INSTITUTE OF HEALTH ECONOMICS
ALBERTA CANADA

PAYING FOR WHAT WORKS

Comparative Effectiveness of Health Technologies and Programs

December 2, 2008, Edmonton, Alberta, Canada

Event Proceedings

IHE Innovation Series
Forum I

About the IHE

The Institute of Health Economics (IHE) is a not-for-profit organization committed to producing, gathering, and disseminating health research findings relating to health economics, health policy, health technology assessment and comparative effectiveness. This work supports and informs efforts to improve public health and develop sustainable health systems. Founded in 1995, the IHE provides services for a range of health-sector stakeholders, and is governed by a Board that includes representatives from government, academia, health-service delivery, and industry organisations:

Industry

Mr. Terry McCool
Vice President, Corporate Affairs, Eli Lilly Canada Inc.

Ms. Patricia Massetti
Vice President Public Affairs and Patient Access, Merck Canada

Dr. Bernard Prigent
Vice President & Medical Director, Pfizer Canada Inc.

Mr. Grant Perry
Vice President Public Affairs and Reimbursement, GlaxoSmithKline Inc.

Mr. William (Bill) Charnetski
Vice President, Global Government Affairs & Public Policy, AstraZeneca

Other

Mr. Doug Gilpin
Chair, IHE Audit and Finance Committee

CEO

Dr. Egon Jonsson
Executive Director and CEO, Institute of Health Economics and Professor, Universities of Alberta and Calgary

Board Secretary

Mr. John Sproule
Senior Policy Director, Institute of Health Economics

Board Treasurer

Ms. Allison Hagen
Director of Finance, Institute of Health Economics

Preface

In December 2008, the Institute of Health Economics launched a series of semi-annual innovation forums whose goal is to bring together senior public and private sector decision-makers to address policy issues of importance in the health care system, not just in Alberta, but to all of Canada and the international community, as well.

Emceed by Andy Greenshaw, Associate Vice President of Research at the University of Alberta, the first session considered the following theme: Paying for What Works— Comparative Effectiveness of Health Technologies and Programs. A variety of perspectives examined established approaches and those under development to compare different technologies, programs, and systems of care.

Lorne Tyrrell, the IHE Board Chair, led things off with a look at the four pillars of health research, focusing on a number of specific advances that have revolutionised the care of certain patients over the past several decades. He noted the importance of nurturing innovation and pointed out that it's not all about being the first to get to market. Money and time are key ingredients to introducing innovative therapies. While there is no doubt that costs can be high, it is also true that introducing new technologies early on can be cost-effective in addition to improving outcomes. Dr. Tyrrell concluded by pointing to the need for a multi-disciplinary approach to healthcare innovation and he emphasised that the full spectrum of related research should be given the attention it deserves. We must be careful not to prejudge what technologies and may or may not lead to success because prediction is a tricky business.

Those opening remarks set the tone for a series of presentations by four eminent experts whose views span both sides of the Canada-US border. Gail Wilensky commented on the importance of getting costs under control and in paying for what works, stressing that what is needed in the US is a centre to study comparative clinical effectiveness. She feels too much emphasis is placed on costs associated with pharmaceuticals, which account for only ten percent of healthcare dollars. The real savings are to be found in areas related to treatment. If interventions are to be covered, payers require evidence that the innovations work better than what already exists. This may be a barrier to the introduction of therapeutic products but it is a necessary part of reigning in skyrocketing health-sector expenditures.

Newell McElwee examined the various decisions that impact the life cycle of a pharmaceutical product, walking the audience through the evidence required to successfully develop a therapeutic product and discussing why there is a need for more consistent and better evidence. He described in detail a process that starts with the investment decision to advance from phase II to phase III, then moves to the regulatory decision—which determines the product's marketability; the third step is the adoption/diffusion decision, which is heavily influenced by the payer. Finally, there come treatment decisions, which have typically been the purview of physicians but are increasingly involving patients as well. All these decisions necessitate the development of different types of evidence, and involve different levels of certainty and divergent requirements associated with their variable stakeholders. Dr. McElwee talked of the need to turn around the concept of evidence-based decision-making and begin first with decision-based evidence-making—that is, base the kind of evidence collected on what is required to inform necessary decisions.

