2011

The Impact of Inclusion Criteria in Health Economic Assessments

Richter, Anke; Thieda, Patricia; Thaler, Kylie; Gartlehner, Gerald

Adis Data Information

Richter, Anke, et al. "The impact of inclusion criteria in health economic assessments." Applied health economics and health policy 9 (2011): 139-148.
https://hdl.handle.net/10945/71398

PRACTICAL APPLICATION

Appl Health Econ Health Policy 2011; 9 (3): 139-148
1175-5652/11/0003-0139/$49.95/0

The Impact of Inclusion Criteria in Health Economic Assessments

Anke Richter,[1] *Patricia Thieda*,[2] *Kylie Thaler*[3] and *Gerald Gartlehner*[3]

1 Naval Postgraduate School, Monterey, California, USA
2 University of North Carolina – Chapel Hill, Chapel Hill, North Carolina, USA
3 Danube University, Krems, Krems, Austria

Contents

Abstract

The debate surrounding whether the findings of efficacy studies are applicable to real-world treatment situations is ongoing. The issue of lack of applicability due to a lack of clinical heterogeneity could be addressed by employing less restrictive inclusion criteria. Given that health economic assessments based on cost-effectiveness measures are required by many governments and insurance providers, the impact of this choice may be far reaching.

The objective of this article was to explore the use of a pilot study to examine the impact of inclusion criteria on cost-effectiveness results and clinical heterogeneity. A health economic assessment was conducted using QRISK®2 and simulation modelling of different population groups within the pilot study in Lower Austria. Patients were referred by their family physicians to 'Active Prevention' (Vorsorge Aktiv), a community-based lifestyle intervention focused on exercise and nutritional programmes. Cardiovascular risk factors were recorded before and after the intervention and translated to cardiovascular events.

As expected, enforcing restrictive inclusion criteria produced stronger and more irrefutable computations – in the expected number of events, the number of deaths, the incremental cost per life-year saved and in the 95% confidence interval. These findings provide insight into the issues surrounding clinical heterogeneity and the need for restrictive inclusion criteria. This is not a full health economic assessment of the intervention.

While inclusion criteria provide stronger results by limiting populations to those who would benefit the most, they must be enforced, both within and

outside the clinical trial setting. Enforcement has costs, both monetary and arising from unintended negative consequences of enforcement mechanisms. All these considerations will affect the results realized by the payer organization.

A pilot study can reveal whether an intervention may be cost effective 'enough' without restrictive inclusion criteria and can enable researchers to search for population subgroups in which the intervention remains cost effective. When the pilot study does not indicate sufficiently strong cost-effectiveness results, the broader trade-offs between clinical heterogeneity and the strength of the submission package to the reimbursement agency can be discussed by all parties. Payer concerns about the ability to generalize the results beyond the clinical trial can also be discussed at this time. Applicability then depends on the ability to enforce inclusion criteria similar to those used in the trials in the real world.

Clinical trials sponsored by large pharmaceutical firms or national agencies often pay physicians for participation in (and referrals to) the clinical trial. Strict inclusion criteria are enforced, and acceptance into the trial is restricted to only those meeting these criteria. Since these trials typically involve a new substance that is being introduced to the human organism, there are ethical requirements and legal barriers enforcing this design and implementation, as outlined by the Declaration of Helsinki and other ethics guidelines. The overall rationale of such studies, often termed efficacy or explanatory trials, is to minimize bias and to assess whether an intervention yields beneficial effects under ideal, highly controlled circumstances.[1]

There has been much discussion in the literature that inclusion criteria in such trials are too restrictive and do not adequately mirror the patient population that is likely to receive a new intervention.[2-5] A common criticism of such trials is that enrolled populations are highly selected and unrepresentative of the general population affected by the condition under consideration.[6] Recruitment frequently uses stringent eligibility criteria to ensure adherence, minimize adverse events, and lessen the potential for non-response.[7] The range of inclusion rates is also highly variable among trials. Some trials reportedly enrolled every person screened for eligibility; others screened as many as 68 people for each person finally enrolled.[8] Furthermore, physicians conducting clinical trials are often better trained and have access to better equipment than average physicians applying the same intervention in daily clinical practice. Consequently, the extent to which the findings of efficacy studies are applicable to broader, more diverse populations and real-world treatment situations remains unclear. The issue of lack of applicability and lack of clinical heterogeneity could be addressed by employing less restrictive inclusion criteria. Given that health economic assessments based on cost-effectiveness measures are required by many governments and insurance providers, the impact of this choice may be far reaching.

