should:

### **Data abstraction**

- ⇒ Consider using one or more people to abstract the data directly from tables and text in the published articles [ $\pm$  independently of others], using suitable software and a predefined protocol. Resolve differences in interpretation of the data for abstraction by discussion, with third party adjudication where needed.
- ⇒ For RCT quality, ideally enter points for each item [ $\pm$  independently of others] and then calculate grade and Jadad score (0-5) according to the algorithms described above. Assessment can be masked or unmasked. Resolve any differences in assigning points by discussion, with third party adjudication where needed.

### **Effectiveness**

- ⇒ Analyse by intention-to-treat. Where intention-to-treat analysis has not been reported, ideally recalculate effectiveness rates by subtracting from reported numbers of patients responding all patients who withdraw/drop-out/are otherwise lost to follow-up.
- ⇒ For effectiveness, ideally combine studies using standard meta-analytic statistical techniques to calculate odds ratios (with 95% confidence limits), using suitable software.
- ⇒ Ideally also calculate person-year weighted incidence rates, using these to calculate weighted rate ratios (relative risks) (aRR) and weighted relative risk reductions (aRRR), using the formulae described in the Appendix.
- ⇒ Also ideally calculate quality-weighted pooled odds ratios and quality-weighted adjusted baseline incidence rates, as described in the Appendix, and thus quality-weighted adjusted rate ratios, etc.

### **Withdrawals and continuations in randomised controlled trials**

- ⇒ Similarly combine RCTs as with effectiveness, calculating weighted absolute risk reductions.

## Assessment of missing data and possible publication bias

Ideally and if possible:

- ⇒ Use funnel plots to assess how complete the published or identified evidence is. This can be done using standard techniques, described elsewhere<sup>43</sup> (<http://bmj.bmjournals.com/cgi/content/full/315/7109/629>).

*Notes:*

*Funnel plot asymmetry may mean there is publication bias, along with other explanations. Funnel plots can be used to examine small study effects, i.e. whether smaller studies in a meta-analysis are tending to show larger treatment effects (which may be due to reasons other than publication bias).*

*Further details on funnel plots are available in the Cochrane Handbook (<http://www.cochrane.dk/cochrane/handbook/hbook.htm> Section 8.11.1 Publication bias and funnel plots).*

*There are also ready methods for testing for heterogeneity, e.g. the  $I^2$  test<sup>44</sup> (<http://bmj.bmjournals.com/cgi/content/full/327/7414/557>)*

*Beware however of important limitations with both funnel plots<sup>45</sup> and tests for heterogeneity<sup>46</sup>, well described by Bandolier in 2000 (<http://www.jr2.ox.ac.uk/bandolier/band81/b81-5.html>). Be cautious with their interpretation, as detailed in the Cochrane Handbook.*

## **Additional clinical opinion and clinical contacts**

Where expert opinion is used, applicants should:

- ⇒ Present expert opinion as an attachment to the proposal or as a technical document.
- ⇒ Justify the need for any expert opinion.
- ⇒ Describe the methods used to obtain and collate the opinions, by following the structured approach provided below. Then summarise the opinion obtained together with the extent of any variability in the opinions. Indicate how the opinions have been used in the application.
- ⇒ Provide the following details:
  - The criteria for selecting the experts;
  - The number of experts approached;
  - The number of experts who participated;
  - Whether a declaration of potential conflict(s) of interest was sought from all experts;
  - Medical specialty groups whose opinions were sought;
  - The background information provided and its consistency with the totality of the evidence provided in the submission;
  - The method used to collect the opinions;
  - The medium used to collect the opinions;
  - The questions asked;
  - Whether iteration was used in the collation of opinions and, if so, how it was used;
  - The number of responses received for each question;
  - Whether all experts agreed with each response, and, if not:
    - The approach used to finalise the estimates; and
    - The approach used to present the variability of the opinions.
- ⇒ Make sure any clinicians providing expert advice declare any potential conflict of interest. This may include financial interest in the development of the technology, likely financial gains arising from the proposed technology, research funding, etc.
- ⇒ Make sure there are no data supplied able to identify individual patients.