Stuart MacLeod provided a clinician's perspective, noting the importance of balance. While acknowledging the importance of empirical evidence and randomised control trials, he pointed out that such data should not form the entire base upon which to make decisions about effectiveness, which means different things to different stakeholders. Room needs to be made for observational data, as well. Two types of analysis are required to assess the value of introducing innovative technologies, cost-effectiveness and comparative-effectiveness. Each should be conducted separately before being combined to inform decision-making with comprehensive evidence. The challenge is how to balance the results. Some evaluations are dependent on societal values, which can change according to what part of society is involved. This can create dilemmas for decision-makers. When faced with complex situations, leaders

sometimes choose to do nothing—the safe approach. "Best" can be the enemy of "better." Sometimes one should opt for the improvement that is available right now rather than waiting to develop the perfect solution.

Terry Klassen spoke of some of the research conducted by his organisation, Alberta Research Centre for Health Evidence and one of its two key components, the University of Alberta Evidence-Based Practice Centre. He gave examples of how important it is to include results from robust studies when making treatment decisions. He noted that variations in sampling size and other factors can require additional investigation to ensure valid results. He concluded by reiterating the need for solid evidence. He warned, though, that requiring additional studies can be a copout. Ultimately, the decision-making process should involve all relevant stakeholders engaged in a transparent process.

Tom Feasby kicked off the general panel discussion by noting the need to balance evaluations of comparative-effectiveness with assessment of appropriateness. That is, decision-makers must ask, "Does it work," not just, "Can it work?" He cited a couple of studies evaluating the use of long-standing treatments, one for stroke, the other a blood product for immune and infectious diseases. In both instances, it appeared the therapy was misused about half the time, evidence that overall efficacy is a function not just of comparative-effectiveness but also of appropriate application of the technology.

The complete speaker presentations can be found on the IHE website at http://www.ihe.ca/research/innovation-forums/paying-for-what-works-comparative-effectivness-in-healthcare/.

Table of Contents

PAYING FOR WHAT WORKS

Master of Ceremonies: Andy Greenshaw, Associate Vice President (Research), University of Alberta

Welcome and Opening Remarks

ANDY GREENSHAW: Friends, colleagues, and distinguished guests, it is a great pleasure for me to welcome you here today. For those of you who don't know me, my name is Dr. Andy Greenshaw. I am Associate Vice President (Research) at the University of Alberta and I have the privilege and great pleasure of being one of the members of the Board of the Institute for Health Economics (IHE). I have been given the very pleasant task today of introducing the speakers and acting as master of ceremonies for what, I believe, is a very important event, the first of the IHE innovation forums.

IHE board chair, Dr. Lorne Tyrrell, will be speaking in a few moments; but, first, I would like to call on Glenn Monteith, Assistant Deputy Minister, Alberta Health and Wellness, to bring greetings on behalf of the Province of Alberta.

GLENN MONTEITH: As the House is still sitting, the minister asked if I would welcome everybody on his behalf to this very exciting forum. We're looking forward to it. The minister will be attending the dinner this evening, and so will be able to mingle with you after the informative sessions we are going to enjoy today. Again, on behalf of the Province of Alberta, I welcome and thank you for coming on this nice snowy day.

ANDY GREENSHAW: Thank you very much, Glenn. It is wonderful to see so many senior people from the province and distinguished guests from elsewhere.

I now have the great pleasure of introducing Dr. Lorne Tyrrell, Chair of the Board of the Institute for Health Economics, and CIHR/GlaxoSmithKline Chair in Virology at the University of Alberta. Lorne is a distinguished dean and emeritus professor at the University of Alberta and a world-renowned researcher in virology. He has played key roles in the area of health in the broadest sense, many of them at the provincial and national levels, and continues to make a great contribution. Lorne has done an exceptional job of chairing the IHE board, and I would like to take this opportunity to congratulate him and his team on making the institute what it is today. IHE is a tremendous success story, one characterized by innovative and relevant activities. A key word here is 'relevant.' We all hang our hats on excellence in the academy. Infrequently, we are able to hang our hats on relevance, and I think that Lorne does an excellent job of driving the interface between the academy, government, and the private sector for the public good. Please join me in welcoming Lorne Tyrrell.

Innovation: Incremental or Big Leaps?

Lorne Tyrrell, Chair, Institute of Health Economics, and CIHR/GSK Chair in Virology

It is a great pleasure to see the number of people here, many of whom are key leaders in our healthcare system at the government level and in the new Alberta Health Services. It is very important for us to have you here.

If there is some credit that goes to the Institute of Health Economics, our Chief Executive Officer, Egon Jonsson, deserves much of that credit. It is sometimes hard to keep up with the new ideas that he puts forward, and it was his idea that we have this Innovation Forum as the first in a series. I think it will be a very, very good initiative. And, certainly, Senior Policy Director John Sproule has been very innovative as well. It is a good team.

