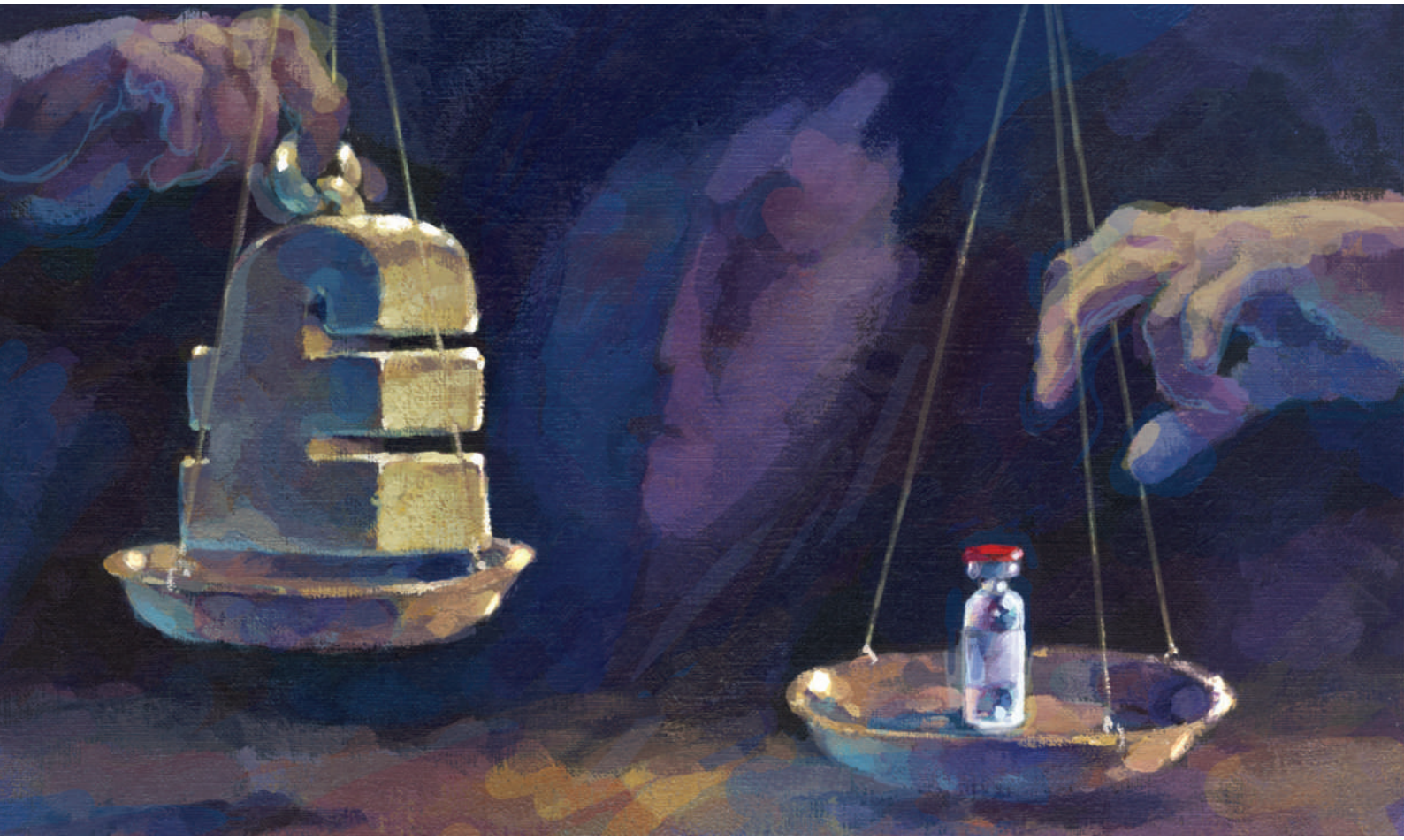

LIFE IN THE BALANCE

How do researchers and policy-makers decide on the value of health? **Daniel Cressey** looks at Britain's National Institute for Health and Clinical Excellence.

In May 2008, a research report concluded that it would be too expensive to keep some British people with kidney cancer alive.

