Dominic Tilden
Director
THEMA Consulting Pty Ltd

14th September 2015

A SUBMISSION TO THE PHARMACEUTICAL BENEFITS ADVISORY COMMITTEE GUIDELINES REVIEW FROM THEMA CONSULTING

Disclaimer: The views in this submission are personal and do not necessarily reflect those of any of our clients.

About THEMA Consulting

I have been working in the preparation of PBAC submissions since 1999 and I established THEMA Consulting in 2006. THEMA employs a multi-disciplinary team of professional staff with tertiary qualifications which include economics, health economics, public health, biochemistry, applied science, epidemiology, biostatistics and pharmacy. THEMA have contributed to in excess of 100 PBAC submissions since the inception of the 2006 Guidelines.

I offer this background because, as an organisation, we believe we would be one of the most prolific users of the PBAC Guidelines. As such we hope to provide, by way of this submission, some insight to the Review about the use of the Guidelines in practice. Please accept that these notes are in no way meant as a criticism of the people who developed the current Guidelines. They are meant as opinions about how the Guidelines could be improved/streamlined based on experience of working with the Guidelines. It is much easier to see the types of anomalies I describe with hindsight.

General, not specific, solutions

At various places I think the Guidelines are guilty of trying to second guess specific situations and provide solutions for those specific situations. This results in inconsistencies with respect to what the Guidelines are trying to achieve overall, or even inconsistencies with some other part of the Guidelines.

I think a good example of this is the three separate Section B's that exist in the Guidelines. Each Section B appears to have evolved by trying to solve the same problem (i.e. what is the comparative effective of this treatment relative to its comparator) with three specific methods. As soon as these specific methods are outdated or not possible in a given submission the Guidelines themselves are outdated. Providing general solutions to general problems will minimise the extent to which the Guidelines can become outdated.

Another example is the skipping of Section C for cost-minimisation analyses. This is a specific instruction which does not necessarily have a generalisable justification and is therefore not always appropriate. The executive summary specifically asks for the results of "the three steps" of the stepped economic evaluation. Just about every economic model ever presented to PBAC would be more nuanced than what can be described in three steps.

PBAC practice versus international practice

THEMA's work is predominantly PBAC and we are not necessarily the most qualified to comment on how PBAC practice compares with international practice. I am sure the review will consider these comparisons systematically. However, based on our experience working with local pharmaceutical companies and their international counterparts I think the following represent at least some areas in which PBAC practice differs from international practice:

- the lack of use of mixed treatment comparison / network meta-analysis (MTC/NMA)
- the lack of use of probabilistic sensitivity analysis (PSA) in economic evaluation
- mapping of utilities from disease specific quality of life instruments
- the "stepped" economic evaluation" in PBAC submissions
- the discount rate

Mixed treatment comparison / Network meta-analysis

In December 2008, the Indirect Comparison Working Group (ICWG) report concluded "There is insufficient basis to consider moving from the approach of Section B(i) of the current 2007 PBAC Guidelines" (Section 6.4; page 52). This advice was, at least in part, predicated on the assumption "that PBAC primarily assesses each proposed drug against its main comparator, which can usually be done through pairwise comparisons." (Section 6.3; page 52).

I presented a critique of the ICWG and the role of MTC/NMA to the ARCS conference in 2013 (available upon request). In this presentation I noted – based on information in PSDs – that 40% (9/23) of major submissions to the March 2013 PBAC meeting had more than one comparator and in these nine submissions, there was an average of 3.9 comparators.

I will consider the proliferation of comparators elsewhere in this submission. However, if the number of comparators in submissions grows, but the methods for presenting comparative evidence do not keep pace, then it gives the impression submissions are being set up to fail.

I don't think it is the role of the Guidelines (or its associated appendices and working group reports) to deliberately recommend for or against particular methodologies. Rather, the Guidelines should be used to describe the conditions under which these methodologies would be more/less preferred by the PBAC.