*Notes:*

*Expert opinion cannot substitute for sound scientific evidence, but will help interpret data, particularly results' relevance and potential impact. "Expert opinion" can be defined as opinion from groups with any relevant expertise in the area of concern, for example, consumer support groups or specialist professional societies.*

*Particular areas where expert opinion may be useful are:*

1. *Setting the context of the economic evaluation by defining the place of the proposed technology in treatment (the main indication and the main comparator);*
2. *Adjusting to meet local conditions, patterns of resource use and, very rarely, the clinical outcomes measured in randomised trials conducted in different settings, such as in other countries;*
3. *Predicting which resources will be used and how often each will be used to manage outcomes reported in the randomised trials but not followed up; and*
4. *Determining the potential impact of an outcome from a patient perspective where this information has not been captured in the trials.*

## Overall summary of benefits and harms; balance sheets

Applicants should:

- ⇒ Provide a summary table of the quality and strength of the studies identified. Where available, tables should include estimates of the absolute risk reduction (with 95% confidence limits):

Hierarchy Grade	Study ID	Critical Appraisal Score (incl modified Jadad score)	Outcomes	Relevance grade	Magnitude of effect (ARR with 95% CI)
[1++ to 4]	[author(s), year published]	[1 to 4] / 4	[state] [state]	[I to III] [I to III]	x% (y-z%) x% (y-z%)
[1++ to 4]	[author(s), year published]	[1 to 4] / 4	[state] [state]	[I to III] [I to III]	x% (y-z%) x% (y-z%)

Notes:

*Summary tables bring together the assessments of the quality and strength seen in studies identified by systematic literature search. Strength is measured by magnitude of effect, with precision measured by 95% confidence intervals.*

*It may also be useful to produce a balance sheet summary of all the clinical benefits and harms of the new technology that can be expected in a group of typical New Zealanders. For example, for a series of patient (e.g. 100 patients) estimate how many can be expected to benefit and how many can be expected to experience some form of adverse outcome from the new technology. Examples of balance sheets and their use can be found at [http://nzgg.org.nz/tools/balance\\_sheet.cfm](http://nzgg.org.nz/tools/balance_sheet.cfm).*

## **Review process**

Applicants should:

- ⇒ Describe the review process used to determine inputs into any cost-effectiveness, need and budgetary impact models you provide.

This might include the external review by New Zealand or international experts in clinical medicine, epidemiology, clinical pharmacology, primary health care, pharmacy, economics, the New Zealand healthcare system, and public policy.

## Appendix: Calculating person-year weighted incidence rates, weighted rate ratios and relative risk reductions

Person-year weighted incidence rates, weighted rate ratios (relative risks) (aRR) and weighted relative risk reductions (aRRR) can be calculated as follows:

nt	= no. of patients in treated group responding,
Nt	= no. patients in treated group,
nc	= no. of patients in control group responding,
Nc	= no. patients in control group
crude response rate for treated patients	= $\frac{\sum nt}{\sum Nt}$
crude response rate for control patients	= $\frac{\sum nc}{\sum Nc}$
crude rate ratio (relative risk, RR)	= $\frac{(\sum nt / \sum Nt)}{(\sum nc / \sum Nc)}$
crude odds ratio (OR)	= $\frac{(\sum nt / (\sum Nt - \sum nt))}{(\sum nc / (\sum Nc - \sum nc))}$

Adjusted baseline event rates (aEc) are derived by weighting control arms according to (t.Nc),

where	t	= study duration,
	Nc	= no. patients in control group, and
	t.N	= risk exposure, in person-year equivalents

and where the standard error (similar to variance) of a proportion p is inversely proportional to  $\sqrt{Nc}$ ,

ie	SE	= $\sqrt{[(p.(1-p)) / Nc]}$
where	p	= $nc/Nc$

Pooled (adjusted) odds ratios for all studies (aOR) are weighted according to the inverse variance of individual RCTs' odds ratios, with associated confidence limits (fixed effects, Peto one-step method)