The 290-page review and economic evaluation had been put together to assess the effectiveness of four new treatments — bevacizumab, sorafenib, sunitinib and temsirolimus — for renal-cell carcinoma. The research team concluded that there was some evidence that patients could benefit from these treatments, in some cases even doubling the time that they lived without their disease worsening from five to ten months. However, "the probability that any of these interventions would be considered cost effective" in Britain's nationalized health system, it said, "is zero"[1]. A few months later, an expert panel from Britain's National Institute for Health and Clinical Excellence (NICE) confirmed that prediction, stating in an August decision based heavily on the report that the four drugs "are not recommended as treatment options for advanced and/or metastatic renal cell carcinoma".

The critics erupted. The drug manufacturers and patients were furious. "They are saying to me, as a cancer patient, that I don't deserve the right to live, whether it's for one month, two months, six months or even years," a person with kidney cancer told the BBC. Britain's leading cancer charity Cancer Research UK said that it was "very disappointed" with the decision, and that "NICE needs to consider how it can reconcile making recommendations so clearly at odds with current clinical opinion". "We were not aware at the start that this was going to be controversial," says Jo Thompson Coon, a researcher at the Peninsula Medical School in Exeter, UK, and the corresponding author of the research report.

This is the brutal reality of NICE, the body that decides which medical treatments the nation can afford to buy. To some, NICE is a world-leading body — one at the forefront of 'comparative effectiveness' research, which compares one treatment with another, and exemplary for the tough decisions it makes on cost effectiveness. To others it is a loathed body ready to deny people life-giving treatment with little justification.

Which of these portraits is more accurate has implications for people far beyond the United Kingdom. NICE's decisions and its decision-making methods have already been adopted by several other countries (see 'NICE abroad', overleaf). Some health economists say that the United States — where the debate over health-care reform has reached fever pitch — will have to incorporate techniques similar to those that NICE uses if the country is serious about reining in its more than $2-trillion health-care

ILLUSTRATIONS BY DAVID PARKINS

bill. But to many others in the United States and elsewhere the idea of a NICE-like body is unfathomable. Former Republican vice-presidential candidate Sarah Palin summed up the feelings of many when she warned in a statement in August this year that health-care rationing would lead to a "death panel" for patients in which "bureaucrats can decide … whether they are worthy of health care".

In the United Kingdom, bureaucrats and health-care rationing are a way of life. What is sometimes lost in the midst of the debate is the research itself, which in the case of NICE involves synthesizing evidence and building models that attempt to assign hard values to the quality and cost of each additional month by which a treatment extends life. The researchers who do this work, however, are well aware that their results can have profound effects for the population. "You are obviously always very aware that this is a very important piece of work," says Thompson Coon. "It is not going to sit on a dusty shelf for no one to read. On the other hand, you have to keep it evidence-based."

"What about the people whose health care gets displaced because of the cost?"
— Karl Claxton

Health arbiter

NICE was born in 1999 amid concerns that patients in Britain's National Health Service (NHS), which provides the vast majority of the nation's health care, were not being given the most effective treatments. Established by the Labour government, its purpose is to make sure that the NHS spends its budget — raised by taxation — wisely, using a transparent decision-making process that is based on the best evidence available.

For many conditions, there is scant evidence to show doctors and patients which of two or more medical alternatives — be it treatments, diagnostic techniques or prevention methods — is likely to be more successful. Comparative-effectiveness research aims to find out which one is best, either by pulling together existing research or by commissioning new studies.

On its own the research is relatively uncontroversial. Where NICE gets into hot water is when money is added to the mix. Even if a treatment is more effective clinically than a rival, NICE committee members can decide that the gain in health that it bestows is not worth the additional cost. NICE issues guidance on issues ranging from medical procedures to public health, and the ones that tend to be most controversial are its 'technology appraisals', which tell the NHS under what conditions treatments — particularly new ones — must be offered.

The decision over bevacizumab, sorafenib, sunitinib and temsirolimus was one of the most divisive so far. In 2007, the Department of Health asked NICE to assess the drugs, made by Roche, Bayer, Pfizer and Wyeth, respectively. NICE turned, as it routinely does, to an external assessment group, in this case the group of seven health economists led by Thompson Coon.