Probabilistic sensitivity analysis in economic evaluation

As above, I don't think the PBAC Guidelines should be used to dismiss methodologies the rest of the world is using. Rather, the Guidelines should be saying things like: This is what PSA is; this is how we would prefer to be seeing it done; these are examples of some of the circumstances in which we believe it will add value to a submission. Such guidance would be very helpful because whilst exceedingly rare, we have had circumstances where the evaluation or the ESC have said "you should have or could have performed a PSA".

Mapping of utilities from disease specific quality of life instruments

Appendix 7 of the Guidelines is dedicated to the measurement of utility weights. This is anecdotal but it appears there has been a reasonable proliferation of these types of studies across a range of therapeutic areas over the last decade. I can think of instruments in asthma (St George's Respiratory Questionnaire), eye disease (VFQ-25) and oncology (EORTC QLQ-30) to name a few. From a practitioners point of view I think these mapping instruments can be really useful in the absence of direct MAUI data form a clinical trial and do have a place in the hierarchy of evidence. Some insight in to the PBAC's view of these methodologies would be a worthwhile addition to the Guidelines.

Also, Appendix 7 seems to offer only two ways of extracting utility data; (i) straight from the trial by comparing utility values across the randomised groups or, (ii) completely separated from the trial by use of scenario based utility valuation. We – and I am sure we are not alone – have used alternate methodology whereby utility weights are measured across patients achieving different surrogate outcomes (or residing in different health states) during the trial. As the Guidelines acknowledge, measuring differences in utility weights across treatment groups of randomised trials can be difficult because of a lack of power and a lack of sensitivity in the instruments. However, what the MAUIs are much better at detecting is quality of life improvements/differences

amongst patients who achieved certain health outcomes versus those who do not. I think this approach is relatively widespread and could be incorporated within the Guidelines.

The "stepped" economic evaluation" in PBAC submissions

This is an example of a practice we use that the rest of the world does not appear to. I describe below why I believe the stepped economic evaluation is no longer necessary and could be replaced with model to trial validation instead.

Discount rate

I mention this here mainly because I think it is a feature of our reimbursement system which could be subject to comparison with other jurisdictions. Our discount rate has been 5% for at least the 16 years I have been using it and it might just be stuck at 5% because of convention and not actually reflect our society's time or risk preferences. What other discount rates are used internationally and how are these rates justified?

Identification of comparator(s)

My experience, at least anecdotally, and to some extent supported by the evidence on the number of comparators per submission above is the identification of comparators in PBAC submissions is becoming increasingly difficult. We find our submissions can include multiple comparators, in multiple appendices, for fear of nominating the "wrong" comparator.

The fact that some submissions have upwards of five comparators points to a failure of the Guidelines to provide adequate definitions of what constitutes the comparator for purposes of PBAC decision making. Naming additional comparators in a submission is easy because it is almost always true that a new drug could, would or should replace more than one treatment. Actually presenting a submission with multiple comparators is not so easy. There are few economies of scale (especially when MTC/NMA cannot be used) associated with adding comparators in a submission. A submission with two comparators is going to be almost twice as big as a submission with one. Such submissions also lack the clarity of submissions with single comparators. As I mentioned earlier in this submission, submissions to PBAC with multiple comparators are just about set up to fail from the outset.

Solving this is likely to prove difficult. However, I do believe the Review should be considering aspects of comparator selection.

Perhaps there could be a comparator hierarchy spelt out explicitly. For example, there is a PBS listed pharmacological analogue. Therefore this is the comparator.

Perhaps there could be greater reliance on PBAC decision making precedents. For example, new drug "A" has three potential comparators B, C and D all of which are PBS listed. Drug B has previously been accepted as cost-effective over C and D which were cost-minimised against each other. Drug B is the comparator because if A is cost-effective to B then A is, by definition, cost-effective to C and D. So rather than a full PBAC submission against three comparators we have only one main comparator together with some evidence to confirm the original conclusions of the PBAC with respect to B, C and D are still valid. Such an approach may be difficult to put in practice when the prices of drugs change differently over time. In fact, it might be this phenomenon which has contributed to the recent proliferation of comparators. Therefore, it would be sensible for the Guidelines to consider the differential pricing of comparators in their updated recommendations with regard to appropriate comparator selection.