A pilot study may provide an opportunity to decide whether to use broad or narrow inclusion criteria in the full trial and may provide quantitative inputs to the cost-effectiveness discussions surrounding such a decision. We had the opportunity to examine the potential insights from using a pilot study in this manner. While the pilot study examined was not originally designed to facilitate such an analysis, the circumstances surrounding patient inclusion lent themselves to such a study and provided unexpected insights and benefits.

The State of Lower Austria runs various community-based programmes aimed at reducing cardiovascular disease (CVD) risk factors in the general population. One of these programmes, 'Active Prevention' (Vorsorge Aktiv), recently completed a pilot study to quantify the health

impacts of the intervention. Participating physicians were recruited from the community. However, many of the participating physicians felt that this intervention could be beneficial for many of their patients who did not meet the pre-defined inclusion criteria. As there were no enforcement mechanisms (either positive reinforcement such as payment to the physicians for referrals that met the inclusion criteria or negative reinforcement such as refusal to allow patients to participate in the programme), a mix of patients were referred to the trial centres. This is not unusual; prior studies have documented that physicians do not necessarily believe that potentially beneficial programmes (especially those that they believe will provide tangible benefits to their patients) should be restricted to only patients who fulfil the inclusion criteria.[9]

The end result was that the pilot study was essentially conducted without restrictive inclusion criteria. This allowed us to examine the impact of clinical heterogeneity by investigating the impact on cost-effectiveness calculations within different sub-populations. A broader discussion was then possible regarding the costs and benefits of requiring the more restrictive inclusion criteria to be met in the full trial that followed. This analysis was conducted from the perspective of the state healthcare agency (Niederosterreichischen Gesundheits- und Sozialfonds) responsible for both the clinical trials and the implementation of the intervention in the population if it were deemed successful.

1. Methods

A preliminary health economic assessment of the pilot study was conducted to provide a quantitative assessment of the impact of the inclusion criteria on the discussions surrounding cost-effectiveness results and general applicability of the intervention. The assessment compared the effectiveness of the intervention seen in the various population groups within the pilot study. If the cost-effectiveness results for the general population were favourable enough, there may be no need for inclusion criteria. If they were not, a broader discussion would need to be initiated. A modelling approach was necessary to translate the observed changes in cardiovascular risk factors to actual cardiovascular risk and then to actual events over a 10-year time horizon. The model predicted the number and types of CVD events pre- and post-intervention and the incremental cost per life-year gained of the intervention for each of three populations: (i) all participants who started the trial (intent-to-treat population; ITT); (ii) all those who completed the trial (completers population; 'completers'); and (iii) those who both completed the trial and met the inclusion criteria for elevated cardiovascular risk (inclusion criteria population; IC).

1.1 Pilot Study

The pilot study focused on a 6-month lifestyle intervention conducted in Lower Austria. A total of 124 participants were referred by family physicians to a community-based physical activity programme comprising 24 professionally supervised intensive sessions focused on developing an exercise routine and a nutritional regimen. The original intent was to enrol individuals with a high risk of CVD events: defined as a cardiovascular risk of >5% according to the New Zealand Cardiovascular Risk Calculator, or four or more points on the American Heart Association Risk Calculator. However, the protocol was not enforced and many physicians opted to enrol a wide variety of patients, including those with very little to no risk according to the chosen, established risk metrics.