The researchers considered the clinical effectiveness question first, namely, what is the effect of the four drugs compared to current standard treatment for renal-cell carcinoma, a particularly nasty type of kidney cancer that is diagnosed in around 6,000 people a year in Britain. One existing treatment option was to remove the tumour entirely by taking out the kidney. If the cancer had spread, or metastasized, then immunotherapy with interferon-α was commonly used. But even with treatment, only 10% of people whose cancer has metastasized survive for at least five years.

The team initially identified 888 potentially relevant studies through electronic searches of various databases. Many of these were excluded for being reviews, reanalyses of the same data or studies that did not meet the researchers' experimental standards, such as being placebo controlled. This left them with 13 relevant papers from 8 clinical studies. For this study "there was actually quite a reasonable amount of evidence", says Thompson Coon. "Because they're new drugs there isn't always that much data available. If you were doing a systematic review of a more established technology you'd probably expect to find more."

In a number of patient groups and disease scenarios, the research team found that the new drugs were better than existing treatments. For example, one of the studies[2] they included showed that people with untreated metastatic renal-cell carcinoma who were given bevacizumab and interferon-α lived without their cancer progressing for just 10.2 months, compared with 5.4 months for those given a placebo and interferon-α.

But a drug that extends life is not much good if the individuals who take it spend that time in crippling pain. So the trick for the assessment team is to assign a somewhat abstract quality of life value — called a 'utility value' — to these additional months or years. Typically, perfect health scores a one. Dead is zero. Putting values in between is the complicated bit.

One way to measure quality of life is using a 'time trade-off' method. Individuals are asked: if you were going to be in a particular health state for ten years, how many years of life would you be prepared to forfeit to be in perfect health? To avoid constant pain, for example, someone might be prepared to give up eight years of life for two in perfect health: this state would be rated 0.2. To avoid a more minor condition a person might only be prepared to forfeit six months, rating 0.95 on the scale. "What's very important is the measure of quality must be based on actual choices that people are willing to make," says Karl Claxton, a health economist at the University of York, UK, who has served on various NICE appraisal committees.

Ideally, the quality of life of a person taking, say, cancer drugs, is measured from patients in a clinical trial. These can be translated

J. SHORT/UPPA/PHOTOSHOT

NICE's refusal to approve spending on new kidney cancer drugs caused outrage among patients.

NICE abroad

The National Institute for Health and Clinical Excellence (NICE) assesses treatments and procedures from a very British viewpoint. Its decisions are informed by NICE's Citizen's Council, which represent the views of the UK population, and the values it assigns to children, elderly people or those who are terminally ill reflect national attitudes and culture.

Yet NICE has influence across the world. Countries including Azerbaijan and Brazil have adapted the institute's guidance. Other countries, such as Australia, have their own NICE-like systems. In 2008, NICE started a non-profit consultancy arm — called NICE International — which does contract work in Canada, Bosnia and other countries looking to use its guidance or to establish similar bodies. "The techniques we use are universal," says Michael Rawlins, chairman of the institute. "Whether a country wants to take them on and in what form is very much related to the culture of the health-care system."

Many health economists think that every country will eventually need some kind of NICE. But the decisions made by the institute are easier to implement in countries that have finite spending on health, such as through a national health service, and therefore a clear basis on which to accept or reject treatments as being affordable or too costly.

The United States has no such limit on health-care spending. The nation's Food and Drug Administration approves a treatment for use if it is deemed to be both safe and effective for the licensed condition — cost is not a factor. The United States did set aside $1.1 billion for comparative-effectiveness research in its economic stimulus package earlier this year. Policy-makers have not yet worked out whether or how to incorporate such research into decisions on cost. "Our job is the research," says Carolyn Clancy, director of the US Agency for Healthcare Research and Quality, the US body that researches health-care quality. "The policy-makers are going to have to struggle with 'how is this research applied'."

"We have had quite a lot of dialogue with the current American administration," Rawlins says. He is unsure how far the United States will go down the clinical and cost-effectiveness road, but he agrees with a view that seems widespread among health economists. US citizens "don't get their money's worth and they can't afford it any longer. They can't go on as they are," he says.