Adjusted rate ratios (aRR) are derived from adjusted baseline event rates and pooled odds ratios, with associated confidence limits, according to the formula<sup>iv</sup>:

$$aRR = \frac{1 - \frac{(1-aEc).(1-aOR)}{1 - [aEc.(1-aOR)]}}{1 - [aEc.(1-aOR)]}$$

where:	aRR	= adjusted rate ratio (ie adjusted relative risk)
	aEc	= adjusted baseline event rate (ie control incidence rate, weighted according to (t.N))
	aOR	= pooled (adjusted) odds ratio (weighted according to inverse variance)

Adjusted relative risk reductions (aRRR) are derived from adjusted rate ratios, where  $aRRR = 1 - aRR$ , according to the formula:

$$aRRR = \frac{(1-aEc).(1-aOR)}{1 - [aEc.(1-aOR)]}$$

where:	aRRR	= adjusted relative risk reduction
	aRR	= adjusted rate ratio
	aEc	= adjusted baseline event rate
	aOR	= pooled (adjusted) odds ratio

If adjusted baseline event rates are considered relevant to the New Zealand population, adjusted absolute risk reductions (aARR) are derived from adjusted baseline event rates and adjusted rate ratios, according to the formula:

$$aARR = aEc.aRRR$$

where:	aARR	= adjusted absolute risk reduction
	aEc	= adjusted baseline event rate
	aRRR	= adjusted relative risk reduction

Similarly, if adjusted baseline event rates are considered relevant to the New Zealand population, adjusted treatment event rates (aEt) are derived from adjusted baseline event rates and adjusted rate ratios, according to the formula:

$$aEt = aEc.aRR$$

where:	aEt	= adjusted treatment event rate (treated incidence rate)
	aEc	= adjusted baseline event rate
	aRR	= adjusted rate ratio (relative risk, derived from pooled odds ratio)

[NB Odds ratios derive from relative risks according to the formula:

$$OR = \frac{RR.(1-Ec)}{1 - (RR.Ec)}$$

$$= \frac{Et .((1/Ec)-1)}{1 - Et}$$

where:	OR	= odds ratio
	RR	= rate ratio (ie relative risk)
	Ec	= baseline (control) event rate
	Et	= treatment event rate

]

<sup>iv</sup> algebraic transformation by PHARMAC of formulae in Sackett D, Straus S, Richardson WS, Rosenberg W, Haynes B. Evidence-based medicine: how to practice and teach EBM, 2<sup>nd</sup> edition. Oxford: Churchill Livingstone, 2000. p136 Table 5.1 Formulae to convert odds ratios (ORs) and relative risks (RRs) to NNTs.

To account for the quality of contributing RCTs, weight each RCT according to its Jadad score (0-5). Combine these quality-based weights with the above [variance-based weights for odds ratios] and the [exposure-based (t.N) weights for adjusted baseline incidence rates], giving quality/variance weights and quality/exposure weights. Use these quality-containing weights to then calculate quality-weighted pooled odds ratios and quality-weighted adjusted baseline incidence rates, using the above formulae, and thus quality-weighted adjusted rate ratios, etc.