Demographic and risk-specific baseline characteristics were captured via a survey form during the initial and follow-up doctors' visits. The study focused on changes in cardiovascular risk factors, including smoking, body mass index (BMI), total cholesterol, high-density lipoprotein (HDL) cholesterol and systolic blood pressure. The pre- and post-intervention measurements were carefully cleaned using the following steps. Missing data were imputed using the 'last observation carried forward' method. Individuals who reported being smokers at the conclusion were assumed to have been smokers at the beginning, although smokers were allowed to quit. Additional inputs of interest,

not captured in the initial pre-trial data forms but required for the cardiovascular event model, were obtained from a sub-sample of 74 patients and included rates of atrial fibrillation, chronic renal disease, rheumatoid arthritis, current treatment of hypertension and family history of coronary heart disease. All trial participants were assumed to be Caucasian as these data were not collected and this is the approximate ethnic distribution in Lower Austria. Furthermore, all patients were assumed to have no previous personal history of coronary heart disease.

1.2 Preliminary Cost-Effectiveness Model

To evaluate preliminary cost-effectiveness estimates for the different populations in the pilot study, we constructed a Monte Carlo simulation model to predict the number of cardiovascular events over a 10-year time horizon. A Monte Carlo simulation was necessary because of the small sample size in each population in the pilot study. The simulation provided more robust results that permitted determination of variations in the population results and their statistical significance. The simulation model was run twice for each of the three populations under consideration – once using the pre-intervention individual risk characteristics and again using the post-intervention individual risk characteristics. The differences between the two runs provided an estimate of the potential impact of the intervention in that population.

While it is possible to examine each pilot study population directly, because of the small number of individuals and the low overall cardiovascular risk of many of them (only approximately one-quarter of the participants met the definition of 'high risk'), the number of predicted events was very small. Having very few expected events makes interpretation of the simulation results difficult because the random variation of the simulation may be larger than the visible effects due to the intervention. To counter this issue and provide an estimation of the impact of the intervention in the full trial, we generated three populations of 300 individuals using the simulation package @Risk™ in Microsoft® Excel (i.e. one set of 300 individuals mimicking the ITT population, one set of 300 mimicking the completers and one set mimicking the IC population). We chose 300 individuals because this is the population size expected in the follow-on clinical trial. The three populations were generated using @Risk™ to ensure that the risk factors in the generated populations were distributed in the same manner as in the pilot study. This implies that each generated individual was given an age, sex, ethnicity, BMI level, systolic blood pressure, cholesterol ratio, smoking status, hypertension treatment status, diabetes status, family history of CVD status, chronic renal disease status and rheumatoid arthritis status. The distributions from which these values were simulated were created from the original pilot study data.

To estimate individual 10-year cardiovascular risk given the individual's risk factors, we chose to implement the QRISK®2 risk calculator.[10] We chose this risk model over the Framingham model because it has been shown that the latter over-predicts CVD in European populations.[11] In addition to requiring a model that more closely mirrored the European population, it was necessary to have a model that used risk factors that would be affected by the intervention – in particular BMI and smoking, as well as cholesterol, blood pressure and diabetes – and that predicted events, not just deaths. The QRISK®2 calculator is available for free as a web-based application[12] or for a fee as a batch processor. For each individual, the risk factors are typed into the model's input section and the software provides the 10-year probability of a CVD event. The exact calculations are documented in publications by the QRISK creators.[10]

After calculating the pre- and post-intervention 10-year cardiovascular risk for each individual, we simulated the populations over 10 years to determine the actual number and type of events. The model allowed individuals to experience up to three CVD events. The probabilities of follow-on events were based on the CAPRIE (Clopidogrel versus Aspirin in Patients at Risk of Ischaemic Events) trial.[13] For illustrative purposes, this event information was combined with the cost data obtained from published European studies

and expert opinions of Austrian physicians to estimate potential cost-effectiveness results.[14]

All costs were converted to year 2006 values, and a discount factor of 3% was applied to both costs and effects. Model parameters are available in the Supplemental Digital Content 1, http://links.adisonline.com/APZ/A37. The cost for the intervention was €1000 per person. The simulations were run 10 000 times to capture the variation in CVD events within these populations over the 10-year time horizon. The outcomes of interest for this pilot study included the number of deaths, the number of events, the life-years lost, estimated overall costs and projected incremental costs per life-year gained. Since the population sizes were the same, comparisons between populations as well as comparisons within populations are valid.