Maynard argues that "the American health-care system rations on the basis of whether people have private insurance or not. Forty-six million people haven't got insurance." He adds that private health insurers ought to adopt assessments of clinical and cost effectiveness to make more rational decisions about what to pay for. "Private insurers are poor purchasers of health care because they don't use health-technology assessment to ration what they put in their benefit package," he says. **D.C.**

into standard utility values by, for example, attaching to them 'reference' measures of time trade-off from the general UK population.

For NICE assessments, research teams then combine the quantity of life that a treatment buys, with the quality of that time into the Quality Adjusted Life Year (QALY), a measure widely used by health economists. This is calculated by multiplying the utility value of a health state by the length of time spent in that state. One year spent in perfect health, for example, gives a QALY of one. Three years spent in a health state with a utility value of 0.5 equals 1.5 QALYs, equivalent to 1.5 years of perfect health.

Complex calculations

For most conditions, working out the QALY is more than a back-of-the-envelope calculation. Clinical trial results cover only the relatively short span of the trial, but researchers are trying to extrapolate from this a QALY for the rest of a patient's life, during which their health may change. For this they turn to mathematical models into which they put all the information on the effectiveness of the drugs from the literature search, the changes the treatments are likely to make to patients' quality of life, how the disease progresses and the timescales involved.

And the answer? The model developed by Thomson Coon's assessment team showed that treating renal-cell carcinoma with interferon-α gives 1.19 QALYs to the average patient. Adding one of the new drugs, bevacizumab, to interferon-α increased this to 1.45 QALYs. Sunitinib was even better, producing 1.62 QALYs. If money were no object, then sunitinib would be the obvious choice.

But for the NHS, money is crucial — and when cost was added the outcome was much less clear-cut. A course of treatment with interferon-α alone cost £8,438 (US$13,786) at the time of the assessment, whereas one with sunitinib was £39,623. This means the additional 0.44 QALYs — effectively just over five extra months of healthy life — costs an extra £31,185. NICE typically calculates the difference as the cost per QALY, which was £71,462 for sunitinib and £171,301 for interferon with bevacizumab.

"NICE probably puts too many things on to the approved list for the NHS, but it's a system that is explicit."
— Alan Maynard

How well the QALY system actually reflects patient preferences is still debated. Stephen Birch, a health economist at McMaster University in Ontario, Canada, says that it fails at this most fundamental of levels. A key problem, as Birch sees it, is the assumption built into calculation of QALYs that it is possible to separate time and health. "These two things are not separable," he says. A person who has been sick for a long time might rate their quality of life differently to someone who has only just become sick. "The QALY doesn't have that," Birch says.

Despite these criticisms and others, many health economists consider the QALY to be the best option available in comparative-effectiveness research. "The argument at a theoretical level is whether the QALY is a good measure of utility or whether it's a crude measure of length and quality of life that doesn't bear close relationships to people's satisfaction," says Alan Maynard, a health economist at the University of York. "It's not strongly theoretically based. It's pragmatic best practice. That's where we're at."

Once calculated, the QALYs and their costs are returned by the assessment team to NICE, where a separate appraisal committee composed of researchers, medical practitioners and laypeople decide what to do with it. The committee produces a draft decision — called an appraisal consultation document — before accepting comments from interested parties and then releasing a final decision.

Generally anything coming in with a cost per QALY gained of under £30,000 is approved for use — and none of the four renal cancer drugs came even close. More than 80% of drugs are approved for use in some form, however (see table). The £30,000 threshold has been criticized for being somewhat arbitrary, but health economists say that it is broadly in line with other spending decisions taken in the NHS, such as those made by health authorities in the absence of NICE guidance. "You can argue that NICE is too generous," says Maynard. "It probably puts too many things on to the approved list for the NHS, but it's a system that is explicit."

An immediate solution for the cancer drugs would be to bring down the cost of the drug. But NICE is not allowed to engage in direct debates with drug companies about the cost of their products, although there are various schemes that allow manufacturers — who are keen to have their drug approved by NICE — to offer 'discounts' to the NHS. Pfizer, for example, offered a deal in which the first cycle of sunitinib would be free to the NHS. NICE's chairman, Michael Rawlins, says it "gets up my nose a bit" when the agency is criticized for rationing when pharmaceutical companies escape criticism for their pricing.

