ARE N-OF-1 TRIALS AN ECONOMICALLY VIABLE OPTION TO IMPROVE ACCESS TO SELECTED HIGH COST MEDICATIONS? – THE AUSTRALIAN EXPERIENCE

Running title: Economic evaluation of N-of-1 trials

1. Paul A Scuffham (PhD, PG Dip, BA), Professor of Health Economics, School of Medicine, Griffith University

2. Michael J Yelland (MB BS, PhD, FRACGP, FAFMM, Grad Dip Musculoskeletal Med), Associate Professor of Primary Health Care, School of Medicine, Griffith University

3. Jane Nikles (MBBS), Senior Research Fellow, School of Medicine, The University of Queensland

4. Eva Pietrzak (PhD), Senior Research Officer, School of Population Health, The University of Queensland,

5. David Wilkinson (BSc(Hons), MBChB MSc MD PhD DSc; Fellowships FRACGP, FAFPHM, MRCP), Head & Professor of Primary Care, School of Medicine, The University of Queensland

Address for correspondence

Michael J Yelland (PhD), Associate Professor of Primary Health Care, School of Medicine, Logan campus, Griffith University, University Drive, Meadowbrook Qld, Australia m.yelland@griffith.edu.au

Ph +61-7-3382 1358 Fax +61-7-3382 1338

STATEMENT OF COMPETING INTERESTS

General Practice Evaluation Program and the Department of Health and Aged Care provided funding but had no involvement in any facet of the research. GlaxoSmithKline provided medication and placebos, and funding; participated in initial planning discussions; and viewed the manuscript prior to publication, but had no involvement in the collection, analysis or interpretation of data, writing the report or the decision to submit the final manuscript for publication. Michael Yelland has been paid by GlaxoSmithKline to present the results of aspects of this research at two conferences.

words: n-of-1 trials, cost-effectiveness, celecoxib, gabapentin, osteoarthritis, chronic neuropathic pain.

ABSTRACT

OBJECTIVES. To explore the economic viability of N-of-1 trials for improving access to selected high cost medications in Australia.

METHODS. Cost and effectiveness estimates were derived from two N-of-1 trials conducted by The University of Queensland from 2003-2005 - celecoxib vs sustained-release paracetamol for osteoarthritis in a general practice setting and gabapentin vs placebo for chronic neuropathic pain in a hospital setting. Effectiveness was determined by the proportion of responders to each medication. The costs of trials were offset against the savings generated by subsequent changes in prescribing. Decision analysis models with semi-Markov processes were used to compare different scenarios of N-of-1 trials versus usual care.

RESULTS. The fixed cost of performing N-of-1 trials was approximately $A23,000 for each trial and the variable cost was approximately A$1300 per participant. Clinical outcomes favoured celecoxib over paracetamol in 17% of participants and gabapentin over placebo in 25% of participants. Modelling these results showed that the cost-offsets from efficient use of medications were less than the cost of running a trial; however, the incremental costs per QALY gained were A$6896 and A$29,550 for the gabapentin/placebo and celecoxib/paracetamol trials respectively over a five-year horizon. Key factors affecting the viability were the time horizon modelled, the variable cost per participant, the probability of response to the intervention medication, and rates of use in non-responders and the usual care alternative.

CONCLUSIONS. The N-of-1 strategy offer a realistic and viable option for increasing access to selected high cost medications where the medications are used for the symptomatic treatment of chronic disease, have rapid onset of action, and clinical response is unpredictable without a trial.

INTRODUCTION

Over the last 10 years there has been increasing concern about the high psychosocial, economic and health costs of inappropriate medication use. In Australia, the Pharmaceutical Benefits Scheme (PBS) is the government scheme that subsidises common prescription medications. The growth of Government expenditure on subsidised medications listed on the PBS is of increasing concern. For the year ending 30 June 2004 PBS expenditure totalled $5 billion, a 9.3 per cent increase compared with the previous year [1]. In the same period, total PBS prescription volumes increased only by 4.3 per cent.

There is a need to control government expenditure while still making medications available to those in whom they will be effective. Consequently, recommendations for listing of medications on the PBS by the Pharmaceutical Benefits Advisory Committee (PBAC) are heavily based on high-level evidence from randomised controlled trials. Typically these trials define the mean response in a group of participants; however, the variation in response within these groups may be large and it may be impossible to predict response at the individual level. It is well known that drug trials can be performed on younger and / or healthier sufferers of a particular disease and there can be differences between the trial groups and the ultimate target groups of a given medication, e.g. [2,3]. Evidence from population-based trials may result in the long-term prescription of expensive medication to individuals for whom it is not effective or even harmful, leading to suboptimal cost-effectiveness and wastage of resources.

When applying economic data within a formal approvals process, such as that used by the PBAC, an overall estimate of the incremental cost per unit of health benefit gained is provided in the form of the incremental cost-effectiveness ratio – ICER, effectively averaging across responders and non-responders. In reality the cost to gain a unit of health benefit amongst responders will be very low while that among non-responders will be very high. What options then exist for more effective "targeting" of drug treatments, especially in the era of very high cost medications, which may offer marginal benefits to a large group (but very high benefit to small numbers of patients), and may also have significant side effects?

Current targeting strategies may be less than optimal as they rely on identifying characteristics of groups of patients who are likely to have a higher response rate. Non-responders are still included in this target group and potential responders may be excluded. For a few medications, genetic testing to identify potential responders is becoming important; however, this approach is currently limited. For some medications, such individualised targeting can only be obtained through N-of-1 trials. N-of-1 trials provide empirical data of individual responses to treatment. These are within-patient randomized, double-blind, cross-over comparisons of two treatment regimens (one of which may be placebo), in which patients act as their own controls [7]. The N-of-1 trial offers the highest level of evidence for the individual on the effectiveness of long-term medications used for symptomatic treatment of stable chronic conditions [8].

The use of N-of-1 trials in routine prescribing could potentially produce financial savings, depending upon five factors: the cost of the medication, the duration of use of the medication, the proportion of responders, the proportion of non-responders who continue medication and the cost of undertaking the trials. Potential financial savings to the PBS are likely to be greater where response rates are relatively low. If non-responders cease their medication, morbidity and mortality from medication-related adverse events, and costs to the health care system would be greatly reduced. Early identification of non-responders has the potential to reduce the cost per unit of effect gained in responders, improving the cost-effectiveness of the medication. It shifts resource allocation decisions to the level of the individual patient – to the margin [9].

A thoughtful paper by Karnon discusses the potential of N-of-1 trials in the estimation of individualized cost-effectiveness and uses hypothetical examples [9]. Larsen et al ran a N-of-1 trial service from 1991-1993 in the US [7]. Crude estimates of resource consumption including time-motion studies estimated that total costs per trial would likely be $400 to $500 in 1990 US dollars. They did not calculate savings that might result from the trials that led to discontinuation of chronic medications or the avoidance of side effects. Guyatt (1990), who also offered an N-of-1 trial service in the US, found that a substantial proportion of trials resulted in discontinuation of medication that would otherwise have been continued for months or years [10].

Until now there has been only one report of a formal economic evaluation of N-of-1 trials. This is a recent randomised controlled trial from Canada which showed that N–of-1 trials with diclofenac/misoprostol produced slightly better health outcomes (as measured by the Health Assessment Questionnaire for pain and disability, WOMAC scales, and physician global assessments), although were more time consuming and expensive over a relatively short time horizon (6 months), than standard treatment in osteoarthritis (OA)(cessation of NSAID with recommencement if symptoms worsened) [11]. This small study had some important limitations. The numbers were small - there were 25 patients in one group and 24 in the other and there was a high drop out rate from the N-of-1 trial group (18 out of 24 did not complete 3 crossovers). The higher costs for the N-of-1 group (from US$60 to $160 more per patient) were due to higher costs of medications, nurse and physician time, and travel costs.

Our study aimed to determine the economic viability of N-of-1 trials of medications for symptomatic treatment of stable chronic conditions in the Australian context, in both hospital and community settings.

METHODS

In 2003-5, the N-of-1 trial service at The University of Queensland conducted two N-of-1 trials. The first compared celecoxib with sustained release paracetamol (SR paracetamol) for osteoarthritis and was performed in general practices throughout Australia. The second compared gabapentin with placebo for chronic neuropathic pain and was performed in outpatient clinics in two Australian hospitals.