I appreciate including explicit advice within the Guidelines might make it difficult for the PBAC to be sufficiently flexible with respect to comparator selection. However, this flexibility does come at a cost. Perhaps the Guidelines could be re-worded, or the emphasis changed to say the PBAC will accept the definition of the comparator according to the Guidelines unless they find compelling evidence to deviate from them.

Unregistered comparators / Cost-ineffective comparators

I am not sure whether this is just personal experience or something that needs to be addressed more broadly. From my experience some guidance on how to approach the cost-effectiveness analysis when the comparator is not registered or not considered cost-effective could be an important inclusion to the Guidelines.

The PSDs provide some precedents in that they do say at various times the comparator itself must be shown to be cost-effective before the new drug itself can be considered cost-effective by way of a comparison to it. This much is true (and analogous to my B, C and D example above). However, neither the PSDs nor the Guidelines go on to say what the approach should be in the circumstance where the comparator is not considered cost-effective. It is tempting to say the comparator should be "placebo" or something similar to establish the new treatment is cost-effective in a general sense. However, if the cost-ineffective comparator is really expensive and doesn't work, and the new drug is better and cheaper, but still not cost-effective versus placebo, should PBAC be recommending or rejecting this drug?

Some ideas for consideration

To help streamline the PBAC Guidelines (a stated objective of the review) and therefore make preparation and assessment of submissions shorter, clearer and more consistent I provide a few ideas and observations about how the Guidelines are used within each section of the Guidelines. I hope these ideas are helpful.

Executive summary

The structure of the executive summary could have better alignment with the structure of the submission document itself.

Section A

In my experience and from informal discussions with current and former PBAC members the one consistent theme that marks a successful submission is a clear communication of the clinical need for the treatment.

However, the Guidelines do not have an explicit section where the clinical need for the treatment which is the subject of the submission can be described. There are places where the clinical need can be "shoe horned" in to the submission. For example, in a preamble to the submission, or in Section A.5 where the algorithm is described. However, I suspect this is done differently in different submissions, with a different context and without describing all the aspects of the clinical need at the same time.

I understand the PBAC would not necessarily want such a section to turn in to an opportunity for a misdirected or completely undirected "pitch" for the new treatment (I am sure my pharmaceutical company colleagues will forgive me for acknowledging this potential – I have been guilty of this myself). However, if done well, I believe this section could help tie together all the main elements or "moving parts" of a submission. It could provide context for the general purpose and direction of the submission, why the restriction is written the way it is, why a particular comparator is more/less relevant, why particular clinical trial endpoints are more/less important to patients, why the health states of an economic model have been formulated the way they have been and so on. Having such a section in the Guidelines about how to describe the clinical need for a product would enable the PBAC to tell those preparing submissions what it is about the clinical need for a product which is most compelling to them.

Sections A.1 and A.2 are repetitive, both ask for Requested PBS listing. Perhaps Section A.1 could include two main subsections. Section A.1.1 presents what the drug is (pharmacological class and action, ATC classification etc.) and Section A.1.2 summarises why we need it (as per the missing clinical need section I refer to above). Then Section A.2 can focus on the proposed PBS restriction.

Section A.3 is asking the submission to describe in detail other relevant therapies even before we have determined the comparator or even looked at the clinical management algorithm. If it is necessary perhaps it

belongs at the end of Section A. That is, first identify what the relevant therapies are, then described them in detail.

Sections A.4 (Main comparator) and A.5 (Clinical management algorithms) also feel as though they are the wrong way around. More often than not, the main comparator is justified with reference to the clinical management algorithm. I think it would make submissions just that little bit easier to prepare if we didn't need to refer forward in to a document to justify the main comparator.

