National Clinical Guideline Centre Cost-Effectiveness Assessment for the National Institute for Health and Clinical Excellence

David Wonderling, MSc; Laura Sawyer, MSc; Elisabetta Fenu, MSc; Kate Lovibond, MSc; and Philippe Laramée, DC, MSc

The National Clinical Guideline Centre (NCGC) develops evidence-based clinical guidelines on behalf of the National Institute for Health and Clinical Excellence (NICE) in the United Kingdom. The U.K. Department of Health has commissioned NICE to make recommendations on the basis of both clinical effectiveness and cost-effectiveness. This article describes how cost-effectiveness is evaluated and accounted for in NCGC guidelines. Six recent case studies are presented, in which consideration of cost-effectiveness has informed recommendations in various ways for clinical guidelines on alcohol use disorders, chronic obstructive pulmonary disease, glaucoma, lower urinary tract symptoms, non-ST-segment elevation myocardial infarction and unstable angina, and venous thromboembolism prophylaxis. Some of the challenges we face in trying to account for cost-effectiveness in clinical guidelines are outlined, as well as some of the difficulties in adapting cost-effectiveness guidelines for other settings.

Why Consider Cost-Effectiveness in Guidelines?

No health care system has enough resources to provide every clinically effective intervention to all persons who could benefit. Health economists assert that to maximize population health, health services should prioritize clinical practices (diagnostic, therapeutic, or palliative) that yield a relatively high health gain per dollar spent over those that give a smaller health gain per dollar spent (prioritize interventions that are the most cost-effective). Clinical guidelines give advice on which treatments to initiate under which circumstances and are therefore ideal vehicles for promoting cost-effective clinical practice.

The role of cost-effectiveness analysis in health care decision making is a contested issue, particularly in the United States. The 2009 report from the Institute of Medicine highlights this issue (2), but notably concludes that cost-effectiveness analysis is a useful tool for comparative effectiveness research. Reluctance to consider cost-effectiveness is often due to misconceptions, some of which we try to dispel here.

Myth 1: The Aim of Health Economics Is to Save Money

No. The aim is to get as much health gain as possible for the population, given available resources. Health economists are also interested in achieving a fairer distribution of health (although this is a particularly controversial area, because what one person considers fair, another may not).

Myth 2: Costly Interventions Are Not Cost-Effective

Not necessarily. The cost of the intervention may be partially or completely offset by cost savings due to a reduction in clinical events or side effects. Alternatively, if the intervention produces large health gains, this may justify the extra cost.

Myth 3: The Main Aim of Cost-Effectiveness Analysis Is to Aid Implementation of Clinical Practice Recommendations

Not really. Although a cost-effectiveness analysis can inform implementers of resource requirements (how much it will cost), the main aim is to inform decisions about which practices to recommend (whether it gives good value). This sometimes means that a clinically effective treatment is not recommended because it does not provide sufficient benefit to justify the cost.

Cost-Effectiveness Analysis

Cost-effectiveness analysis is a method for assessing which interventions are the most cost-effective. The simplest scenario includes the standard therapy and a new therapy, but more than 2 therapeutic options are often available. In cost-effectiveness analysis, we calculate the mean health outcome and mean health care cost for each intervention being compared. Health care costs include not
only the cost of the drug or procedure but also the cost of the whole patient pathway (for example, preventing deliri-um could reduce both the length of hospital stay and the number of admissions to long-term care homes). The health outcome should be suitable for comparing the interventions in the analysis; for example, reduction in blood pressure for comparing antihypertensive medications. However, an outcome that can be applied more generally should be used when prioritizing across specialties. The life-year (mean survival) has been used for this purpose, although it does not capture differences in quality of life. The quality-adjusted life-year (QALY) incorporates both survival and health-related quality of life. To estimate QALYs for a particular intervention, we multiply the mean life-years by a utility, a measure of health-related quality of life on a scale on which 1 is equivalent to full health and 0 is considered no better than death. Thus, 10 years at full health is equal to 10 QALYs and is also equivalent to 20 years at 50% health-related quality of life. Other approaches that incorporate both length and quality of life are also available, such as the disability-adjusted life-year (3) and the healthy-year equivalent (4). The Figure illustrates the decision rules of cost-effectiveness analysis (5).

Decision Rules for Only 2 Comparators

If the new treatment both improves health outcomes and reduces cost (Figure, bottom right quadrant in the top panel), then we usually recommend that the new treatment be offered to patients. If the new treatment both reduces health and increases cost (Figure, top left quadrant in the top panel), we consider the new treatment to be dominated and recommend staying with the standard treatment. However, if the new treatment is both more effective at increasing health and more costly (Figure, top right quadrant in the top panel), then we have to judge whether the health gain is large enough to justify the additional cost. This is usually done by calculating the incremental cost-effectiveness ratio (ICER) and comparing it with a pre-specified cost-effectiveness threshold. The ICER is the difference between the mean costs of each strategy divided by the difference in mean health outcome (the slope of the line that connects the strategies).