The duration of each trial was 12 weeks, comprising three 4-week cycles with a fortnight on each medication assigned in a random order (e.g. 2 weeks of celecoxib followed by 2 weeks of SR paracetamol). The choice of three pairs of treatment periods is regarded as the best compromise between statistical certainty and patient acceptability. With one pair, the chance of a false positive or false negative result is up to 50%, but with three pairs this reduces to 12.5%. More than three cycles makes the trials unacceptably long and expensive. Trials could be stopped early in cases of severe adverse reactions or if symptoms became intolerable.

Patients were randomised to their starting medication in blocks of four. Five variables were monitored in each cycle – pain, stiffness (for celecoxib/paracetamol trials) or sleep interference (for gabapentin/placebo trials), functional limitation, frequency of adverse events and preferred medication. Differential responses in pain, stiffness, sleep interference and functional limitation responses were determined by minimum clinically detectable differences [12-14] for adverse events by a lower frequency on one medication in at least two cycles and for medication preference by a preference for one medication in at least two cycles. The overall response status of each patient was then based on an equal weighting of the differential response for each variable [15,16]. Participants and their doctors were sent a comprehensive report of the results within two weeks of completion of the last cycle. This was used to inform, but not dictate, future medication usage. Ethics approval for both trials was obtained from The University of Queensland's Medical Research Ethics Committee, and additionally, for the gabapentin vs placebo trial, from the ethics committees of the participating institutions, Princess Alexandra Hospital, Brisbane and the Port Kembla Hospital. Due to the increased risk of sudden death from celecoxib, in late 2004 the Therapeutic Goods Administration instructed all trials using celecoxib to cease – including ours.

The economic evaluation of these trials is primarily a cost-minimisation analysis. This approach was chosen because medication and resource use were routinely measured during follow-up, but health outcome data were monitored during the trial period only; thus a full economic evaluation was not possible. Nevertheless, we address this shortcoming by undertaking cost-effectiveness and cost-utility analyses by extrapolating the pain scores for responders and non-responders from the trial for the time horizon modelled. The pain index was recorded on a visual analogue 0-10 scale with 0=no pain and 10=extreme pain; scores were categorised and converted to a utility weight to estimate quality-adjusted life years (QALYs). Although not ideal, this approach does give some relativity to the cost minimisation analysis by estimating the incremental cost per point decrease on the pain scale and the incremental cost per QALY gained.

Cost of N-of-1 trials

We estimated the costs of setting up and running a fully-funded university-based N-of-1 trials service based on our experience with the staff time and costs of running the N-of-1 trials funded from research grants from 2003 to 2005. Costs included a fixed component for trial establishment that is independent of the number of patients in a trial, and a variable component per patient for each N-of-1 trial. The fixed establishment costs were estimated at $23,280 per trial. These comprised staff costs for protocol development ($8,130), funding applications ($4,730), ethics applications and business agreements ($3,000), preparation of forms and questionnaires ($4,350), database development ($2,200) and design/preparation of medication packs ($870) The costs of N-of-1 trials per patient included the costs of recruitment, administration, data collection and analysis, generation and feedback of results and 12 months follow-up of health and economic outcomes. These totalled $610 for each patient for the celecoxib/paracetamol trial and $577 for each patient for the gabapentin/placebo trial.

Because the N-of-1 trials are an ongoing service, only a proportion of the total fixed costs can be assigned to those who have participated so far. Therefore, for the fixed costs, we have assumed the protocol will be valid for 200 participants in each trial, and calculated a fixed cost per patient of $116 based on this number. The fixed costs and patient trial costs were subject sequentially to the university infrastructure charge of 50%, a surcharge of 20% for internal infrastructure and then the statutory goods and services tax (GST) of 10%. In this scenario, the total cost per patient is set at $1373 for celecoxib/paracetamol trials and $1438 for gabapentin/placebo trials. These costs, and the number of participants, are later varied in a sensitivity analysis.

Modelling

Structure of the model

A decision analysis model with two arms: 'N-of-1 trial' and 'No trial' was constructed using TreeAge Pro Software [17]. Within the N-of-1 arm, there is a chance (probability) that the intervention medication (celecoxib and gabapentin) is better than the alternative medication. The alternative medication for each trial is the set of other medications used by participants, for example, analgesics (including paracetamol), opioids, and non-steroidal anti-inflammatory drugs. These alternatives are used for the economic evaluation, rather than placebo in the gabapentin

trial and paracetamol in the celecoxib trial, because participants used these medications for pain rescue during the trial and used these before and/or after the trial during follow-up. Within each arm of the model, there are five "health states" based on the observed pattern of medication use and the probability of survival. The health states are: intervention medication only, intervention plus other medications, other medications only, no medications, and dead. Using a semi-Markov process with monthly cycles, patients transition between medications or can die (at the Australian age-specific all-cause mortality rate [18]. This approach allows "leakage" of those shown to be non-responsive to the intervention medication to use the intervention medication, and *vice versa*. The "No trial" comparator follows the same pathway but without testing of which medication is best. The model used for the evaluation of both trials is summarised in Figure 1.

FIGURE 1 GOES ABOUT HERE

Assumptions used in the base case model

1. Response status is defined by an equal weighting of the differential response five outcomes: pain, functional limitation, adverse event frequency and preferred medication for both trials as described earlier in the methods. Responders are defined as those showing a superior response to celecoxib compared with SR paracetamol or gabapentin compared with placebo.
2. Those participants showing no difference between medications or a superior response to SR paracetamol or placebo are defined as non-responders and would continue on SR paracetamol or no additional medication respectively.
3. Response rates used in estimates are those from the series of completed trials conducted for this project (i.e. 17.1% for celecoxib and 24.0% for gabapentin)[15,16].
4. The use of the intervention medication by Responders is that observed in the 12 months of follow-up from the trials; the rates of use of the intervention medication only, intervention medication in combination with other medication, and other medication are constant.

5. Non-responders are allowed to continue use of celecoxib or gabapentin in the N-of-1 arm but, because participants are informed of their response outcome, they discontinue use of the intervention medication at a rate 50% annually.

6. Medication use in the no-trial arms are estimated from the trial data. For the celecoxib/paracetamol analysis, the pre-trial data on medication use was used; for the gabapentin analysis, we assumed the use of gabapentin was double that observed for non-responders over the 12-months follow-up. This assumption was made because pre-trial data indicated very low use of gabapentin. The expected use of gabapentin in the absence of a trial is much greater than reported for the pre-trial period as all patients are expected to use gabapentin for an unknown period (NB: patients were recruited into the N-of-1 trial when the clinician considered a trial of gabapentin). In the absence of a trial some would respond and others would persist with using it even in the absence of a true response.

7. An annual discontinuation rate of 25% of the intervention medication is used in the no-trial arm.

8. Compliance with medications is 100%.

9. The full dispensing price for all medications (as per the PBS Schedule) is used.

10. The pain scores used for responders, and consequently utility weights, are the mean scores reported by responders when using the intervention medication. For non-responders, the mean scores of intervention and comparator medication are used, and for the no-trial arm, the mean of the all scores (intervention and comparator medication for responders and non-responders) is used.

11. The pain scores were mapped to the EQ-5D by categorising scores as no pain or discomfort <= 1; moderate pain or discomfort <=6, and extreme pain or discomfort >6. To approximate a utility weight, the EQ-5D scoring algorithm for the UK was used [19] with the assumption that there were no problems with mobility, personal care, usual activities or anxiety/depression.

12. The pain score and utility weight are constant for the duration of the time horizon modelled.

Variables used in the models

The proportions using each medication are calculated from the N-of-1 trials using the means over 12-months follow-up. Upper and lower values were calculated from the 95% confidence intervals. Similarly, costs for each medication were calculated from observed medication usage over 12 months. This approach implicitly accounts for different daily doses. The mean costs, and 95% confidence intervals are used for the upper and lower values. The time horizon used in the base case model is five years and is varied in sensitivity analyses up to end of life. All costs (and QALY's) are discounted at 5% [20]. Variables used in the models are presented in Tables 1 and 2.

TABLES 1 & 2 GO ABOUT HERE

Sensitivity and scenario analyses

One-way sensitivity analyses were performed on all variables to identify the variables that have the greatest impact on the incremental costs and the incremental cost per QALY gained for both models, including assessing the effects of scale on cost-effectiveness. A scenario was developed to assess a policy where high cost pharmaceuticals require that an individual's response must be proven through an N-of-1 trial before the medication is approved for government subsidy to that individual. In addition, an analysis of responders compared with non-responders is presented. This analysis is equivalent to an evaluation of the efficacy and costs of an intervention medication compared with the control. Incremental costs and QALYs are estimated for the intervention medication.