Section B

Section B is too prescriptive and too dense. We do not need three alternative Section B's for different levels of evidence. Rather, the objective of Section B should be clearly stated in the Guidelines in a way which is generalisable to all submissions. For example, "Section B seeks to provide a complete systematic evaluation of the best available evidence for the proposed intervention relative to its main comparator(s)"

If that best available evidence happens to be direct randomised trials or indirect randomised trials or non-randomised trials (as per each of our Section B's at the moment) then so be it. The current Guidelines suggest the different types of Section B should be completed for Sections B.3 to B.8. I believe this can be dramatically streamlined. Sections B.3, B.4 and B.5 probably do not need to change for one type of evidence base to the next. The way in which the results are presented or conclusions drawn from the presented evidence doesn't need to be specific to the nature of the comparisons (i.e. direct or indirect) until Section B.6. That is: present the details of your evidence (Sections B.3 to B.5), present the results of these studies (Section B.6). Then present the results of these studies in a manner which allows comparative therapeutic conclusions to be drawn (with if necessary, references to appendices in the Guidelines for the latest or preferred methodology with respect to indirect comparisons and use of non-randomised evidence).

Section B.7 probably does not need to be specific to the type of evidence presented in Sections B.3 to B.5. In fact, it appears to be explicitly designed for the presentation of evidence collected outside the usual frameworks. Section B.8 should definitely not be specific to the type of evidence collected. It is a **general** conclusion which relies upon the strength of the underlying evidence. Everything about Section B should be leading to this one conclusion because everything which follows from Section B.8 relies upon this conclusion. The Guidelines should not allow for multiple Section B.8's.

Section C

I believe the inclusion of Section C in the Guidelines in 2006 was a worthwhile introduction. It forced us to think about the quality and applicability of all the evidence making up an economic evaluation. However, when it comes to the practice of preparing submissions Section C is probably in the wrong location which leads to repetition and can lead to inconsistent consideration of different pieces of evidence.

In terms of repetition the nature and design of a pre-modelling study is a function of both the clinical trial data we have and the economic model structure we want to fit the data in to. This means we often have to describe the economic model we will be using in order to justify why we are conducting the pre-modelling study we are. However, we don't get the opportunity to describe and justify the model structure until Section D.3.

In terms of inconsistent consideration of different pieces of evidence, Section C exposes the clinical trial data to special consideration of its applicability, its use over time (extrapolation) in an economic evaluation and its transformation to patient relevant metrics – as it should be. However, other variables in an economic evaluation – which may or may not be just as important to the results of an evaluation – are not necessarily subject to such scrutiny. A good example of this inconsistency is the recommendation in the Guidelines that Section C is skipped in the case of cost-minimisation analysis. Why is the applicability of trials used to conclude non-inferiority not relevant? Why can we ignore extrapolation issues with respect to a claim of non-inferiority based on a short term trial?

Perhaps the pre-modelling studies should become part of Section D. In the case of cost-effectiveness analysis the pre-modelling studies will be required in Section D.4 where the variables of the economic model are

presented and justified. The methods for conducting pre-modelling studies as they are described in the Guidelines are relatively sound and should be maintained. In the case of cost-minimisation analysis issues of applicability, extrapolation and transformation could be addressed in Section D.1 before the calculation of equi-effective doses and the cost-minimisation analysis itself.

Section D

Generally speaking the design and structure of the two alternative Section D's (one for cost-minimisation and one for cost-effectiveness) is reasonable.

Personally, I do not like the stepped economic evaluation. It is often shoe-horned in to the submission at the last minute. Also, we try and manipulate the order of the stepped economic evaluation so as to hide results which are either inconvenient, or which we believe could be unfairly or inappropriately misinterpreted by the evaluation, the ESC or the PBAC. Extrapolation variables are often a good example of how the stepped economic evaluation can be misinterpreted. For example, I could present a stepped economic evaluation which shows no change (or even an increase) in the cost per responder over time, and then I transform to QALYs and I have an incremental cost per QALY. This appears in the stepped economic evaluation as extrapolation having a minimal (or deleterious) impact on the ICER. The same stepped economic evaluation could calculate the incremental cost per QALY from the trial and extrapolate it over time but the ICER will decrease dramatically over time because the value of the response accumulates over time faster than do the costs. This is especially true in the case of economic models using stopping/continuation rules.