2. Results

Figure 1 shows the average number of cardiovascular events and deaths pre- and post-intervention in each of the three populations over the 10-year time horizon. The number of events pre-intervention was almost the same between the ITT (32.5 events) and the completers populations (32.2 events). Completers did experience a greater reduction in events post-intervention than the ITT population, but the difference was not large (an average of 5.2 and 2.6 events, respectively). Typically, completers would be expected to have fewer events since everyone in that population completed the intervention while not everyone in the ITT population would have fully completed the intervention. In contrast, the number of events in the IC population without the intervention was approximately 60% higher than the two other populations (51.2 events) and the post-intervention reduction was doubled (8.4 events). This difference in the intervention effect between the different populations was statistically significant using the Chi-squared test statistic at $p < 0.001$.

The numbers of deaths within the ITT and completers populations were also similar and the post-intervention reductions were an average of 1.0 and 1.9 deaths, respectively. In the IC population, the number of deaths without the intervention increased to approximately 18, compared with approximately 11 in the other two populations, and the reduction post-intervention increased to 3.0 deaths ($p = 0.05$).

Using the illustrative pan-European costs for cardiovascular events, the average incremental cost per life-year saved dropped by about 50%

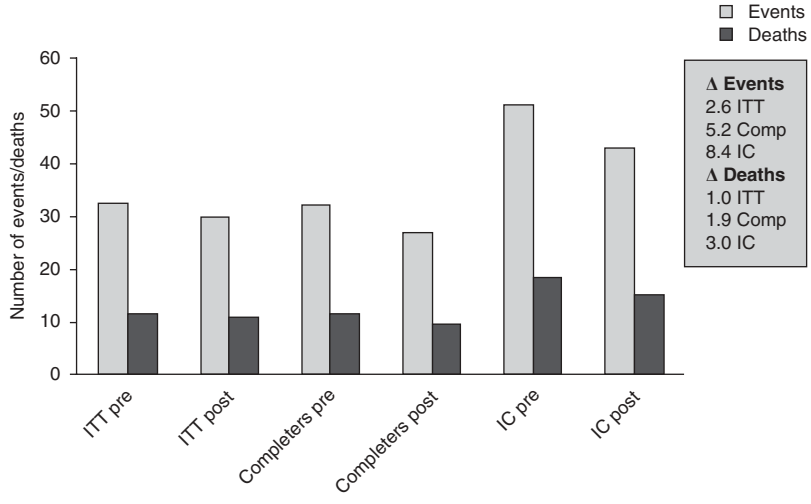

Fig. 1. Average number of cardiovascular events and average number of deaths by population and intervention status (n = 300, 10-year time horizon). **IC** = inclusion criteria population; **ITT** = intent-to-treat population; **post** = post-intervention; **pre** = pre-intervention.