Decision Rules for 3 or More Comparators

Suppose that we have 5 possible comparators: A, B, C, D, and E (Figure, bottom panel). First, we eliminate dominated strategies; in this case, treatment C is dominated by D. Second, we eliminate strategies that are subject to extended dominance, which occurs when a less effective intervention has a higher ICER when both are compared with a mutual comparator. Treatment B is less effective than D but has a higher cost per QALY gained when each is compared with A; adopting B would mean paying a higher cost per QALY than is necessary to achieve that level of effective-ness. The third and final step is to look at each remaining intervention, calculate its ICER compared with the next most effective option, and see which ICER (if any) is below the cost-effectiveness threshold. The most cost-effective strategy will be the most effective strategy that has an ICER below the cost-effectiveness threshold. In this example, 3 strategies (A, D, and E) remain. Although treatment E is the most effective, it has a relatively high ICER compared with D. Therefore, as the most effective treatment with an ICER below the threshold, treatment D is the most cost-effective.

Methods of NICE

Cost-effectiveness is considered at every stage in the production of NICE clinical guidelines, from scoping to assessing the need for an update (Table 1). These methods are described in the Guidelines Manual (6); here we provide an overview of how these methods are applied at the National Clinical Guideline Centre (NCGC).

The NICE Reference Case

Health economists do not always agree about the details of conducting a cost-effectiveness analysis, and different methods are appropriate for different decision contexts. Therefore, NICE has implemented a set of rules for conducting cost-effectiveness analysis, known as the reference case, which pro-
vides a benchmark to allow different guideline development groups (GDGs) to make decisions in a fair and consistent manner (7). These groups sometimes use different methods, usually for sensitivity analyses, but they need a strong rationale for doing so. Table 2 summarizes the NICE reference case. The following aspects should be noted.

First, QALYs should be the primary health outcome used in any economic evaluation. The preferred measure of utility for estimating QALYs in adults is the EuroQol-5D (EQ-5D) questionnaire, a generic, preference-based measure comprising a classification of health states for which a table of values has been derived. Patients (or caregivers) should use the EQ-5D questionnaire to report descriptions of health states at different time points; the 5 dimensions can be scored on 3 levels of severity. The U.K.—based table of values for the EQ-5D states have been elicited from a sample of 3395 members of the general public by using time-tradeoff score (8). The Health Utility Index 2 is the validated measure preferred for measuring utility in children. Alternative measures may be used in the absence of these preferred measures or for sensitivity analyses.

Second, because the aim is to maximize health gained from use of the NHS budget, the costs considered should include only those incurred by the NHS and personal social services, not those directly incurred by patients or other government sectors. Productivity gains and losses (due to time off work) are not considered, to avoid double-counting the benefits already captured in the QALY or giving undue weight to the treatment of patient groups who are of working age.

Cost-Effectiveness Decision Rules

For a treatment strategy to be considered cost-effective, its ICER must be compared with a prespecified

Table 2. Summary of the National Institute for Health and Clinical Excellence Reference Case

<table>
<thead>
<tr>
<th>Element of Technology Appraisal</th>
<th>Reference Case</th>
</tr>
</thead>
<tbody>
<tr>
<td>Comparators</td>
<td>Therapies routinely used in the NHS, including technologies regarded as current best practice</td>
</tr>
<tr>
<td>Perspective on costs</td>
<td>NHS and personal social services</td>
</tr>
<tr>
<td>Perspective on health outcomes</td>
<td>All health effects on individuals</td>
</tr>
<tr>
<td>Type of economic evaluation</td>
<td>Cost-effectiveness analysis</td>
</tr>
<tr>
<td>Synthesis of evidence on outcomes</td>
<td>Based on a systematic review</td>
</tr>
<tr>
<td>Measure of health effects</td>
<td>QALY</td>
</tr>
<tr>
<td>Source of data for measurement of health-related quality of life</td>
<td>Reported directly by patients or caregivers</td>
</tr>
<tr>
<td>Discount rate</td>
<td>An annual rate of 3.5% on both costs and health effects</td>
</tr>
<tr>
<td>Equity weighting</td>
<td>An additional QALY has the same weight regardless of the other characteristics of the persons receiving the health benefit</td>
</tr>
</tbody>
</table>

NHS = National Health Service; QALY = quality-adjusted life-year.
* From reference 7.
cost-effectiveness threshold. To determine the appropriate threshold, one would theoretically assess the ICER for every treatment; arrange the treatments in order of cost-effectiveness; and then implement the treatments, in order of increasing ICER, until the budget is depleted. The ICER associated with the last treatment implemented would represent the cost-effectiveness threshold. In practice, relatively few interventions have a known ICER, so the threshold must be estimated and then modified if it seems too lenient or too strict.