RESULTS

Within the range of parameters chosen for the base-case models, the "N-of-1 trial" option costs significantly more per person than the alternative of "No trial" for both trials conducted. However, the cost of the trial was partially offset by more efficient use of medications through non-responders using less of the high cost medication. These cost-offsets were greater for the gabapentin/placebo trial (A$569) than for the celecoxib/paracetamol trial (A$221), resulting in incremental costs of A$869 and A$1152 for the two trials respectively (Table 3).

TABLE 3 GOES ABOUT HERE

The pain scores, measured during the trials only, show that the average level of pain for participants in N-of-1 trials was reduced by 0.278 and 0.113 points for the gabapentin/placebo and celecoxib/paracetamol trials compared with the "No trial" alternative. Thus, an N-of-1 trial can reduce pain through efficient medication use. Although these reductions in pain score may appear small, they contain the scores from responders and non-responders. Converting the pain scores into utility weights, and then to QALYs, the N-of-1 trials lead to an additional 0.126 QALYs for the gabapentin/placebo trial and 0.039 QALYs for the celecoxib/paracetamol trial over five years compared with no trial.

The incremental cost per one point reduction on the 10-point pain scale is A$3125 and $10,199 for the gabapentin/placebo and celecoxib/paracetamol trials respectively, and the incremental cost per QALY gained is A$6896 and A$29,550 respectively.

Sensitivity analyses

Varying the time horizon modelled shows that greater cost-offsets are achieved with longer horizons (i.e. the incremental cost is smaller) for both trials; similarly, the QALYs gained also increase with increased time horizons. Using a horizon to the end of life, the ICERs were A$1725 and A$10,278 per QALY gained.

Tornado diagrams are used to depict the change in expected costs from changing individual parameter values within their expected range (Tables 1 and 2); the factor that has the greatest effect is depicted at the top of the diagram down to those with the smallest effect on results. Tornado diagrams are presented for the incremental costs for each trial (Figures 2 and 3) and the incremental cost per QALY gained (Figures 4 and 5) using the 5-year time horizon.

FIGURES 2, 3, 4 AND 5 GO ABOUT HERE

For both N-of-1 trials, the rates of use of the intervention medication in the no-trial alternative and the non-responders were highly important factors affecting the incremental costs and ICERs. That is, the probability of use of the intervention medication by non-responders and without a trial makes a substantial difference to the costs.

The variable costs to conduct N-of-1 trials were a key driver of the costs and ICERs in all four tornado diagrams. Assumption of economies of scale were obtained, the variable costs could be reduced by approximately 50% for the celecoxib/paracetamol trial and 40% for the gabapentin/placebo trial; this would result in incremental costs and ICERs that are 47% and 44% lower compared with the base case. The costs of the intervention medications had relatively little effect on the incremental costs or ICERs; more important was the cost of the alternative medications used – both in combination with the intervention medication and without the intervention medication. When lower "other" medication costs were used, the incremental costs and the ICERs increased as the difference in costs between running a N-of-1 trial and the alternative were greater. As such, the benefits of N-of-1 trials were reduced.

Changes in the probabilities of responding to celecoxib in the celecoxib/paracetamol trial had little effect on the incremental cost but a large effect on the ICER; this large effect on the ICER occurred when the additional benefit of celecoxib compared with other medications approached zero, resulting in a very large ICER. In contrast, the probability of responding to gabapentin was a key factor affecting the incremental costs with a relatively smaller effect on the ICER. Because gabapentin is required to be prescribed by specialists, lower response rates tend to reflect decreased usage by non-responders and hence, lower incremental costs. Variation in the utility weights used and the pain scores had substantial effects on ICERs. Changes to other factors, such as the initial age and discount rates, had relatively little effect on results.

Overall, from examining both tornado diagrams, the important features of N-of-1 trials affecting the ICER were the assumptions around the no-trial alternative (including the use of the intervention drug and other drugs, and the discontinuation rate). Changes in health outcomes, including the utility weights, were key factors, as was the probability the intervention drug was better than the comparator. Features of N-of-1 trials that were less important were the fixed costs, discount rates and the mix of drugs used by responders and non-responders.

A scenario where prescribing of the high cost intervention medication is restricted to those who have shown a positive response in a N-of-1 trial would reduce the costs of the N-of-1 trial only; that is, the pain scores and QALYs would remain unchanged as the use of the medication in a non-responder has no effect on these outcomes (Table 4). Thus, this restriction has some benefit to controlling pharmaceutical budgets and better use of medications.

The analysis of responders compared with non-responders within the N-of-1 trial is used to identify the incremental cost and cost-effectiveness of the intervention medication compared with the alternative medication (opposed to whether an N-of-1 trial provides value for money compared with no trial). The difference in medication costs for responders vs non-responders in the celecoxib vs SR paracetamol trial were relatively small, whereas the difference in costs between responders and non-responders in the gabapentin vs placebo trial were substantially greater (Table 5). In contrast, responders in the celecoxib trial had a slightly greater reduction in pain score compared with responders in the gabapentin trial. This difference in pain reduction translated into greater difference in QALYs between responders and non-responders in the two trials. From this responder analysis for gabapentin, an ICER of A$8764 per point reduced on the pain scale and A$28,439 per QALY gained were obtained. For the celecoxib trial, ICERs were A$389 and A$522 per point of pain reduced and QALY gained respectively.

TABLE 5 GOES HERE

DISCUSSION

This study provides important evidence that within the Australian context N-of-1 trials offer a realistic option for increasing patients' access to high-cost medications, under specific circumstances. In particular, the characteristics of the gabapentin trial appeared more economically viable than the celecoxib trial. Gabapentin is relatively more expensive than celecoxib, and reducing its use in non-responders results in greater cost-offsets compared with the cost-offsets obtained from reducing use of celecoxib in non-responders. In addition, the proportion of responders in the gabapentin trial were greater than those in the celecoxib trial, and the reduction in pain score by gabapentin responders was greater than the pain reduction by celecoxib responders. Therefore, the gabapentin trial resulted in substantially greater QALYs than those obtained from the celecoxib trial.

Our analysis of the costs of developing, setting up and running N-of-1 trials showed a relatively high cost per individual trial. However, the modelling suggests that these costs may partially recouped if the cost differential between the two therapies is sufficiently high, as illustrated by the gabapentin/placebo trials. In contrast, when the

cost-differential is smaller, as in the celecoxib/SR-paracetamol trials, it was cheaper to allow every patient to use celecoxib. This concurs with the findings of Pope et al. [11]. The additional costs of the N-of-1 trials reported here were A$869 and A$1152 for the gabapentin/placebo and celecoxib/paracetamol trials. Moreover, N-of-1 trials improve clinical and health outcomes. We estimated modest QALY gains of 3.8% and 1.3% for two trials, which when combined with costs, produced incremental cost-effectiveness ratios of A$6896 and A$29,550 per QALY gained for the two trials. However, over and above any costs or cost-offsets, there is value to both clinicians and patients in having knowledge that they are being prescribed medications to which they will respond positively, and thus, these analyses may understate the full economic benefit of N-of-1 trials.

Some of the assumptions chosen for the base case models were simplifications and may be unlikely to be observed in the real world. For example, persistence with long-term therapy is generally poor (approximately 50%), particularly after the first six months [21]. However, persistence is expected to be greater when the patient is given the knowledge that the medication is effective for them, and is expected to be much lower when the patients are told the medication does not work for them. A reduction in the number of responders persisting with the more expensive medication (or the duration of use by responders) would have the effect of reducing the ICERs (assuming quality of life was maintained). Conversely a 'leakage' effect of non-responders resuming the more expensive medication will reduce cost-effectiveness. The long-term rates of the opposing effects of persistence and leakage are unknown; however, we have attempted to capture these effects in our model using data reported from 12 months of follow-up.

Additional cost-offsets not included in our analysis may come from a reduction in physician time for medication review and treatment of adverse reactions for those in the N-of-1 trial; however, the size of this effect is unknown. Another limitation of the analysis is the need to extrapolate health outcomes from the 12 week trial for the modelled time horizon. In future, we recommend that follow-up data collection should include the clinical measures used to identify responders and the use of a utility instrument such as the EQ-5D. Similarly, the "No trial" group modelled in the analysis was, effectively, a hypothetical cohort. Ideally, testing the effectiveness and cost-effectiveness of N-of-1 trials requires participants to be randomised to standard care or an N-of-1 trial.