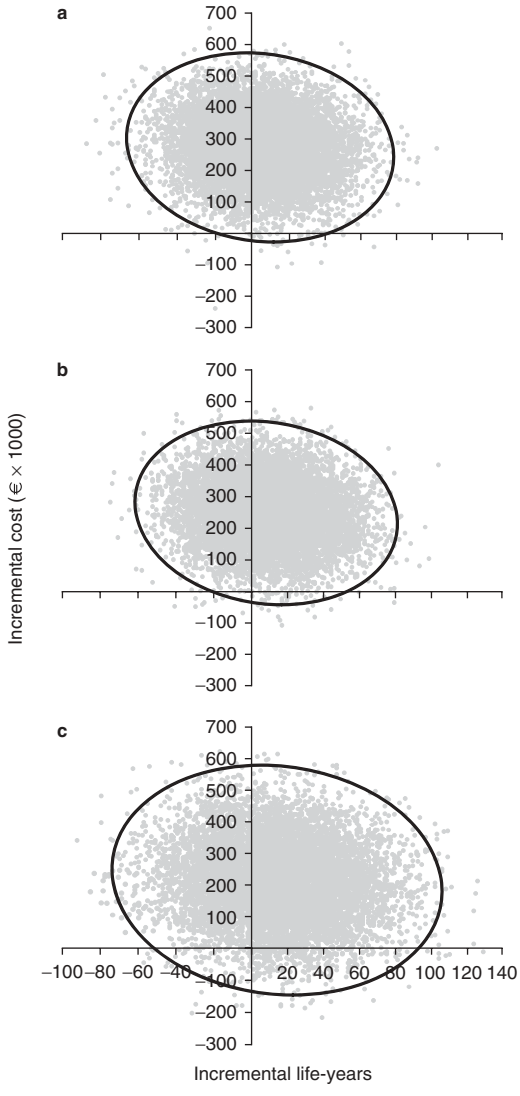

Fig. 2. Comparison of cost-effectiveness space by population with 95% confidence ellipse: (**a**) intent-to-treat population; (**b**) completers population; (**c**) inclusion criteria population.

from the ITT population (approximately €60 000 per life-year saved) to the completers population (approximately €28 000 per life-year saved). It dropped by about another 50% for the IC population (approximately €15 000 per life-year saved). This drop in average incremental cost effectiveness is remarkable but the variations around these average values are also important.

Since the analysis was a Monte Carlo simulation, it is possible to plot the incremental cost per life-year saved for each simulation run (10 000 in all) in one graph of cost-effectiveness space for each population. Figure 2 shows the cost-effectiveness space for each of the three populations. The ellipse represents the 95% confidence interval. In the ITT population, the scatter plot is almost centred around the y-axis (58% of the time, the post-intervention population experiences life-year gains). With the completers and the IC populations, the scatter plot shifts further to the right of the y-axis (post-intervention populations gained life-years 65% and 69% of the time, respectively). The IC population shows distinctly larger gains than the completers. In the ITT and completer populations, the incremental costs are almost uniformly greater than zero (99.7% and 99.4% are above the x-axis, respectively), so there are very few instances of cost savings. The costs are spread much further apart with the IC population, for which there are several instances (355 times; i.e. 4%) of cost savings.

3. Pilot Study Discussion

The pilot study results lead to an interesting discussion on the importance of more/less restrictive inclusion criteria. The 75% reduction in the incremental average cost-effectiveness results strongly favours more restrictive inclusion criteria. However, they require an educational campaign targeting physicians and an enforcement campaign (which costs time and money, including the potential for unintended negative consequences of enforcement tactics, on either patients or physicians) to ensure that all patients meet the trial inclusion criteria. In addition, if the cost-effectiveness results are to be realized by the payer organization, then the inclusion criteria must be enforced and monitored in the general population – costing time and money. Whether campaigns to enforce strict inclusion criteria are worthwhile may draw parallels from the discussions on measures to reduce over-use of medical technology.[15] This is worrisome, as physicians clearly wish to (and did) enrol patients outside of the restrictive inclusion criteria.

Table I. Comparison of broader/tighter inclusion criteria – Austrian example

Less restrictive inclusion criteria	More restrictive inclusion criteria
Advantages	
Addresses clinical heterogeneity	75% reduction in the average incremental cost per life-year saved
Results likely to be seen when administered to the larger population	Cost-effectiveness result that is clearly positive Greater proportion of patients benefitting from the intervention
Payer can be more confident of return on investment	
Benefits not captured in life-years or CVD events may be discovered for lower risk population	
Disadvantages	
Incremental cost per life-year saved is rather high	Need to enforce inclusion criteria in clinical trial (costs, physician cooperation)
Some patients clearly not benefitting from the intervention Secondary benefits not given as much weight in most submissions	Need to find ways to limit recipients of intervention upon launch to larger population (cost of implementation, cost of monitoring and enforcement) Secondary benefits not examined

CVD = cardiovascular disease.

While cost-effectiveness results seen with the less restrictive inclusion criteria were rather high (although still below the threshold suggested by WHO), it is more likely that these results would be seen in a nationwide implementation of the intervention, and the payer organization could be more confident of their return on investment. Another consideration is that, with the less restrictive criteria, there are clearly patients who are receiving the intervention but not experiencing any benefits in terms of cardiovascular events avoided. With an intervention consisting primarily of exercise and nutrition, this may not be as disconcerting as it would be with a medical or pharmaceutical intervention that has the potential for negative side effects.