Decisions about whether to recommend interventions are made by a GDG, which comprises clinical experts, patients or caregivers, a project manager, a systematic reviewer, an information scientist, and a health economist. The GDG is expected to consider not only cost-effectiveness but also other factors when developing guidance, including the need to distribute health resources in the fairest way across society as a whole. The principles that GDGs should consider when judging whether an intervention offers good value are set out by NICE (9, 10). In general, an intervention is considered to be cost-effective if it dominates other relevant strategies or costs less than £20 000 ($30 000; 2010 purchasing power parity, $1 = £0.657 [11]) per QALY gained compared with the next best strategy.

If the GDG recommends an intervention that is estimated to cost more than £20 000 per QALY gained, or does not recommend one that is estimated to cost less than £20 000 per QALY gained, the reasons for this decision must be discussed explicitly in the full version of the guideline. The discussion should refer to issues regarding the plausibility and certainty of the estimate of incremental cost-effectiveness, as well as such factors as whether the intervention adds demonstrable and distinct nonhealth benefits or the requirements to prevent discrimination and promote equality (10). For interventions with ICERs of £30 000 ($46 000) per QALY or greater, the GDG must make an increasingly stronger case for recommending the intervention as an efficient use of NHS resources with respect to these factors.

**Review of Health Economics Literature**

Published economic studies relevant to any of the topics being addressed in the guideline are sought through a literature review. The aim is to include studies that are helpful for decision making in the context of the guideline and the setting (the United Kingdom), because it is not possible to undertake new analyses for all areas of the guideline.

The population and intervention inclusion criteria for economic studies are the same as for the clinical review of the literature. Noncomparative cost studies, including cost-of-illness studies, are excluded because they are not helpful in assessing whether something is cost-effective.

Economic studies are assessed for applicability and methodological quality by using the NICE economic evaluation checklist (Appendix Table 1, available at www.annals.org). Studies are prioritized for inclusion in the guideline on the basis of their relative applicability and methodological limitations. For example, if a high-quality analysis that is directly applicable to the United Kingdom is available, other, less relevant studies may be excluded. A study is considered applicable if it covers the appropriate population and interventions and complies with the NICE reference case. The methodological limitations of a study are judged by how well the model structure reflects the nature of the condition; whether the time horizon is sufficiently long enough; and whether the estimates of relative treatment effects, resource use, and unit costs are from sources with a low risk for bias.

Included economic studies are summarized for the GDG, usually in a standardized table called an economic evidence profile (12). This parallels the Grading of Recommendations Assessment, Development, and Evaluation profile tables (13) used to summarize the clinical evidence. The profile for each economic study shows the assessment of applicability and methodological quality, with footnotes that indicate the reasons. It also shows incremental costs, incremental outcomes (such as QALYs), and the ICER, as well as information about the assessment of uncertainty in the analysis (such as CIs or the results of sensitivity analyses).

**Modeling**

In addition to a review of the economic literature, 1 or more decision models are developed for almost every guideline to assess cost-effectiveness in high-priority areas. New or updated cost-effectiveness analyses cannot usually be done for all areas of the guideline; therefore, the health economist and the GDG decide which areas are most important for modeling at the beginning of the guideline development process. Prioritized topics are those for which current practice varies, uncertainty about the cost-effectiveness of interventions is high, or a change in practice may have substantial cost implications (because of a big difference in cost per patient or because the question covers a large patient population). Analysis may not be necessary, for example, if a high-quality and recent estimate of cost-effectiveness relevant to the United Kingdom is identified in the economic literature review, or clinical trial evidence shows that the intervention with the lowest acquisition cost is also the most clinically effective. A systematic review of the clinical evidence will show how one treatment compares with another with respect to different outcomes but usually cannot show which treatment is best overall, because some outcomes will favor one intervention and others will favor a comparator. Decision analysis is invaluable for systematically bringing together all of the evidence to estimate which treatment is best overall. Decision models provide a structured framework in which to evaluate the clinical and economic consequences of choosing one strategy over another, given the inherent probabilities of events (14, 15). Useful features of these models include the ability to synthesize randomized and observa-
cost-effectiveness analysis, with the possibility of extending clinical data beyond the length of follow-up or extrapolating intermediate clinical end points to final outcomes.