## References/endnotes

- 
- <sup>1</sup> Metcalfe S, Dougherty S, Brougham M, Moodie P. PHARMAC measures savings elsewhere to the health sector. *N Z Med J.* 2003;116:U362.
- <sup>2</sup> Pharmaceutical Benefits Scheme (PBS). 2002 Guidelines for the Pharmaceutical Industry on Preparation of Submissions to the Pharmaceutical Benefits Advisory Committee including major submission involving economic analyses.
- <sup>3</sup> Pharmaceutical Benefits Scheme (PBS). 1995 Guidelines for the Pharmaceutical Industry on Preparation of Submissions to the Pharmaceutical Benefits Advisory Committee including major submission involving economic analyses.
- <sup>4</sup> Interim Document to accompany the Guidelines for the Pharmaceutical Industry on Preparation of Submissions to the Pharmaceutical Benefits Advisory Committee: including major submissions involving economic analyses, April 2000.
- <sup>5</sup> NHS Centre for Reviews and Dissemination. Undertaking Systematic Reviews of Research on Effectiveness: CRD's Guidance for those Carrying Out or Commissioning Reviews. CRD Report Number 4 (2nd Edition). York: University of York, March 2001.
- <sup>6</sup> National Institute for Clinical Excellence. Guide to the Methods of Technology Appraisal (reference N0515). May 2004. [http://www.nice.org.uk/pdf/TAP\\_Methods.pdf](http://www.nice.org.uk/pdf/TAP_Methods.pdf)
- <sup>7</sup> Guidelines for economic evaluation of pharmaceuticals: Canada. 2nd edition. Canadian Coordinating Office for Health Technology Assessment (CCOHTA), 1997. Appendix 5. <http://www.ccohta.ca/>
- <sup>8</sup> Center For Drug Evaluation And Research. Manual of policies and procedures: Clinical Review Template. MAPP 6010.3. FDA 2004. <http://www.fda.gov/cder/mapp/6010.3.pdf>
- <sup>9</sup> The Burden of Disease and Injury in New Zealand. Public Health Intelligence Occasional Bulletin No. 1. Wellington: Ministry of Health, 2001.
- <sup>10</sup> Selected morbidity data for publicly funded hospitals 1997/8. Wellington: New Zealand Health Information Service, Ministry of Health, 1999. Tables 1 and 2.
- <sup>11</sup> Mental health data 1993. Wellington: New Zealand Health Information Service, Ministry of Health, 1996 (Tables 5, 5a, 9, and 9a).
- <sup>12</sup> Mortality and demographic data 1997. Wellington: New Zealand Health Information Service, Ministry of Health, 2000 (Tables 4 to 5a).
- <sup>13</sup> <http://www.ncbi.nlm.nih.gov/entrez/query.fcgi>
- <sup>14</sup> Selected morbidity data for publicly funded hospitals 1997/8. Wellington: New Zealand Health Information Service, Ministry of Health, 1999. Tables 1 and 2.
- <sup>15</sup> Mental health data 1993. Wellington: New Zealand Health Information Service, Ministry of Health, 1996 (Tables 5, 5a, 9, and 9a).
- <sup>16</sup> Mortality and demographic data 1997. Wellington: New Zealand Health Information Service, Ministry of Health, 2000 (Tables 4 to 5a).
- <sup>17</sup> <http://www.stats.govt.nz/>
- <sup>18</sup> Seeking the Evidence: a protocol. School of Health and Related Research (ScHARR), University of Sheffield.
- <sup>19</sup> Haynes RB, Wilczynski N, McKibbon KA, Walker CJ, Sinclair JC. Developing optimal search strategies for detecting clinically sound studies in MEDLINE. *J Am Med Inform Assoc* 1994;1:447-58

---

<sup>20</sup> The TRIP database at <http://www.tripdatabase.com> amalgamates 61 databases of 15,000 hyperlinks from 'Evidence-based' sites internationally: ACP Journal Club [ACP]; Agency for Healthcare Research and Quality AHRQ - Evidence-based practice reports; Annals of Internal Medicine [AIM]; Annals of Rheumatic Diseases; Archives of Dermatology; Archives of Disease in Children (including Fetal and Neonatal Edition; Archives of Family Medicine; Archives of General Psychiatry; Archives of Internal Medicine; Archives of Neurology; Archives of Ophthalmology; Archives of Otolaryngology; Archives of Pediatrics and Adolescent Medicine; Archives of Surgery; ARIF; ASERNIP-S; Bandolier; British Journal of Ophthalmology; British Medical Journal; CMA Guidelines; Cochrane Database of Systematic Reviews; Critically Appraised Papers; Critically Appraised Topics (CAT) Bank; DARE; DEC; Effective Health Care Bulletins; Effectiveness Matters; eMedicine.com; Evidence Centre Reports-pdf; Evidence-Based Medicine; Evidence-Based Mental Health; Evidence-Based Pediatrics; Family Medicine Research Reviews; Family Practice; Guideline [PRISE]; Gut; Health Evidence Bulletins Wales; Health Technology Assessment (HTA) Database [HTAD]; Health Technology Assessment Programme [HTAP]; Heart; JAMA; Journal Club on the Web; Journal of Epidemiology and Community Health; Journal of Medical Genetics; Journal of Neurology, Neurosurgery and Psychiatry; MD Digests; The American National Cancer Institute; National Guideline Clearinghouse; National Institute for Clinical Excellence [NICE]; NEJM; NHS Economic Evaluation Database; NIH Consensus Statements [NIHcon]; POEMs; Postgraduate Medical Journal; PRODIGY; Scottish Health Purchasing Information Centre [SHPIC]; SIGN; South and West R&D Briefing Papers [SWRDBP]; Swedish Council on Technology Assessment; Thorax; Virtual Hospital Family Practice