There is also the issue of secondary benefits, not captured in the traditional cost-effectiveness analysis. As seen in the general feedback survey of the pilot study, the intervention was much appreciated by those not currently at high risk for CVD events.[16] Even if the intervention is beneficial (at some level) for low-risk patients, it cannot be observed in the predicted level of cardiovascular events. While CVD events are the obvious and most defensible endpoint for a cost-effectiveness analysis of a cardiovascular intervention, the implications of this choice should be considered. In a study using less restrictive inclusion criteria, it may be very important to look for secondary benefits. In a study with more restrictive inclusion criteria, the primary benefit will be seen more clearly so there may be less interest in examining the secondary benefits. This is a complex choice, as there are potential quality of life and health benefits outside the cardiovascular field (e.g. impacts on diabetes, arthritis and other chronic conditions). These potentially important considerations raise the issue of primary versus secondary outcomes. If the intervention is found not to be statistically significantly different on its primary outcomes of reduced cardiovascular morbidity and mortality due to its lack of enforced inclusion criteria, can this loss be overcome by positive secondary outcomes?

The final, and arguably most important, issue is the question of the interplay of all of these factors into the approval process of the country. If excellent cost-effectiveness scores are the main criteria for approval, then there is a strong argument for more restrictive inclusion criteria. If clinical heterogeneity and payer impact are more important factors, then the less restrictive inclusion criteria would be favoured. This deliberation is summarized in table I.

4. General Discussion

While we examined a specific pilot study, there are broader implications. Clinical trials with very restrictive inclusion criteria will be faster and less expensive to conduct than those with broader

inclusion criteria. Restrictive inclusion criteria will also provide stronger clinical results and, consequently, stronger cost-effectiveness results, and fewer side effects or adverse events. The problem will be the applicability of these results to the real world, both in terms of clinical benefits obtained and cost implications for the payer organization. The more heavily restricted the clinical trial population, the less likely the same results will be seen in practice.

The key is the approval process. If the approval process is based primarily on the achievement of a series of effectiveness and cost-effectiveness goals, then companies have an incentive to meet these, even if it means restricting their trial populations. If the approval process requires more inclusive trials, then companies need to determine whether they feel that their product can meet the expected achievement levels with the broader population. The providers' natural incentives are to find a (restricted) population wherein their intervention is most likely to succeed, and present those results. They are already balancing the potential size of the market with the risk of regulatory failure. Sufficiently large increases in the potential market may lessen their opposition to a broader trial; however, it is not reasonable to expect providers of interventions to accept the risk of having a multi-million dollar clinical trial fail due to issues surrounding inclusion criteria.

Further complicating this decision process is the lack of clarity surrounding approval levels and the expectations of the bodies governing the approval process. For example, there may be no clear standard regarding the relative importance of the presentation of cost effectiveness within the submission package and/or what level of cost effectiveness is required in order for the intervention to be adopted. Some countries have well accepted standards (e.g. £25 000–35 000 per QALY in the UK).[17] The WHO has suggested that the value of a gain of one perfectly healthy life-year could be worth as much as three times the average GDP; however, it is not clear that governments necessarily feel bound by these recommendations.[18] Given a known level of required cost effectiveness, simulations after the initial pilot study can help determine the necessity of restrictive inclusion criteria. If a mixed population provides sufficient evidence that the intervention meets the prescribed cost-effectiveness threshold, then the full trial can proceed with a mixed population, and the needs for addressing clinical heterogeneity and approval barriers are simultaneously met. If a mixed population does not provide sufficient evidence to convince decision makers of an intervention's cost effectiveness then, if inclusion criteria are not enforced, the clinical trial will not be considered economically worthwhile, despite positive clinical results.