Expected costs and outcomes are calculated for each strategy on the basis of defined parameters, such as probabilities, estimates of resource use, costs, and utilities. Data used to estimate these parameters should come from the best available sources. A meta-analysis of randomized trials is generally considered the best source for effect estimates, but a network meta-analysis may be required when several direct and indirect comparisons are available. Meta-analyses and network meta-analyses are often conducted by the guideline team, although sometimes a relevant and current analysis from the literature is used. To attach costs to the resource units, the most reliable figures are obtained from national sources in the United Kingdom, such as the NHS reference costs for hospital stays and procedures or the British National Formulary or NHS drug tariff for cost of drugs. Utilities associated with the condition of interest are retrieved through systematic searches and selected on the basis of the NICE reference case. When no published data are available for a model parameter, expert opinion is sought and assumptions are made.

To test the robustness of the model to any uncertainty about parameters, structure, and assumptions, additional sensitivity analyses are conducted. A one-way sensitivity analysis, in which a single parameter is varied within its plausible range, is the simplest form. The intrinsic uncertainty around point estimates is dealt with by using a probabilistic sensitivity analysis, in which all parameters are entered into the model as distributions. The model is run thousands of times, and values are sampled from the defined distributions in each simulation. This allows us to assess how confident we are in the results of the model.

In addition to peer review of the guideline during the stakeholder validation phase, the models at the NCGC are systematically validated by testing the model parameters to ensure that the results change in a plausible way, having a second health economist perform quality assurance of all data and formulae in the model, and disaggregating results to allow the GDG to assess the plausibility of the model (for example, by breaking down costs into resource categories).

Case Studies

To demonstrate the various ways that cost-effectiveness is considered in guidelines, we present 6 case studies from recently published NICE guidelines. Each case study provides an example of how the previously outlined methods were used to address a particular decision problem in a guideline.

Case Study 1: Alcohol Use Disorders

For the guideline on alcohol use disorders, the clinical review of the literature identified a single randomized trial that compared surgical drainage of the pancreatic duct with endoscopic drainage in the treatment of patients with obstructive chronic pancreatitis. Surgery was associated with better pain control and improved quality of life but was a more costly procedure. The guideline team estimated the overall effect on costs and QALYs by using data from this trial, along with standard NHS unit costs. The results of the cost-effectiveness assessment showed that surgical drainage was both more effective and cost-saving than endoscopy. Although no surgical deaths occurred in the trial (which had a relatively small sample size), the GDG was concerned about the potential risk for death with surgical drainage. A separate review of case series indicated a risk of about 1%. In a sensitivity analysis, the model demonstrated that even with an implausibly high mortality rate of 2%, total QALYs would still be increased by the considerable improvements in quality of life and surgery would still be cost-effective.

Case Study 2: Chronic Obstructive Pulmonary Disease

In contrast to case study 1, in which the more effective option was cost-saving, an analysis for the guideline on chronic obstructive pulmonary disease demonstrated that the more effective strategy was also more costly. This cost-effectiveness analysis compared therapy with a long-acting muscarinic agonist (LAMA), a long-acting β-agonist (LABA) concurrent with inhaled corticosteroids (ICS), or all 3 agents (LAMA, LABA, and ICS) for initial maintenance therapy in patients with an FEV1 less than 50%. The primary analysis found that either LAMA or LABA plus ICS was the most cost-effective option, depending on the clinical data used. Triple therapy was more effective than either but was not considered a cost-effective initial therapy because its ICER was more than £90 000 ($137 000). Therefore, the GDG only recommended triple therapy as a subsequent option for patients who remain symptomatic after first-line therapy with LAMA or LABA plus ICS.

Case Study 3: Unstable Angina and Non–ST-Segment Elevation Myocardial Infarction

Developing an economic analysis to answer a review question is sometimes unnecessary, because a recent evaluation of cost-effectiveness applicable to the NHS is available from the published literature. For the guideline on unstable angina and non–ST-segment elevation myocardial infarction, the systematic literature review identified a published analysis that addressed the question of whether patients should be managed with an early invasive or a conservative management strategy. The study was assessed as being fully applicable and having only minor limitations. This question was considered a high priority for economic analysis by the GDG, and the NCGC would have conducted a new analysis to inform recommendations had the study not been identified.