The method of patient selection for these trials may have affected response rates and therefore estimates of cost-effectiveness. Requests for the celecoxib/paracetamol trials were mostly patient-generated and therefore attracted people who were uncertain about the effectiveness of their medication. The gabapentin/placebo trials were initiated by doctors who were uncertain of gabapentin's effectiveness in patients on long-term gabapentin, who had shown some response to a two to three week trial of gabapentin.

The minimum and maximum doses used in the gabapentin trial were below those recommended for the treatment of neuropathic pain (900-3,000mg/day) [22]. Both of these factors could have reduced the number of responders. A lower number of responders results in lower total costs in the N-of-1 trial option. In turn, this requires a shorter duration of continuous treatment for the trial arm to become more cost-effective than the usual care arm.

The responder analysis, comparing the intervention medication with the set of "other medications", is a spin-off benefit of N-of-1 trials. The responder analysis can allow efficacy, costs and cost-effectiveness of an intervention medication to be established and/or confirmed in addition to identifying the suitability of a medication for an individual. It is possible that non-random biases may occur in the N-of-1 trial process, despite double blinding, if patients know that the results will influence their treatment decisions. For example, patients may exaggerate benefits in the periods when they believe they are taking the more expensive medication [9].

Many of the practical issues of conducting post and telephone N-of-1 trials throughout Australia with minimal demands on prescribing doctors have been addressed by the N-of-1 trial service. However, the trials were performed in relatively small numbers of consenting, motivated patients. The provision of trials in a timely fashion to large numbers of community or hospital patients would present considerable logistical challenges, but would not be impossible. One option would be an extension of the current authority-prescribing system, where doctors request approval for specific subsidised medications under a special authority. Funding of the extension of such a service could come from those who would benefit, namely government, pharmaceutical companies and patients. Government stands to gain by making savings to the PBS budget. Pharmaceutical companies could benefit if the N-of-1 trials are a condition of listing for previously unlisted medications or if they are offered a premium price for

medications in responders. (That is, pharmaceutical companies could show a greater benefit by reporting the effect in responders compared with an alternative treatment. The price of the pharmaceutical could be increased to the point where the ICER was bordering on the acceptable threshold for government subsidy/funding, and thus, a higher price could be obtained with the restriction that the medication is prescribed for responders only.) Patients may benefit from clinical decisions based on individualised evidence; however, there may be strong resistance to pay for a test that may exclude them from taking a medication they think they need.

We analysed a scenario where mandatory participation in an N-of-1 trial is a condition for getting affordable access to a desired medication and showed greater cost-offsets could be achieved; however, mandatory participation in an N-of-1 trial might also be considered a significant ethical or moral issue. An appeals mechanism might be necessary for those who refuse to consent to an N-of-1 trial, and for non-responders disaffected by the denial of affordable access to medication. Furthermore, N-of-1 trials are double-blind studies for which informed consent is a standard requirement. Ethics guidelines state that refusal to consent to entry in a trial should never compromise the doctor-patient relationship [23]. However, clinicians commonly conduct informal (and methodologically inadequate) N-of-1 trials when they try a medication in a patient and judge a clinical response; and this occurs in the face of the issues outlined above.

Several questions remain unanswered about the broader application of N-of-1 trials. How can they be administered successfully in patients with cognitive impairment or for whom English is a second language? Should there be an item number or a Practice Incentive Payment for the effort required by doctors requesting them? Will they be accepted by doctors as a tool for rational prescribing or just perceived as another barrier to prescribing? Will they be accepted by patients who stand to lose affordable access to a medication that they may believe is helping them? A marketing campaign would be required to explain the benefits of N-of-1 trials as a tool for tailoring treatment to individual patients so these trials are not just seen as a method of cutting costs for the PBS.

Collectively these issues restrict the range of medications suitable for an N-of-1 trial prescribing strategy (Table 6). However, suitable medications would include expensive medications for the symptomatic treatment of chronic

conditions where there is a low overall response rate but a high and unpredictable variability in individual response. The difference in costs for responders and non-responders is a key factor in being able to make cost-savings from conducting N-of-1 trials. If there is an alternative medication for non-responders, it needs to be relatively cheap. N-of-1 trials may also have a role in targeting the use of medications which have a predictably superior clinical response to alternatives but with an unpredictably worse adverse event profile.

ACKNOWLEDGEMENTS

Thanks to the General Practice Evaluation Program and the Department of Health and Aged Care for funding; GlaxoSmithKline Consumer Healthcare for supply of medication, placebos, and funding and for their invaluable support; Prof Paul Glasziou, Prof Christopher Del Mar, A/Prof Geoffrey Mitchell for advice on design and conduct of the trials the N-of-1 trial service staff Norma McNairn, Grace McBride, Robyn Brown and, Donna-Marie Preston, for estimates of costs and collecting the data; Dr Peter Pillans and Jo Sturtevant from Princess Alexandra Hospital and Dr Chris Poulos from Port Kembla Hospital for estimates of costs and assistance with trials; Rachel Ngui for assistance with data analysis; and the doctors and patients who took part in the trials. In addition, the authors are grateful to the four anonymous referees who provided very constructive comments which helped shape this article.

REFERENCES

1. Summary of PBS processing Year Ending 30 June 2004. In. Health Insurance Commission Pharmaceutical Benefits Schedule Item Statistics December 22 2004. http://www.hic.gov.au/statistics/dyn_pbs/forms/pbs_tab1.shtml. Accessed 20/1/05.

2. Cameron HJ, Williams BO. Clinical trials in the elderly. Should we do more? Drugs Aging. 1996;9:307-10.

3. Rothwell PM. External validity of randomised controlled trials: "to whom do the results of this trial apply?" Lancet. 2005;365:82-93.

4. Sharpe N. Clinical trials and the real world: selection bias and generalisability of trial results. Cardiovasc Drugs Ther. 2002;16:75-7.

5. The Oxford League Table of Analgesic Efficacy. Bandolier website 2006. http://www.jr2.ox.ac.uk/bandolier/booth/painpag/Acutrev/Analgesics/lftab.html.. Accessed 12 April 2006.

6. Sculpher, M. The cost-effectiveness of preference-based treatment allocation: the case of hysterectomy versus endometrial resection in the treatment of menorrhagia. Health Econ. 1998; 7: 129-42.

7. Larson EB and Ellsworth AJ. An evidence based approach to individualising treatment. JAMA 1993;70: 2708-2712.

8. Guyatt GH, Haynes RB, Jaeschke RZ, et al. Users' Guides to the Medical Literature: XXV. Evidence-based medicine: principles for applying the Users' Guides to patient care. Evidence-Based Medicine Working Group. JAMA. 2000 Sep 13;284(10):1290-6.

9. Karnon J, Qizilbash N. Economic evaluation alongside N-of-1 trials: getting closer to the margin. Health Econ 2001;10:79-82.

10. Guyatt GH, Keller JL, Jaeschke R, et al. The n-of-1 randomized controlled trial: clinical usefulness. Our three-year experience. Ann Intern Med. 1990 Feb 15;112(4):293-9.

11. Pope JE, Prashker M, Anderson J. The efficacy and cost effectiveness of N-of-1 studies with diclofenac compared to standard treatment with nonsteroidal antiinflammatory drugs in osteoarthritis. J Rheumatol 2004;31:140-9.

.

12. Ehrich EW, Davies GM, Watson DJ, Bolognese JA, Seidenberg BC, Bellamy N. Minimal perceptible clinical improvement with the Western Ontario and McMaster Universities osteoarthritis index questionnaire and global assessments in patients with osteoarthritis. J Rheumatol 2000; 27:2635-2641.

13. Zisapel N, Nir T. Determination of the minimal clinically significant difference on a patient visual analog sleep quality scale. J Sleep Res 2003;12(4):291-8.

14. Chatman AB, Hyams SP, Neel JM, et al. The Patient-Specific Functional Scale: measurement properties in patients with knee dysfunction. Phys Ther 1997; 77:820-829.

15. Yelland MJ, Nikles CJ, McNairn N, et al. Celecoxib compared to sustained-release paracetamol for osteoarthritis: a series of n-of-1 trials. Rheumatology 2006 Jun 15; Epub ahead of print.
)

16. Yelland MJ, Nikles CJ, McNairn N, et al. Single patient trials of gabapentin for chronic neuropathic pain. In: IASP World Pain Congress; 2005; Sydney; 2005.