In October, after two of the cancer-drug manufacturers submitted new clinical data or comments, NICE commissioned another review of the evidence, eventually concluding in its final appraisal that the cost of sunitinib "could be less than £50,000 per QALY gained"[3].

Citizens advice

But in the end, redoing the calculations was not what made the difference. The outcry over NICE's decision was such that at the start of January, NICE made a very significant move, effectively rewriting its procedures for drugs such as those for renal cancer. Based in part on the recommendations of its Citizens Council, NICE issued additional guidance for appraising treatments that "may be life-extending for patients with short life expectancy, and which are licensed for indications affecting small numbers of patients with incurable illnesses".

For these drugs, says NICE, appraisal committees can consider "giving greater weight to QALYs achieved in the later stages of terminal diseases". It was effectively telling its committees to be lenient about approving certain treatments that fall above its £30,000 threshold.

In its final appraisal a month later, the committee decided that sunitinib should be used in some circumstances despite its cost. The committee concluded that "in this case there was a significant step-change in treating a disease for which there is currently so little to offer patients"[3]. This August, it confirmed it was rejecting use of the other three drugs.

INSIGHT-VISUAL UK/REX FEATURES

NICE's chairman, Michael Rawlins, says the agency's decisions reflect the desires of the British public.

"The difficulty with these drugs was they were in a disease area where there wasn't much else," says Peter Littlejohns, NICE's clinical and public-health director. "Probably that was an issue that we weren't sensitive enough to at that time. That was why there was a huge backlash." Littlejohns says that the additional end-of-life guidance and the approval of sunitinib "represented NICE responding to scientific and public concern". Rawlins agrees, saying that the guidance change reflects the beliefs and desires of the British public.

The new procedure does not sit well with everyone. "I'm very worried about the notion of weighting different people's QALYs differently," says Claxton. He worries that health gained by those on sunitinib will be health lost by others in the NHS when treatments they will need in future can no longer be afforded. "Of course the people who stand to benefit from a technology might have characteristics that we believe ought to give them a particular social weight, but what about the people whose health care gets displaced because of the cost?" he asks.

As the renal-cancer story illustrates, whatever hard numbers the models spit out, the decisions about what to do with these figures — and hence the outcome of the appraisal process — are influenced by many other social, economic and political factors. Claxton is not the only one who thinks that two groups could well come up with different — and wholly legitimate — answers to the same cost-effectiveness question. Rawlins likens each decision made by an appraisal committee to being tried in court by a jury. "You might get off with one and not with the other," he says. "We try to make sure they're as consistent as possible but at the end of the day there are [scientific] judgements and then of course on top of that these social judgements too."

"Ideologically this isn't very pleasant," sums up Maynard, "but as I tell my medical students, there are two certainties in life, one is death and the other one is scarcity of resources." ■

Daniel Cressey is a reporter based in *Nature*'s London office.

FIVE MOST RECENT NICE TECHNOLOGY APPRAISALS

Drug	Condition	Cost per QALY range	Approved?
Cetuximab	Colorectal cancer	£26,700–33,300	Yes*
Alitretinoin	Eczema	£15,000–31,000	Yes
Bevacizumab, sorafenib and temsirolimus	Renal-cell carcinoma	Lowest estimate: £53,800	No
Tenofovir disoproxil	Chronic hepatitis B	Less than £20,000	Yes
Rituximab	Chronic lymphocytic leukaemia	Probably less than £30,000	Yes†

*In combination with 5-fluorouracil, folinic acid and oxaliplatin. †In combination with fludarabine and cyclophosphamide.

1. Thompson Coon, J. *et al. Bevacizumab, Sorafenib tosylate, Sunitinib and Temsirolimus for Renal Cell Carcinoma: A Systematic Review and Economic Evaluation* (Peninsula Technology Assessment Group, 2008).
2. Escudier, B. *et al. Lancet* **370,** 2103–2111 (2007).
3. National Institute for Health and Clinical Excellence *Final Appraisal Determination: Sunitinib for the First-line Treatment of Advanced and/or Metastatic Renal Cell Carcinoma* (NIH, 2009).

See Editorial, page 315.