Case Study 4: Preventing Venous Thromboembolism

Decision models are useful when a decision is needed regarding a tradeoff between clinical efficacy and adverse
events. For example, pharmaceutical thromboprophylaxis in hospitalized patients reduces the incidence of venous thromboembolism, but at the expense of increased bleeding, and both conditions are associated with mortality or long-term morbidity in severe cases. For the update to the guideline on venous thromboembolism (25), a network meta-analysis was conducted to synthesize all relevant randomized trial data, including studies identified since the original guideline was published. The output of the network meta-analysis helped to determine which of the strategies was most effective at reducing the incidence of venous thromboembolism and which entailed the most risk for major bleeding events. Using the network meta-analysis evidence, the original decision model was updated with observational data to estimate the overall effect on mortality, quality of life, and cost-effectiveness. The overall effectiveness and cost-effectiveness results were found to vary by baseline risks for both bleeding and venous thromboembolism, and so the recommendations differed by subpopulation (medical inpatients, general surgery, hip replacement, hip fracture surgery, and knee replacement surgery).

Case Study 5: Glaucoma

Separate subgroup analyses should be done when differences in baseline risks, treatment effects, potential gains in quality of life, or costs may affect the relative cost-effectiveness of interventions; what is not cost-effective for patients with mild disease may be cost-effective for those with severe disease. For the guideline on glaucoma (26), clinical evidence indicated that the baseline risk for glaucoma in patients with ocular hypertension was determined by a combination of age, intraocular pressure, and central corneal thickness. By rerunning the decision model for subgroups that differed by each of these factors, the GDG could target drug prevention toward the highest-risk patients and avoid the harms and unnecessary cost associated with long-term treatment for patients at lower risk. The results of the subgroup analyses were reflected in the final recommendation on management of ocular hypertension.

Case Study 6: Lower Urinary Tract Symptoms

A key aim of decision modeling is to account for the variability and uncertainty related to the parameters used in the model. Because probabilities and other model inputs have a level of uncertainty, it is necessary to explore how this uncertainty affects the decision being made. For the guideline on lower urinary tract symptoms in men (27), a model compared transurethral resection of the prostate with holmium laser enucleation. Several parameters in the model had a high degree of uncertainty, and very small changes in some parameters (such as the probability of treatment failure or treatment adverse events) changed the optimal strategy. The probabilistic analysis showed split results, with resection and enucleation being cost-effective in 52% and 48% of the simulations, respectively, when a threshold of £20,000 ($30,000) per QALY was adopted; this made recommending one intervention over the other inappropriate.

Challenges to Developing Cost-Effective Clinical Guidelines

Health economics has been an important input to NICE guidelines since their inception, and although the methods and understanding of cost-effectiveness analysis have improved over time, challenges still remain. Unlike the more established role of systematic reviews and meta-analysis, decision analysis and cost-effectiveness analysis will be new to some members of the GDG. At the start of each guideline, health economists make every effort to dispel the aforementioned myths and introduce key concepts, such that economic considerations are successfully incorporated into the guideline. During this initial period, the GDG must reach a consensus on the highest-priority topics for economic analysis. This can prove challenging, because not enough information may be available to make this decision; furthermore, the health economist’s knowledge of the clinical area and the clinician’s understanding of health economics are both at their weakest. Once topics have been prioritized and analyses are under way, the health economist and the GDG work in close collaboration, designing plausible treatment pathways and identifying the best available data inputs. Some GDGs find it difficult to decide on a single approach or particular data sources and therefore plan sensitivity analyses. Input of clinical and patient members during model development is essential to improve the applicability and accuracy of the model and ensure the credibility of the analysis within the medical community.

When interpreting the evidence, GDGs soon discover the differences between rules used in evidence-based medicine and those used in decision analysis. The latter system endeavors to compare all relevant strategies simultaneously, whereas the former typically compares strategies in a pairwise fashion. While decisions guided by evidence-based medicine have historically focused on statistically significant differences (or lack thereof) between strategies, decision rules in economic analysis are based on society’s willingness to pay for expected health gains. Several interventions may each be effective compared with placebo, but not all may be deemed cost-effective when considered together.

Another area with inherent challenges is quantifying cost-effectiveness, whether through review or original analysis. Recent U.K. cost-effectiveness studies of good methodological quality do not exist for most questions, and updating existing analyses or undertaking new ones is challenging. Inputs to the decision models often come from a range of sources, and each source must be critically appraised to ensure that we are using the data with the lowest risk for bias. Data may also need to be transformed before they can be used in the model (such as from rates to risks).
Estimates of NHS resource use are not always available, and in some cases must be solicited directly from GDG members. In addition, quality-of-life data that are suitable for calculating QALYs are limited for some conditions. Finally, GDGs often face decision problems for which little or no clinical evidence exists, but the cost difference between strategies is substantial. Demonstrating cost-effectiveness (or ineffectiveness) can be difficult in these circumstances, and models are used as a tool to explore different “what-if” scenarios rather than decide on the most cost-effective strategy. When modeling is not feasible or a high-enough priority, a critical discussion of the uncertainty in cost-effectiveness accompanies a weak recommendation.