17. TreeAge Pro Software, Inc. In; 2004. 2005 Suite Release 1.5, Williamstown, MA. 2005.

18. Australian Institute of Health and Welfare (AIHW). Grim (General Record of Incidence of Mortality) Books. Canberra: AIHW, 2005])

19. Dolan P, Gudex C, Kind P, Williams A. The Time Trade-Off method: Results from a general population survey. Health Economics 1996;5:141-154)

20. Commonwealth Department of Health and Ageing. Guidelines for the Pharmaceutical Industry on Preparation of Submissions to the Pharmaceutical Benefits Advisory Committee. Canberra: Publications Production Unit, 2002]

21. Osterberg L, Blaschke T. Adherence to medication. N Engl J Med 2005;353:487-97.

22. MIMS Annual, 2005 v 29. Crows Nest NSW Australia.

23. Declaration of Helsinki. Brit Med J 1964;313:1448-1449.

FIGURE LEGENDS

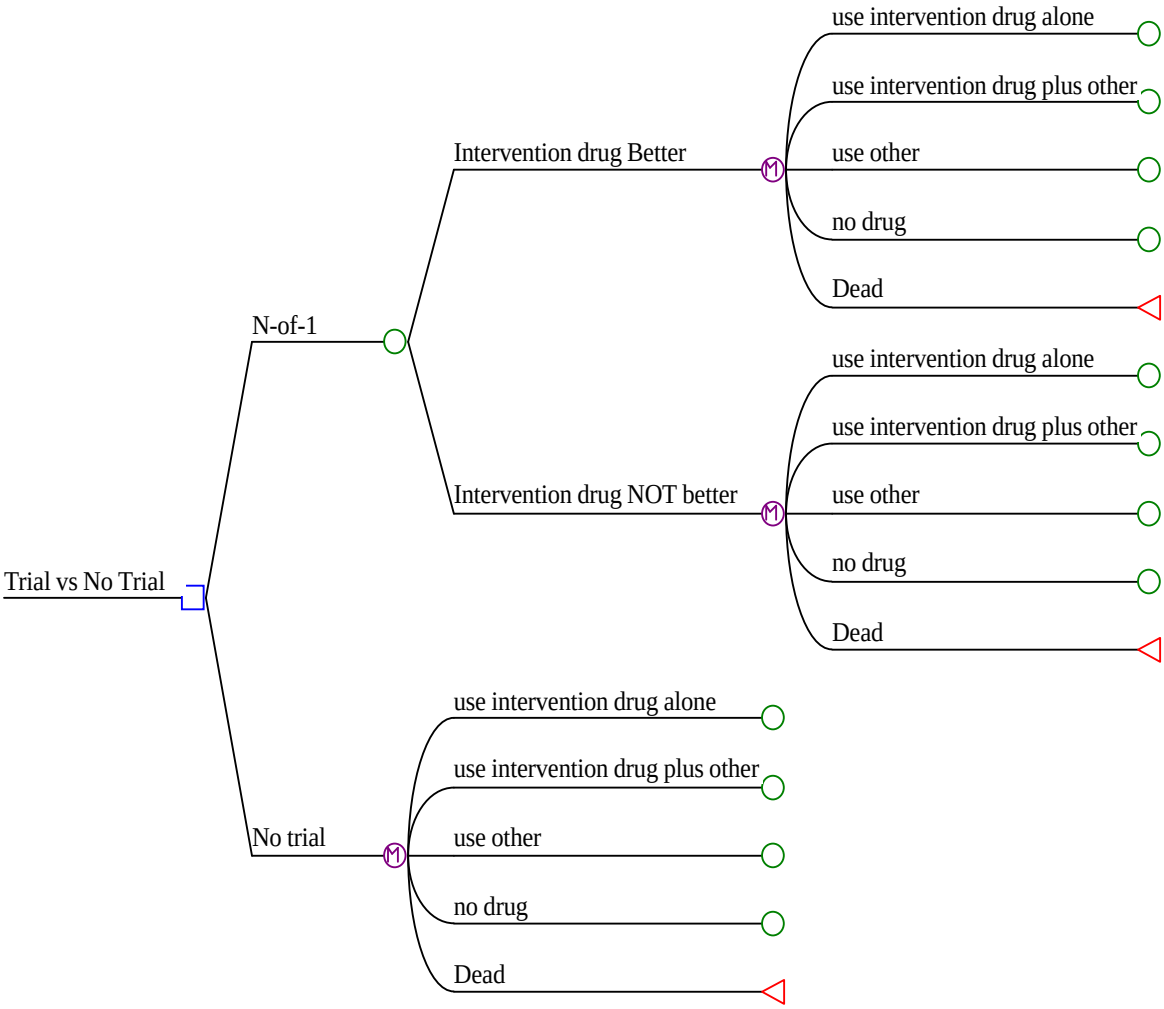
Figure 1. Illustration of the Decision Model with medication use/health states

Figure 2. Sensitivity analysis around incremental costs for the celecoxib/paracetamol model. The baseline incremental cost of the 'N-of-1 trial' compared with 'No trial' was $1,152.

Figure 3. Sensitivity analysis around incremental costs for the gabapentin/placebo model. The baseline incremental cost of the 'N-of-1 trial' compared with the 'No trial' was $869.

Figure 4. Sensitivity analysis around incremental cost per QALY gained for the celecoxib/paracetamol model. The baseline ICER of the 'N-of-1 trial' compared with "No trial" was $29,216.

Figure 5. Sensitivity analysis around incremental cost per QALY gained for the gabapentin/placebo model. The baseline ICER of the 'N-of-1 trial' compared with the "No trial" was $6,896.


Trial vs No Trial
N-of-1
Intervention drug Better
use intervention drug alone
use intervention drug plus other
use other
no drug
Dead
Intervention drug NOT better
use intervention drug alone
use intervention drug plus other
use other
no drug
Dead
No trial
use intervention drug alone
use intervention drug plus other
use other
no drug
Dead