Table II. Comparison of alternative study designs – general

Less restrictive inclusion criteria	More restrictive inclusion criteria	Stratified clinical trial
Advantages		
Effects of clinical heterogeneity visible Better clinical information Clearer picture of impacts in real-world setting Results likely to be obtained by payer; potential for larger market	Smaller, less expensive clinical trial Stronger responses Faster responses	Potential for including multiple patient categories in clinical trial Results more likely to show clinical heterogeneity Results more likely to be obtained by payers
Disadvantages		
Higher cost of trial Longer duration of trial More patients receiving intervention and not benefitting from it Potential for more side effects or adverse events Potential for greater risk of non-approval	Does not address clinical heterogeneity Requires enforcement of inclusion criteria in trial Requires enforcement of inclusion criteria in real-world setting (if similar results are to be obtained); potential for smaller market Results may not be obtained by payer agency Less complete clinical information	Trial will be more expensive Additional time required to run full trial Need to agree on cost-effectiveness reporting requirements for approval process Need to ensure trial sponsors are not penalized for including extra patients

If the approval process goes beyond a simple formulaic assessment to maintain flexibility in the system to address a host of extenuating circumstances and secondary impacts (e.g. clinical issues regarding disadvantaged populations and orphan illnesses; economic issues such as the number of competitor drugs on the market; national issues around supporting local companies), then discussions are necessary to find mutually agreeable targets and populations. While these discussions will require considerable time and effort, they reduce the risks for the provider companies and clarify the trade-offs for all concerned. Table II summarizes these issues.

A possible solution to this dilemma is to have stratified clinical trials.[19] In this case, an unstructured pilot study could be used to identify subgroups of patients in which the intervention is likely to produce benefits as well as subgroups of patients in which the intervention is likely to produce an unfavourable benefit-harm ratio. The successful groups would be included in the main trial. This should include the types of patients most likely to use this intervention in practice as well as the types of patients of greatest interest to the payer organization. The sponsor of the intervention and the board governing the approval process would then need to come to an agreement on which subgroup results need to meet the achievement standards for approval and need to be officially reported. This would avoid penalizing the sponsor for including more patients and more types of patients in the clinical trial and yet provide more complete clinical information on the intervention. In addition, if the resultant trial were larger than the originally planned trial, then discussions could be had about the possibility of subsidizing the portions of the trial that focus most directly on the subgroups of interest to the payer organizations. The results of the stratified study would provide effectiveness and cost-effectiveness results for multiple subsets of patients and would thereby address payer concerns about the generalizability of the trial results. The negotiations necessitated by this process would ensure that the sponsor of the clinical trial would not be penalized, and could potentially even be rewarded, for conducting the larger, more inclusive trial.

5. Conclusions

While inclusion criteria provide stronger results by limiting populations to those who would benefit the most, they must be enforced, both within and outside the clinical trial setting. Enforcement has costs, both monetary and arising from unintended negative consequences of enforcement mechanisms. All these considerations will affect the results realized by the payer organization.

A pilot study can reveal whether an intervention may be cost effective 'enough' without restrictive inclusion criteria and can enable researchers to search for population subgroups in which the intervention remains cost effective. When the pilot study does not indicate sufficiently strong cost-effectiveness results, the broader trade-offs between clinical heterogeneity and the strength of the submission package to the reimbursement agency can be discussed by all parties. Payer concerns about the ability to generalize the results beyond the clinical trial can also be discussed at this time. Applicability then depends on the ability to enforce inclusion criteria similar to those used in the trials in the real world.

Acknowledgements

Financial support for this study was provided by the Niederosterreichischen Gesundheits- und Sozialfonds, Stattersdorfer Hauptstrasse 6A, 3100 St. Pölten, Austria. The funding agency is also the sponsor of the pilot study discussed in this article as well as the full-sized clinical trial that ensued. They placed no restrictions or requirements on the publication of this article. The authors have no conflicts of interest.

The results of this study were presented at two meetings: the European Forum for Evidence-Based Prevention (EUFEP) Congress; Baden (Austria); 2009 June 24–26 and the Institute for Operations Research and Management Sciences (INFORMS) Annual Meeting; San Diego (California, USA); 2009 October 11–14.

The authors thank Mark Helfand, MD, MPH, from the Oregon Health and Science University for his thorough review and excellent suggestions for improving the paper.