Although using QALYs to assess cost-effectiveness has been widely accepted by health economists, some methodological issues remain (28, 29). Many of these relate to how utility is quantified and concerns about fairness. For example, what is the most appropriate instrument for assessing utility? Should QALYs for people with more severe conditions be weighted higher than QALYs for people with less severe conditions? The National Institute for Health and Clinical Excellence regularly reviews its guidance for cost-effectiveness analysis methodology, taking advice from such expert panels and advisory groups as the Citizens Council (30, 31).

Keeping recommendations up to date in an environment in which clinical evidence, medical technologies, and costs are changing can be a challenge. Clinical guidelines are published with the expectation that they will be reviewed and updated as necessary. Change can sometimes be anticipated, such as the probable reduction in cost of a drug when its patent expires, and the effect can be explored preemptively through sensitivity analysis. To cover most cases, however, NICE has a process by which it determines whether a partial or full guideline update is warranted. The decision to update an economic analysis as part of a guideline update will depend on the priority of that clinical area relative to others being covered, using the previously outlined criteria. It could be that the highest priority for economic modeling for the guideline update is an area that was not informed by economic modeling in the original guideline. It is essential that modeling is undertaken where it will most inform recommendations.

In an ideal world, we could develop a single model for a whole disease pathway from diagnosis, incorporating all the different decision points along the way. Looking at a condition in this holistic manner would help to ensure the whole care pathway recommended in the guideline represents the most cost-effective use of resources. This is ambitious, but the feasibility of such modeling is currently being investigated (32).

A final challenge is estimating the net effect of a treatment for patients with comorbid conditions. Generic health-related quality-of-life instruments (such as the EQ-5D) are useful for this purpose, because they indicate whether a patient’s quality of life has improved overall. This is particularly useful when a treatment is effective for one condition but harmful toward a comorbid condition. In the absence of such data, we try to model all health effects. For example, if we looked at smoking prevention in a coronary prevention guideline, we should ideally estimate its effect on the incidence of cancer and stroke, as well as its effect on coronary heart disease. Sometimes we can conduct cost-effectiveness analyses in subgroups that are differentiated by the presence or absence of various comorbid conditions; we do this if there was trial evidence of a differential treatment effect or if another key parameter varied (for example, the baseline risk for a key event—see case studies 3 to 5). However, data are often insufficient to differentiate between subgroups.

**Challenges to Adapting Cost-Effective Clinical Guidelines**

The guidelines from NICE are produced explicitly for implementation in the United Kingdom, but NICE and the NCGC have been approached by foreign governments and health organizations that are interested in adapting them for use in their own health care systems. Both organizations have experience in sharing this expertise through their international departments. However, the transferability of NICE guidance to another health system requires careful scrutiny of between-country differences in underlying epidemiology and clinical practice (33, 34). Population values and preferences could also vary in ways that render utilities not transferable. The component of a cost-effectiveness analysis that is most likely to differ between countries is the unit costs (such as physician salaries or drug prices). Finally, the budget of the health system affects the cost-effectiveness threshold that should be applied; richer health systems will want to provide things that are currently not recommended in the United Kingdom, whereas poorer systems may not be able to afford recommended practices. Differences in any of these areas may individually or collectively alter the expected clinical and cost-effectiveness of the guideline recommendations.

**Quality Standards**

The coalition government, elected in 2010, has initiated a wave of reform for the NHS, including NICE (35, 36). As a consequence, the NICE clinical guidelines program will be developing quality standards with each clinical guideline. These are statements about what constitutes high-quality care in the NHS, which will be used to commission local health services. This should give health care providers greater incentives to follow clinical practices that have been assessed to be cost-effective.

From the National Clinical Guideline Centre, Royal College of Physicians, London, United Kingdom.
Disclaimer: The views expressed in this publication are those of the authors and not necessarily those of the National Institute for Health and Clinical Excellence.

Acknowledgment: The authors thank Jill Parnham, Maggie Westby, and 3 anonymous reviewers for their useful comments on this paper.

Grant Support: By the National Institute for Health and Clinical Excellence.

Potential Conflicts of Interest: Disclosures can be viewed at www.acponline.org/authors/icmje/ConflictOfInterestForms.do?msNum=M10-2763.

Requests for Single Reprints: David Wonderling, MSc, National Clinical Guideline Centre, Royal College of Physicians, London NW1 4LE, United Kingdom; e-mail, enquiries@ncgc.ac.uk.