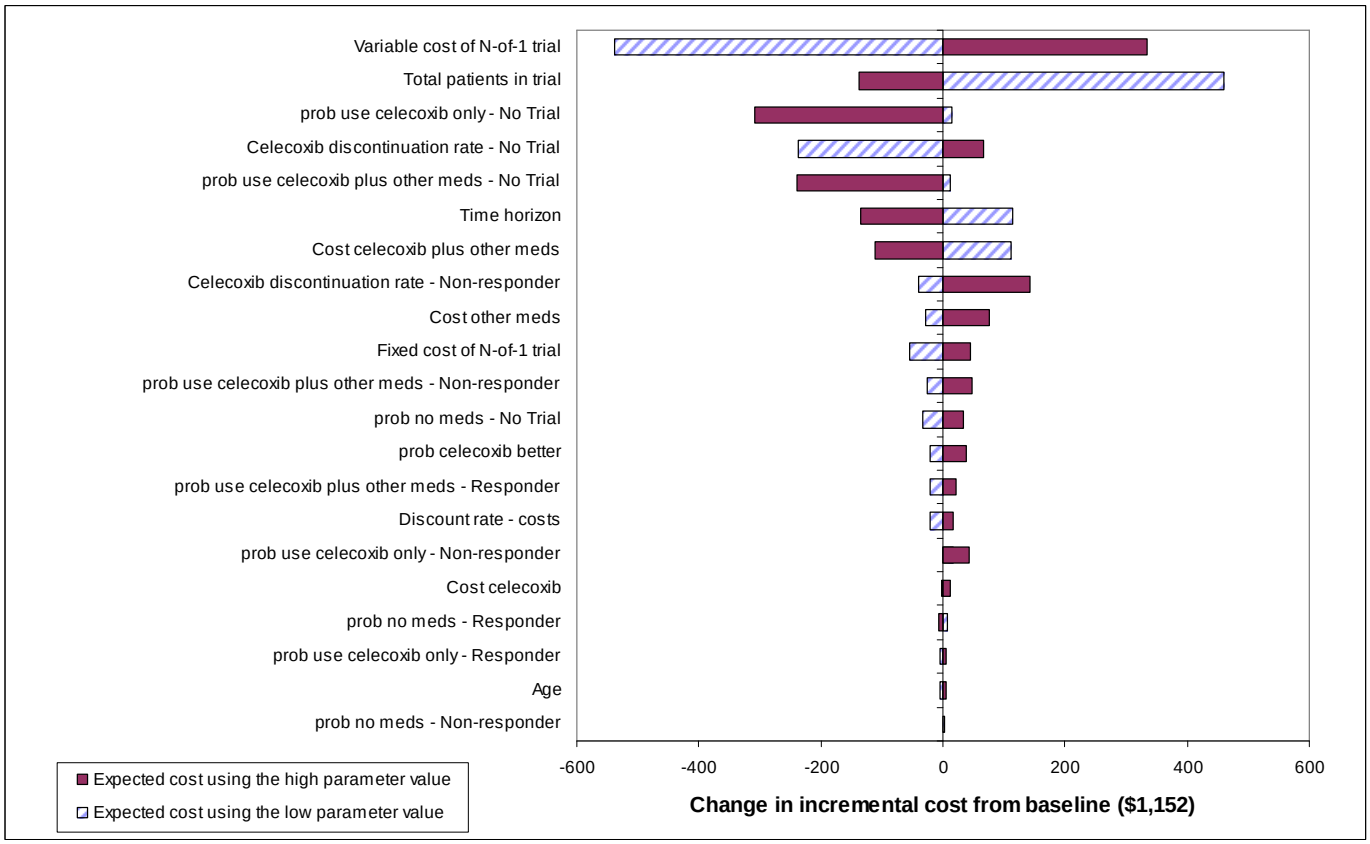

Variable cost of N-of-1 trial
Total patients in trial
prob use celecoxib only - No Trial
Celecoxib discontinuation rate - No Trial
prob use celecoxib plus other meds - No Trial
Time horizon
Cost celecoxib plus other meds
Celecoxib discontinuation rate - Non-responder
Cost other meds
Fixed cost of N-of-1 trial
prob use celecoxib plus other meds - Non-responder
prob no meds - No Trial
prob celecoxib better
prob use celecoxib plus other meds - Responder
Discount rate - costs
prob use celecoxib only - Non-responder
Cost celecoxib
prob no meds - Responder
prob use celecoxib only - Responder
Age
prob no meds - Non-responder
-600
-400
-200
0
200
400
600
Change in incremental cost from baseline ($1,152)
Expected cost using the high parameter value
Expected cost using the low parameter value

Figure 2.

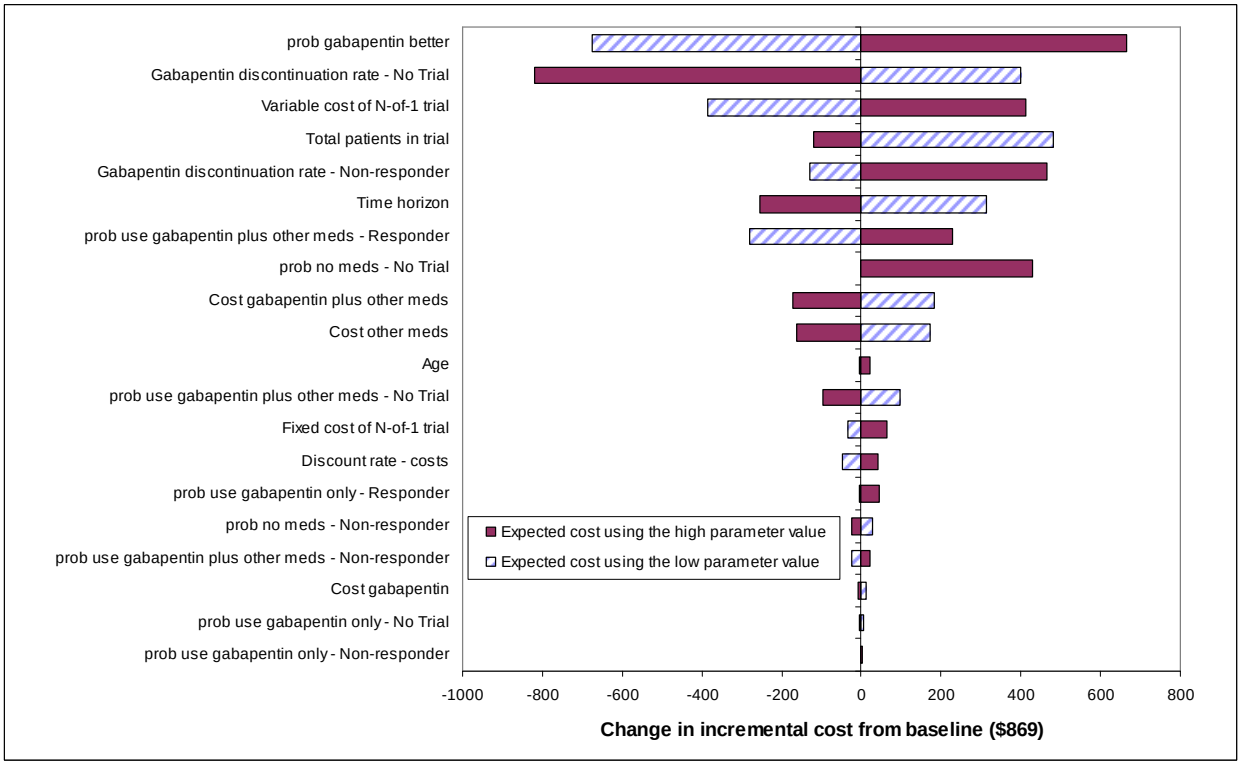

prob gabapentin better
Gabapentin discontinuation rate - No Trial
Variable cost of N-of-1 trial
Total patients in trial
Gabapentin discontinuation rate - Non-responder
Time horizon
prob use gabapentin plus other meds - Responder
prob no meds - No Trial
Cost gabapentin plus other meds
Cost other meds
Age
prob use gabapentin plus other meds - No Trial
Fixed cost of N-of-1 trial
Discount rate - costs
prob use gabapentin only - Responder
prob no meds - Non-responder
prob use gabapentin plus other meds - Non-responder
Cost gabapentin
prob use gabapentin only - No Trial
prob use gabapentin only - Non-responder
Expected cost using the high parameter value
Expected cost using the low parameter value
-1000
-800
-600
-400
-200
0
200
400
600
800
Change in incremental cost from baseline ($869)

Figure 3.