I think the stepped economic evaluation can be replaced with a section dedicated to validating the economic model against the clinical trial data. Does the model predict what is happening in the clinical trial? If so, then good, and the rest of the model can be assessed on its merits. If not, then why not. Is it because the model uses a subgroup? Is it because the model uses a different dose? Is it because the model uses per protocol instead of ITT data? Once this discrepancy has been confirmed, explained and justified the rest of the model can be assessed on its merits.

More generally and related to the stepped economic evaluation is an apparent inconsistency with the way the Guidelines advise on appropriate model time horizon and how they are interpreted by the PBAC. I have not examined this specifically but anecdotally my experience would suggest that if I was to refer to various PBAC decisions on appropriate time horizons I would see a range of recommendations from 1 year to 5 years to 10 years to lifetime models. That variability could be expected for different disease areas – however again my experience suggests these time horizons are selected arbitrarily more than anything else.

The Guidelines determine an appropriate time horizon as relating to "the natural history of the medical condition, the treatment patterns, and an estimation of the time period(s) over which outcomes from the two therapies would be expected to occur". I think the last word, "occur" needs to change to "differ". Specific outcomes might occur over the short term but their impact may last a lifetime.

When dealing with a chronic condition a lifetime model horizon is probably the correct model horizon most of the time (or at least until such time as the treatment is no longer different to the comparator). When dealing with a treatment which saves lives or prevents some other permanent disability (e.g. stroke) a lifetime model is always the correct model time horizon. The Guidelines should be used to make this clear. If in a given submission PBAC have reservations about the quality of the data used to populate a lifetime model then that can be assessed on its merits.

I can understand the PBAC's reticence for models which extrapolate over long time periods and the uncertainty this inevitably brings. However, we have a discount rate to capture at least some of this uncertainty. When arbitrary model time horizons of one year, five year and so on are employed we are essentially assigning a 100% discount rate to all costs and outcomes after this period. The implications of arbitrarily cutting models short could be made explicit in the Guidelines. That is, it will inevitably lead to a bias towards interventions with short term effects.

Section E

Section E could also benefit from a re-ordering of some of the sections. I don't think we need a section (Section E.1) dedicated to "Justification of the selection of sources of data". This should happen as a matter of course whenever data are presented in any analysis. Also, it is difficult to justify a given data source without first providing the context in which it is used.

Perhaps Section E.1 could be dedicated to describing the "structure" of the budget impact analysis (market share or epidemiological approach). Section E.2 would populate this structure with the necessary data and Section E.3 presents the results of the analysis. This is analogous to how one would present an economic model in Section D.

Elsewhere in this submission to the review, I have tried to argue against being too prescriptive. However, I will acknowledge this is one area where the Guidelines could become more prescriptive. It would be very useful if the Guidelines could explicitly present the table or tables which need to be populated. In our submissions to PBAC over the years we have varied the five year time horizon by calendar year, by financial year or by predicted date of listing. Then, following PBAC recommendation this all changes again depending upon which government department uses it.

The spreadsheets provided by the department, whilst helpful, are probably overly prescriptive because they focus on the structure of the analysis. It is difficult to engineer different drugs with treatment modalities with different treatment duration of different diseases in to a specific model structure. However, what is possible is to reach a consistent output – total costs over 5 years (see Attachment for a potential example).

Finally

I hope these comments are informative and helpful as coming from an organisation which uses the Guidelines for the preparation of submissions to PBAC. I appreciate my comments do not always necessarily align completely with the objectives of the review and I apologise for the length of this submission. However, it seems this forum is as good as any to offer this perspective. I would be happy to discuss further.

Kind regards,

Dominic Tilden

Director
THEMA Consulting Pty Ltd
14th September 2015

Attachment: Example of Section E output table

	Calendar year of listing [a]	Full calendar year 1	Full calendar year 2	Full calendar year 3	Full calendar year 4	Full calendar year 5
Total expenditure on DRUG X by PBS/RPBS and patients						
Total expenditure on DRUG X to PBS/RPBS						
Substituted PBS/RPBS expenditure						
Net PBS/RPBS expenditure						
Net non-PBS health care expenditure						
Net health care expenditure						

[a] Predicted listing date to end of that calendar year