Current author addresses and author contributions are available at www.annals.org.

References


Current Author Addresses: Mr. Wonderling, Ms. Sawyer, Ms. Fenu, and Ms. Lovibond: National Clinical Guideline Centre, Royal College of Physicians, London NW1 4LE, United Kingdom. Dr. Laramée: 8 rue Pastourelle, 75003 Paris, France.

Author Contributions: Conception and design: D. Wonderling. Analysis and interpretation of the data: D. Wonderling, L. Sawyer, E. Fenu, K. Lovibond, P. Laramée. Drafting of the article: D. Wonderling, L. Sawyer, E. Fenu, K. Lovibond, P. Laramée. Critical revision of the article for important intellectual content: D. Wonderling, L. Sawyer, E. Fenu, K. Lovibond, P. Laramée. Final approval of the article: D. Wonderling, L. Sawyer, E. Fenu, K. Lovibond, P. Laramée.

Appendix Table 1. Applicability and Quality Assessment of Economic Studies by Using the NICE Economic Evaluation Checklist*

<table>
<thead>
<tr>
<th>Item</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Applicability</td>
<td>An assessment of applicability of the study to the clinical guideline, the current National Health Service situation, and NICE decision making: Directly applicable: The applicability criteria are met, or 1 or more criteria are not met but this is not likely to change the conclusions about cost-effectiveness. Partially applicable: 1 or more of the applicability criteria are not met, and this might change the conclusions about cost-effectiveness. Not applicable: 1 or more of the applicability criteria are not met, and this is likely to change the conclusions about cost-effectiveness.</td>
</tr>
<tr>
<td>Quality</td>
<td>An assessment of the methodological quality of the study: Minor limitations: The study meets all quality criteria, or the study fails to meet 1 or more quality criteria but this is unlikely to change the conclusions about cost-effectiveness. Potentially serious limitations: The study fails to meet 1 or more quality criteria, and this could change the conclusion about cost-effectiveness. Very serious limitations: The study fails to meet 1 or more quality criteria, and this is very likely to change the conclusions about cost-effectiveness. Studies with very serious limitations would usually be excluded from the economic profile table.</td>
</tr>
</tbody>
</table>