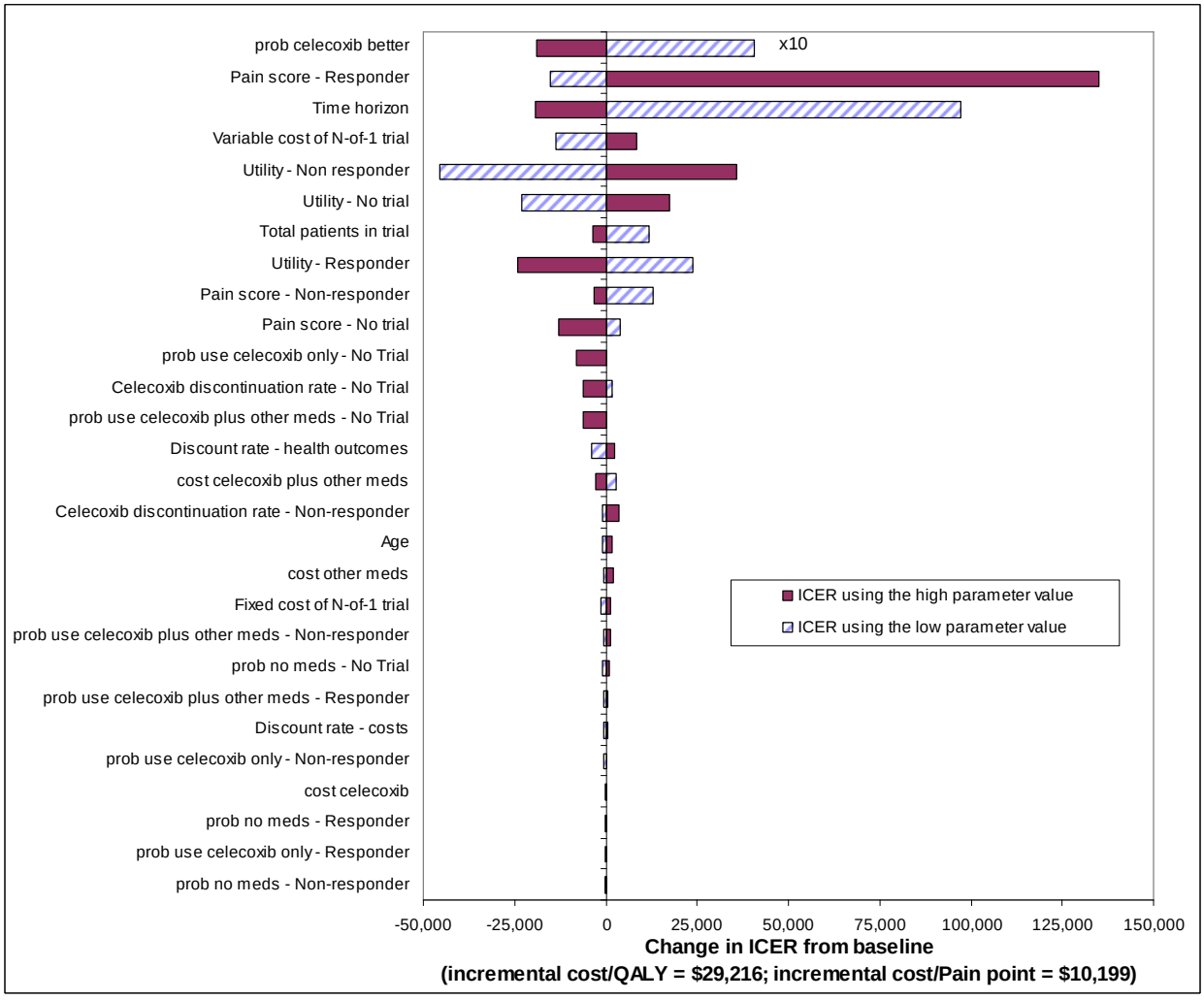

prob celecoxib better
Pain score - Responder
Time horizon
Variable cost of N-of-1 trial
Utility - Non responder
Utility - No trial
Total patients in trial
Utility - Responder
Pain score - Non-responder
Pain score - No trial
prob use celecoxib only - No Trial
Celecoxib discontinuation rate - No Trial
prob use celecoxib plus other meds - No Trial
Discount rate - health outcomes
cost celecoxib plus other meds
Celecoxib discontinuation rate - Non-responder
Age
cost other meds
Fixed cost of N-of-1 trial
prob use celecoxib plus other meds - Non-responder
prob no meds - No Trial
prob use celecoxib plus other meds - Responder
Discount rate - costs
prob use celecoxib only - Non-responder
cost celecoxib
prob no meds - Responder
prob use celecoxib only - Responder
prob no meds - Non-responder
x10
ICER using the high parameter value
ICER using the low parameter value
-50,000
-25,000
0
25,000
50,000
75,000
100,000
125,000
150,000
Change in ICER from baseline
(incremental cost/QALY = $29,216; incremental cost/Pain point = $10,199)

Figure 4.

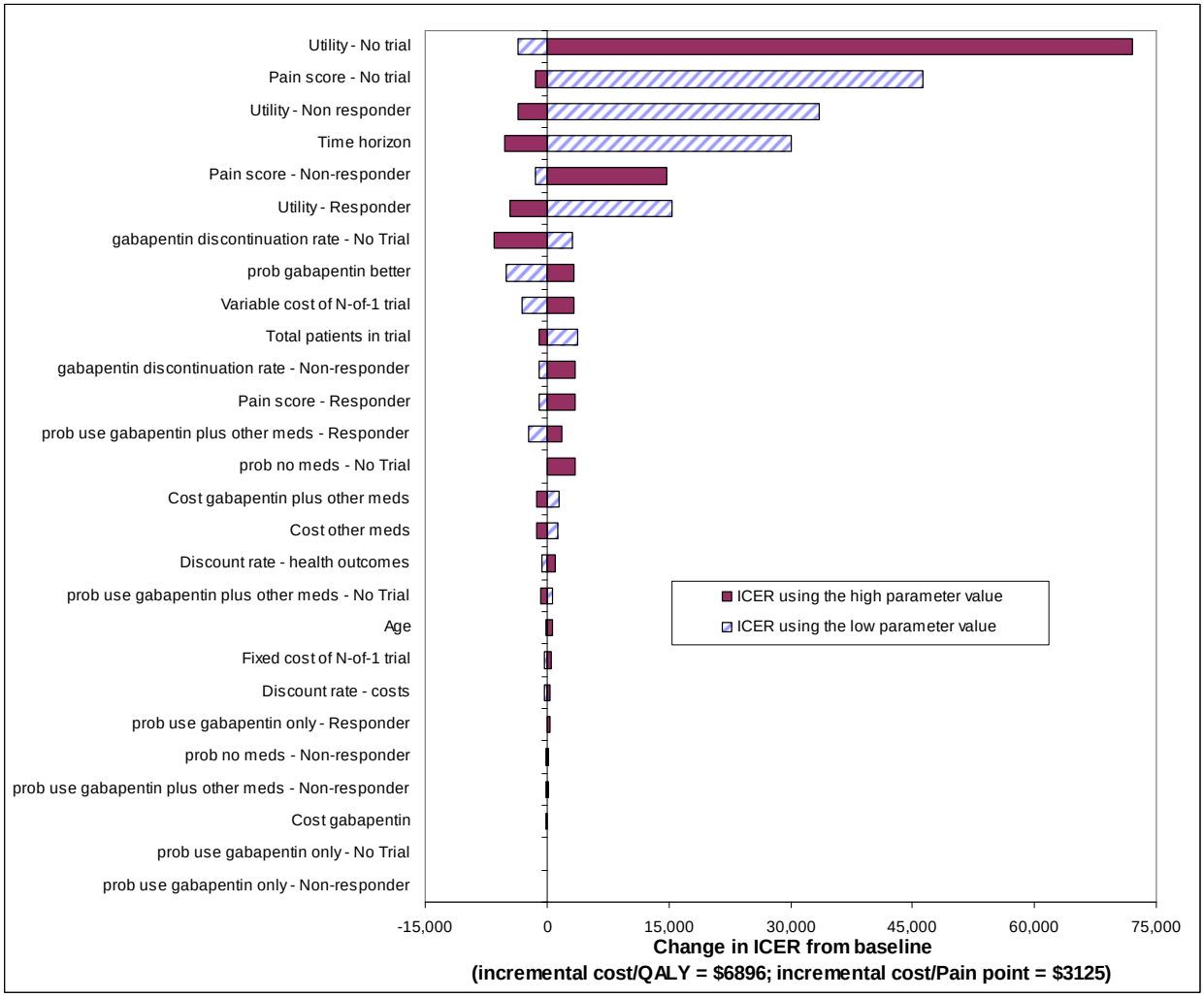

Utility - No trial
Pain score - No trial
Utility - Non responder
Time horizon
Pain score - Non-responder
Utility - Responder
gabapentin discontinuation rate - No Trial
prob gabapentin better
Variable cost of N-of-1 trial
Total patients in trial
gabapentin discontinuation rate - Non-responder
Pain score - Responder
prob use gabapentin plus other meds - Responder
prob no meds - No Trial
Cost gabapentin plus other meds
Cost other meds
Discount rate - health outcomes
prob use gabapentin plus other meds - No Trial
Age
Fixed cost of N-of-1 trial
Discount rate - costs
prob use gabapentin only - Responder
prob no meds - Non-responder
prob use gabapentin plus other meds - Non-responder
Cost gabapentin
prob use gabapentin only - No Trial
prob use gabapentin only - Non-responder
ICER using the high parameter value
ICER using the low parameter value
-15,000
0
15,000
30,000
45,000
60,000
75,000
Change in ICER from baseline
(incremental cost/QALY = $6896; incremental cost/Pain point = $3125)

Figure 5.

Table 1. Variables used in the celecoxib/paracetamol model

Variable	Base case value	Lower value	Upper value	Notes
Cost of trial:				
Total fixed costs	$23,200	$11,800	$39,600	Based on actual staff
Variable cost per patient	$1,257	$628	$1,500	inputs and
N participants	200	50	2000	infrastructure costs
Celecoxib responders				
Proportion Celecoxib better	0.171	0.085	0.320	Celecoxib/paracetamol
Use celecoxib only	0.194	0.158	0.231	N-of-1 trial; lower and
Use celecoxib and other meds	0.278	0.167	0.388	upper 95% CI; mean
Use other meds only	0.333	0.278	0.389	at 12 months f/u
No meds	0.194	0.146	0.243	
Celecoxib Non-responders				
Use celecoxib only	0.083	0.000	0.093	Celecoxib/paracetamol
Use celecoxib and other meds	0.185	0.000	0.206	N-of-1 trial; lower and
Use other meds only	0.667	0.533	0.800	upper 95% CI; mean
No meds	0.065	0.064	0.067	at 12 months f/u
Celecoxib discontinuation rate (annual)	0.500	0.000	1.000	
No Trial				
Use celecoxib only	0.089	0.000	0.691	Celecoxib/paracetamol
Use celecoxib and other meds	0.375	0.309	0.691	N-of-1 trial; lower and
Use other meds only	0.500	0.000	0.750	upper 95% CI; pre-
No meds	0.036	0.000	0.072	trial useage
Celecoxib discontinuation rate (annual)	0.250	0.000	0.500	
Other factors				
Cost of celecoxib	$34.74	$9.89	$59.59	Calculated from
Cost of other meds	$17.90	$8.32	$44.12	observed daily doses
Cost of celecoxib plus other meds	$39.86	$20.90	$58.82	over 12 months f/u, using PBS costs; lower and upper 95% CI
Pain score celecoxib responder	3.89	3.19	4.59	Mean score of responders for periods of gabapentin use during trial; lower and upper 95% CI
Pain score celecoxib Non-responder	4.55	4.20	4.89	Mean score of non-responders during trial; lower and upper 95% CI
Pain score No trial	4.55	4.25	4.85	Mean from all in trial; lower and upper 95% CI
Utility - celecoxib responder	0.78	0.47	1.00	Calculated from
Utility – celecoxib Non-responder	0.67	0.64	0.71	proportion with no,
Utility – No trial	0.68	0.65	0.71	moderate, and extreme pain mapped to the EQ-5D