<sup>21</sup> Jadad AR, Moore RA, Carroll D, Jenkinson C, Reynolds DJ, Gavaghan DJ, McQuay HJ. Assessing the quality of reports of randomized clinical trials: is blinding necessary? *Control Clin Trials*. 1996;17:1-12.

<sup>22</sup> Clarke M, Oxman AD, editors. Assessment of study quality. *Cochrane Reviewers Handbook 4.1.5* [updated April 2002]; Section 6. In: *The Cochrane Library*, Issue 2, 2002. Oxford: Update Software. Updated quarterly.

<sup>23</sup> Bias. *Bandolier* October 2000 (issue no. 80) <http://www.jr2.ox.ac.uk/Bandolier/band80/b80-2.html>

<sup>24</sup> Moher D, Cook DJ, Jadad AR, Tugwell P, Moher M, Jones A, Pham B, Klassen TP. Assessing the quality of reports of randomised trials: implications for the conduct of meta-analyses. *Health Technol Assess*. 1999;3(12):i-iv, 1-98.

<sup>25</sup> Rowe BH, Spooner C, Ducharme FM, Bretzlaff JA, Bota GW. Early emergency department treatment of acute asthma with systemic corticosteroids (Cochrane Review). In: *The Cochrane Library*, Issue 3, 2000. Oxford: Update Software.

<sup>26</sup> Chalmers 1983. Chalmers TC, Celano P, Sacks HS, Smith H. Bias in treatment assignment in controlled clinical trials. *N Engl J Med* 1983; 309:1358-61.

<sup>27</sup> Schulz KF, Chalmers I, Hayes RJ, Altman D. Empirical evidence of bias. *JAMA* 1995; 273:408-12.

<sup>28</sup> Moher D, Pham B, Jones A, Cook DJ, Jadad AR, Moher M, Tugwell P, Klassen TP. Does quality of reports of randomised trials affect estimates of intervention efficacy reported in meta-analyses? *Lancet* 1998;352:609-13.

<sup>29</sup> Colditz GA, Miller JN, Mosteller F. How study design affects outcomes in comparisons of therapy. I: medical. *Stat Med* 1989; 8:441-54.

<sup>30</sup> Schulz et al 1995, *op. cit.*

<sup>31</sup> Scottish Intercollegiate Guidelines Network (SIGN). SIGN 50: A guideline developers' handbook. SIGN Publication No. 50, February 2001. <http://www.show.scot.nhs.uk/sign/guidelines/fulltext/50/index.html>. Section 6.3 Levels of evidence and grades of recommendation. <http://www.show.scot.nhs.uk/sign/guidelines/fulltext/50/section6.html>

<sup>32</sup> Revised SIGN 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, or RCTs with a low risk of bias
- 1- Meta analyses, systematic reviews, or RCTs with a high risk of bias
- 2++ High quality systematic reviews of case-control or cohort or studies; High quality case-control or cohort studies with a very low risk of confounding or bias and a high probability that the relationship is causal
- 2+ Well conducted case control or cohort studies with a low risk of confounding or bias 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 significant risk that the relationship is not causal
3	Non-analytic studies, eg case reports, case series
4	Expert opinion

deriving from four overall levels of evidence and with overall levels 1 and 2 each subdivided into three quality-rated sublevels:

Level 1	Randomised controlled clinical trials
Level 2	Non-randomised controlled analytical studies (non-randomised interventional, observational cohort, case-control)
Level 3	Non-analytic uncontrolled observational studies (cross sectional studies, prospective longitudinal follow-up studies, retrospective follow-up case series, case reports)
Level 4	Expert opinion and/or modelling in absence of empirical data

++	All or most criteria are met; Where criteria are not met, conclusions are thought very unlikely to alter.
+	Some of criteria are met; Where criteria are not fulfilled or are not adequately described, conclusions are thought unlikely to alter.
– (minus)	Few or no criteria are met; Where criteria are not fulfilled or are not adequately described, conclusions are thought likely or very likely to alter.

<sup>33</sup> US Department of Health and Human Services. Agency for Health Care Policy and Research. Acute pain management: operative or medical procedures and trauma. Rockville (Md): The Agency, 1993. Clinical Practice Guideline No. 1. AHCPR Publication No.92-0023. p107.

<sup>34</sup> Hadorn DC, Baker D, Hodges JS, Hicks N. Rating the quality of evidenced for clinical practice guidelines. *J Clin Epidemiol* 1996;49:749-54.

<sup>35</sup> Liddle J, Williamson M, Irwig L. Method for evaluating research and guideline evidence. Sydney: New South Wales Department of Health, 1996.

<sup>36</sup> Canadian Task Force on the Periodic Health Examination: The periodic health examination. *CMAJ* 1979;121:1193-1254.

<sup>37</sup> US Preventive Services Task Force. Guide to clinical preventive services: an assessment of the effectiveness of 169 interventions (ed M Fisher). Baltimore: Williams and Williams, 1989. Appendix A, p388.

<sup>38</sup> Oxman AD. Checklists for review articles. *BMJ*. 1994 Sep 10;309(6955):648-51.

<sup>39</sup> Guyatt GH, Sackett DL, Sinclair JC, Hayward R, Cook DJ, Cook RJ. Users' guides to the medical literature. IX. A method for grading health care recommendations. *JAMA* 1995;274:1800-4.

<sup>40</sup> National Health and Medical Research Council. A guide to the development, implementation and evaluation of clinical practice guidelines. Canberra: NHMRC, 1999.

<sup>41</sup> Sackett D, Straus S, Richardson W, Rosenberg W, Haynes R. Guidelines. Evidence-based Medicine: How to practice and teach EBM. 2nd ed. Edinburgh: Churchill Livingstone, 2000. Cited in Box 8.4 Grading of practice recommendations. In: Khalid S Khan, Gerben ter Riet, Jos Kleijnen. STAGE III Reporting and Dissemination PHASE 8 The report and recommendations. [http://www.york.ac.uk/inst/crd/crd4\\_ph8.pdf](http://www.york.ac.uk/inst/crd/crd4_ph8.pdf). In: Khan KS, Ter Riet G, Glanville J, Sowden AJ, Kleijnen J (eds) for the NHS Centre for Reviews and Dissemination (CRD). Undertaking Systematic Reviews of Research on Effectiveness. CRD's Guidance for Carrying Out or Commissioning Reviews. 2nd Edition. CRD Report No. 4. York: NHS Centre for Reviews and Dissemination, University of York, 2001. <http://www.york.ac.uk/inst/crd/report4.htm>

<sup>42</sup> Centre for Evidence-Based Medicine. Levels of Evidence and Grades of Recommendations. <http://cebm.jr2.ox.ac.uk/docs/levels.html>

<sup>43</sup> Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ*. 1997;315:629-34.

<sup>44</sup> Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ*. 2003;327:557-60.

---

<sup>45</sup> Tang JL, Liu JL. Misleading funnel plot for detection of bias in meta-analysis. *J Clin Epidemiol.* 2000;53:477-84.

<sup>46</sup> Gavaghan DJ, Moore RA, McQuay HJ. An evaluation of homogeneity tests in meta-analyses in pain using simulations of individual patient data. *Pain.* 2000;85:415-24.