NICE = National Institute for Health and Clinical Excellence.
* From reference 6.
<table>
<thead>
<tr>
<th>Guideline Topic</th>
<th>Comparison or Question</th>
<th>Summary of Results</th>
<th>Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alcohol use disorders</td>
<td>Surgical versus endoscopic drainage in patients with obstructive chronic pancreatitis</td>
<td>Clinical data showed that surgical drainage was more effective than endoscopic drainage in terms of pain relief and improved quality of life. The result of the base-case analysis was that surgical drainage of the pancreatic duct dominates endoscopic drainage. The sensitivity analysis showed that the surgical option is cost-effective in every scenario, with the probability of cost-effectiveness for surgery compared with endoscopy exceeding 95% in all analyses at a threshold of £20 000 ($30 000) per QALY gained.</td>
<td>In obstructive chronic pancreatitis, surgical drainage is highly cost-effective compared with endoscopic drainage because of the combination of cost savings and substantial improvements in quality of life. The model also showed that, in terms of QALYs, the large benefits from improved quality of life outweighed the potential risk for death. Surgery was recommended over endoscopic therapy for persons with pain from obstructive, chronic, alcohol-related pancreatitis.</td>
</tr>
<tr>
<td>Chronic obstructive pulmonary disease</td>
<td>LAMA, LABA plus ICS, or triple therapy for initial maintenance therapy in patients with an FEV₁ ≤50%</td>
<td>Clinical studies suggested some potential benefits of triple therapy over other options. Depending on which clinical studies were used as inputs for the model, LAMA or LABA plus ICS was found to be the most cost-effective option. Triple therapy was the most effective option (highest number of QALYs) but was not found to be cost-effective in the primary analyses, with an ICER &gt; £90 000 ($137 000) per QALY gained. Triple therapy was found to be cost-effective in key sensitivity analyses, with an ICER &lt; £15 000 ($23 000) per QALY gained.</td>
<td>The GDG recommended LAMA or LABA plus ICS as first-line treatment in patients with an FEV₁ ≤50%, which reflects the uncertainty in the clinical and cost-effectiveness evidence about which option was superior. The GDG did not recommend triple therapy for first-line treatment but did recommend it as a subsequent option for patients who remain symptomatic after first-line therapy.</td>
</tr>
<tr>
<td>Unstable angina and non–ST-segment elevation MI</td>
<td>Early invasive versus conservative management strategies</td>
<td>Findings were drawn from a published cost-effectiveness analysis undertaken from a United Kingdom National Health Service perspective and based primarily on data from a randomized trial in the United Kingdom. Clinical data supported using an early invasive strategy. Compared with a conservative strategy, an early invasive strategy was found to be increasingly cost-effective as risk for MI and death increased. Cost-effectiveness ratios of £54 000, £23 000, £21 000, £12 000, and £13 000 ($82 000, $33 000, $35 000, $35 000, $32 000; £18 000, and £20 000) per QALY gained were reported for risk groups 1, 2, 3, 4a, and 4b, respectively (1 is lowest and 4b is highest risk for MI or death).</td>
<td>When interpreting the published study, the GDG considered the results of various sensitivity analyses, the likely effect that changes in practice since that randomized trial would have on its results, and the risk profile of the patients included in the trial. An early invasive strategy was recommended for patients at higher risk but not for those at lower risk on the basis of the clinical and cost-effectiveness evidence.</td>
</tr>
<tr>
<td>Venous thromboembolism</td>
<td>LMWH, unfractionated heparin, or fondaparinux versus no prophylaxis in medical patients</td>
<td>The network meta-analysis suggested that LMWH was the most effective at reducing the incidence of deep venous thrombosis but that it increased risk for bleeding. The decision model indicated that LMWH had the highest increase in QALYs compared with no prophylaxis. Compared with other strategies, LMWH was also cost-saving. Results were sensitive to some key assumptions and parameters in the model, especially the baseline risks for venous thromboembolism and major bleeding events.</td>
<td>The GDG recommended LMWH for medical patients but also recommended fondaparinux and unfractionated heparin as options because of uncertainty in the network meta-analysis and decision model. Prophylaxis was not recommended for patients who are expected to recover mobility within 3 days and have no additional risk factors, because they are deemed to be at low risk for venous thromboembolism. Patients at high risk for bleeding should receive mechanical prophylaxis (such as compression stockings) instead of drugs.</td>
</tr>
<tr>
<td>Glaucoma</td>
<td>No treatment or β-blockers versus PGAs for patients with different combinations of risk factors (CCT, IOP, or age)</td>
<td>The clinical evidence showed that baseline risk for glaucoma was determined by age, IOP, and CCT. The results of the model showed that for the general population with ocular hypertension, no treatment is the most cost-effective option. However, treatment with either β-blockers or PGAs was cost-effective for the higher risk groups.</td>
<td>The GDG recommended treatment or no treatment as described in Appendix Table 3.</td>
</tr>
<tr>
<td>Lower urinary tract symptoms in men</td>
<td>Transurethral resection of the prostate versus holmium laser enucleation of prostate</td>
<td>The clinical evidence showed a larger improvement of symptoms (although with wide CIs) and fewer complications with enucleation. In the base-case analysis, enucleation was the dominant strategy. However, a series of 1- and 2-way sensitivity analyses showed that results were very sensitive to several parameters (such as mean difference in symptom improvement, probability of treatment failure for both interventions, and probability of adverse events). At a threshold of £20 000 ($30 000) per QALY, resection was cost-effective in 52% and enucleation in 48% of the simulations.</td>
<td>Although the base case showed enucleation to be more cost-effective, the uncertainty from the sensitivity analyses led the GDG to recommend both interventions. Because of the steep learning curve and high initial purchasing costs associated with the enucleation equipment, the GDG recommended that this intervention be carried out only at centers that specialize in the technique.</td>
</tr>
</tbody>
</table>

CCT = central corneal thickness; GDG = guideline development group; ICER = incremental cost-effectiveness ratio; ICS = inhaled corticosteroids; IOP = intraocular pressure; LABA = long-acting β-agonist; LAMA = long-acting muscarinic agonist; LMWH = low-molecular-weight heparin; MI = myocardial infarction; PGA = prostaglandin analogue; QALY = quality-adjusted life-year.
**Appendix Table 3. Treatment Recommendations for Glaucoma**

<table>
<thead>
<tr>
<th>Intraocular Pressure</th>
<th>Central Corneal Thickness*</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;26 mm Hg</td>
<td>&lt;590 μm</td>
</tr>
<tr>
<td></td>
<td>555 μm to &lt;590 μm</td>
</tr>
<tr>
<td>≥26 mm Hg</td>
<td>&lt;555 μm</td>
</tr>
</tbody>
</table>

- No treatment
- Age <60 y, β-blockers; age >60 y, no treatment
- Age <65 y, PGAs; age >65 y, no treatment
- Age <80 y, PGAs; age >80 y, no treatment

PGA = prostaglandin analogue.

* Rounded to the nearest 5 μm.