Probability of death	age-specific mortality rates			[18]
Start age (years)	64.2	54.2	74.2	Observed in trial; lower and upper 95% CI
Duration (months)	60	12	End of life	
Discount rate (annual)	5.00%	0.00%	10.00%	[20]

Table 2. Variables used in the gabapentin/placebo model

Variable	Base case value	Lower value	Upper value	Notes
Cost of trial:				
Total fixed costs	$23,200	$11,800	$39,600	Based on actual staff
Variable cost per patient	$1,322	$802	$1,600	inputs and infrastructure
N participants	200	50	2000	costs
Gabapentin responders (proportions)				
Gabapentin responders	0.244	0.083	0.392	Gabapentin/placebo N-of-1 trial; lower and upper 95% CI
Use gabapentin only	0.227	0.154	0.300	
Use gabapentin and other meds	0.603	0.406	0.799	
Use other meds only	0.170	0.161	0.180	
No meds	0.000	0.000	0.000	
Gabapentin Non-responders (proportions)				
Use gabapentin only	0.071	0.068	0.075	Gabapentin/placebo N-of-1 trial; lower and upper 95% CI
Use gabapentin and other meds	0.146	0.132	0.159	
Use other meds only	0.614	0.550	0.679	
No meds	0.169	0.161	0.176	
Gabapentin discontinuation rate (annual)	0.500	0.000	1.000	Assumed
No Trial (proportions)				
Use gabapentin only	0.143	0.136	0.149	Assumed 100% greater than non-responder
Use gabapentin and other meds	0.291	0.263	0.319	Assumed 100% greater than non-responder
Use other meds only	0.566	0.601	0.435	Residual
No meds	0.000	0.000	0.097	Min from responders; high value = observed pre-trial
Gabapentin discontinuation rate (annual)	0.250	0.000	0.500	
Other factors				
Cost of gabapentin	$113.78	$51.80	$155.41	Calculated from
Cost of other meds	$83.56	$53.08	$112.55	observed daily doses
Cost of gabapentin plus other meds	$197.34	$104.88	$267.96	over 12 months f/u, using PBS costs; lower and upper 95% CI
Pain score gabapentin responder	3.64	3.03	4.24	Mean score of responders for periods of gabapentin use during trial; lower and upper 95% CI
Pain score gabapentin Non-responder	4.13	3.83	4.43	Mean score of non-responders during trial; lower and upper 95% CI
Pain score No trial	4.29	4.03	4.54	Mean from all in trial; lower and upper 95% CI
Utility – gabapentin responder	0.74	0.53	0.96	Calculated from
Utility – gabapentin Non-responder	0.71	0.68	0.75	proportion with no,
Utility – No trial	0.69	0.66	0.72	moderate, and extreme pain mapped to the EQ-5D

Probability of death	Age-specific mortality rates			[18]
Start age (years)	57.6	41.8	73.4	observed in trial; lower and upper 95% CI
Duration (months)	60	12	end of life	
Discount rate (annual)	5.00%	0.00%	10.00%	[20]

Table 3. Results for the N-of-1 trial prescribing strategies: Costs, outcomes and cost-effectiveness of the N-of-1 Trials vs No Trial

	Gabapentin vs Placebo	Celecoxib vs Paracetamol
5 years (Base case)		
Cost: No trial (A$)	5,654	1,193
Cost: N-of-1 (A$)	6,523	2,346
Incremental Cost (A$)	869	1,152
Pain score: No trial	4.290	4.550
Pain score: N-of-1	4.012	4.437
Reduction in pain (points)	0.278	0.113
QALYs: No trial	3.141	3.045
QALYs: N-of-1	3.267	3.084
Incremental QALYs	0.126	0.039
Incremental cost per point reduction in pain (A$)	3,125	10,199
Incremental cost per QALY gained (A$)	6,896	29,550
12 months		
Cost: No trial (A$)	1,539	326
Cost: N-of-1 (A$)	2,722	1,592
Incremental Cost (A$)	1,183	1,267
QALYs: No trial	0.821	0.801
QALYs: N-of-1	0.853	0.811
Incremental QALYs	0.032	0.010
Incremental cost per point reduction in pain (A$)*	4,254	11,209
Incremental cost per QALY gained (A$)	36,958	126,661
End of life		
Cost: No trial (A$)	15,279	2,791
Cost: N-of-1 (A$)	15,893	3,809
Incremental Cost (A$)	614	1,018
QALYs: No trial	9.040	7.604
QALYs: N-of-1	9.396	7.703
Incremental QALYs	0.356	0.099
Incremental cost per point reduction in pain (A$)*	2,209	9,004
Incremental cost per QALY gained (A$)	1,725	10,278

* The pain score, obtained from the trials, is constant for all time horizons modelled.

Table 4. Scenario where high-cost medication is available to N-of-1 responders only. Costs and cost per QALY gained compared with no trial over 5-years.

	Gabapentin		Celecoxib	
	No trial	**N-of-1**	**No trial**	**N-of-1**
Cost (A$)	5,654	6,251	1,193	2,255
Incremental Cost (A$)		597		1,062
QALYs	3.141	3.267	3.045	3.084
QALYs gained		0.126		0.039
ICER (A$)		4,761		26,934

Table 5. Results for the N-of-1 trial prescribing strategies: Costs, outcomes and cost-effectiveness of Gabapentin vs placebo and Celecoxib vs Paracetamol

	Gabapentin vs Placebo	Celecoxib vs Paracetamol
5 years (Base case)		
Cost: Non-responder (A$)	5,512	2,301
Cost: Responder (A$)	9,806	2,558
Incremental cost per responder (A$)	4,294	257
Pain score: Non-responder	4.130	4.550
Pain score: Responder	3.640	3.890
Reduction in pain (points)	0.490	0.660
QALYs: Non-responder	3.230	3.000
QALYs: Responder	3.381	3.492
Incremental QALYs	0.151	0.492
Incremental cost per point reduction in pain (A$)	8,764	389
Incremental cost per QALY gained (A$)	28,439	522
12 months		
Cost: Non-responder (A$)	2,515	1,586
Cost: Responder (A$)	3,391	1,624
Incremental cost per responder (A$)	876	39
QALYs: Non-responder	0.841	0.789
QALYs: Responder	0.891	0.918
Incremental QALYs	0.050	0.129
Incremental cost per point reduction in pain (A$)*	1,788	58
Incremental cost per QALY gained (A$)	17,525	298
End of life		
Cost: Non-responder (A$)	12,761	3,676
Cost: Responder (A$)	26,067	4,454
Incremental cost per responder (A$)	13,306	778
QALYs: Non-responder	9.302	7.493
QALYs: Responder	9.695	8.723
Incremental QALYs	0.393	1.230
Incremental cost per point reduction in pain (A$)*	27,156	1,179
Incremental cost per QALY gained (A$)	33,859	633

* The pain score, obtained from the trials, is constant for all time horizons modelled.

Table 6. Features of medications suited to an N-of-1 trial prescribing strategy.

Pharmacological features

- Indicated for symptomatic treatment of chronic conditions
- Rapid onset of action
- Short washout period

Clinical features

- Variable clinical response that cannot be predicted without a trial
- Superior clinical response to alternative but with a more serious adverse event profile
- Low overall response rate

Other features

- High differential in price compared with alternatives
- Limited demand at a national level