References

1. Godwin M, Ruhland L, Casson I, et al. Pragmatic controlled clinical trials in primary care: the struggle between external and internal validity. BMC Med Res Methodol 2003 Dec 22; 3: 28

2. Rothwell PM. External validity of randomised controlled trials: "to whom do the results of this trial apply?" Lancet 2005; 365 (9453): 82-93
3. Van Spall HG, Toren A, Kiss A, et al. Eligibility criteria of randomized controlled trials published in high-impact general medical journals: a systematic sampling review. JAMA 2007; 297 (11): 1233-40
4. Steg PG, López-Sendón J, Lopez de Sa E, et al. External validity of clinical trials in acute myocardial infarction. Arch Intern Med 2007; 167 (1): 68-73
5. Malmivaara A, Koes BW, Bouter LM, et al. Applicability and clinical relevance of results in randomized controlled trials: the Cochrane review on exercise therapy for low back pain as an example. Spine (Phila Pa 1976) 2006; 31 (13): 1405-9
6. Jones R, Jones RO, McCowan C, et al. The external validity of published randomized controlled trials in primary care. BMC Fam Pract 2009 Jan 19; 10: 5
7. Gartlehner G, Hansen RA, Nissman D, et al. A simple and valid tool distinguished efficacy from effectiveness studies. J Clin Epidemiol 2006; 59 (10): 1040-8
8. Gross CP, Mallory R, Heiat A, et al. Reporting the recruitment process in clinical trials: who are these patients and how did they get there? Ann Int Med 2002; 137 (1): 10-6
9. Vanderpool HY, Weiss GB. False data and the therapeutic misconception: two urgent problems in research ethics. False data and last hopes: enrolling ineligible patients in clinical trials. Hastings Cent Rep 1987; 17 (2): 16-9
10. Hippisley-Cox J, et al. Predicting cardiovascular risk in England and Wales: prospective derivation and validation of QRISK2. BMJ 2008; 336 (7659): 1475-82
11. Hense HW, Schulte H, Löwel H, et al. Framingham risk function overestimates risk of coronary heart disease in men and women from Germany: results from the MONICA Augsburg and the PROCAM cohorts. Eur Heart J 2003; 24 (10): 937-45
12. The QRISK®2-2010 cardiovascular disease risk calculator [online]. Available from URL: http://www.qrisk.org/ [Accessed 2011 Mar 29]
13. Berger K, Hessel F, Kreuzer J, et al. Clopidogrel versus aspirin in patients with atherothrombosis: CAPRIE-based calculation of cost-effectiveness for Germany. Curr Med Res Opin 2008; 24 (1): 267-74
14. Levy E, Gabriel S, Dinet J. The comparative medical costs of atherothrombotic disease in European countries. Pharmacoeconomics 2003; 21 (9): 651-9
15. Gandjour A, Lauterbach KW. When is it worth introducing a quality improvement program? A mathematical model. Med Decis Making 2003; 23 (6): 518-25
16. Department fur Evidenzbasierte Medizin und Klinische Epidemiologie. Final report: evaluation of the pilot program Vorsorge Aktiv ('Active Prevention') [in German; online]. Available from URL: http://www.donau-uni.ac.at/imperia/md/content/department/evidenzbasierte_medizin/projekte/evaluation/vorsorge_aktiv.pdf [Accessed 2011 Mar 29]
17. Rawlins MD, Culyer AJ. National Institute for Clinical Excellence and its value judgments. BMJ 2004; 329 (7459): 224-7
18. Eichler HG, Kong SX, Gerth WC, et al. Use of cost-effectiveness analysis in health-care resource allocation decision-making: how are cost-effectiveness thresholds expected to emerge? Value Health 2004; 7 (5): 518-28
19. Sculpher M, Gafni A. Recognizing diversity in public preferences: the use of preference sub-groups in cost-effectiveness analysis. Health Econ 2001; 10 (4): 317-24

Correspondence: Dr *Anke Richter*, 699 Dyer Rd., Bldg 234, Naval Postgraduate School, Monterey, CA 93943, USA.
E-mail: arichter@nps.edu