Research has been my passion for a long time, and in introducing this forum, I want to talk briefly about what I see as four pillars of health research. The first is biomedical research, which is the source of discovery. Sometimes we forget how important biomedical research is. Without it, we have no innovation. The second pillar, clinical research, is also extremely important. It is not by accident that the outcomes for heart attack and stroke in centres both here in Edmonton and in Calgary are right up there with the best in the country. It is because of very careful research that has been done to determine how we can improve incrementally the outcomes of these diseases. The third pillar is health services and systems research. We all know the problems in utilization — the overuse, underuse, and misuse — of our healthcare system. The only way that we find solutions to these problems is by conducting research in this area. Finally, the fourth pillar is research in social, cultural, environmental, and population health, and in how we may improve the healthcare system by addressing these other determinants of health.

I would next like to talk about some research innovations and outcomes. I saw the first AIDS patient in Edmonton in 1983, and I cared for AIDS patients all through the 1980s and 90s. At that time, I knew that when I diagnosed AIDS, the patient would die within 11 to 20 months, on average. In 1985, we started treating with AZT [azidothymidine], and a few years later with 3TC [lamivudine]; and the patients were still dying, because one rapidly develops resistance to these drugs. But basic scientists — doing discovery research — were working on the structure of the polymerase of HIV. In 1996, the first protease inhibitor was introduced to the market; and from 1996 to 2008, I never lost another AIDS patient. AIDS is now a chronic disease, not a fatal disease, dramatically changed by innovative discovery research and its translation into clinical research.

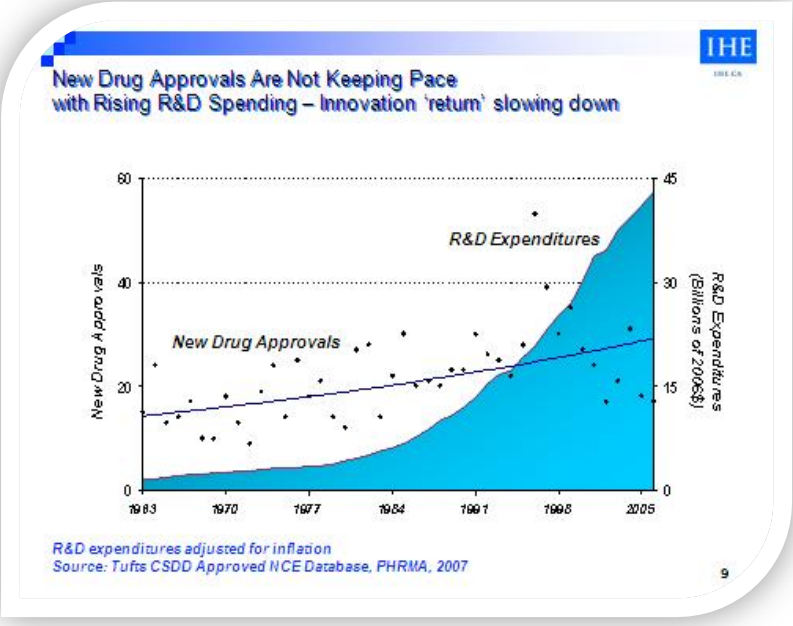
Myelocytic leukemia is another disease that has been dramatically changed by basic research. If you are a follower of the Toronto Maple Leafs, you know that somebody diagnosed with chronic myelocytic leukemia can be placed on Gleevec and not miss a practice or a game, as Jason Blake has done for the last few years while on this drug. Peptic ulcer disease is another example. When I was an intern, vagotomy and pyloroplasty were the most common surgeries. Today, we see less than one percent of the surgery for ulcers that we saw in 1970, because a scientist, Barry Marshall, discovered Helicobacter pylorus and the use of antibiotics to treat ulcers. Some of the newest applications coming from research today are gene therapy to treat congenital forms of blindness and siRNA to treat macular degeneration. Another outcome of discovery research is the dramatic advance we have made in brain and body imaging.

I want to take a second to talk about stem cell research. I sit on board of the Gairdner Foundation in Toronto. It prides itself on picking out Nobel Prize winners before they win the prize: 73 of the recent Nobel Prize winners won a Gairdner Award before they won the Nobel Prize. In 1969, we awarded Till and McCullough the Gairdner prize for their discovery of stem cells. That discovery remained rather quiet for nearly 40 years, but now with the discovery of stem cells in other areas — for example, Sam Weiss in Calgary has discovered stem cells in the neurological system, and John Dick in Toronto has discovered them in cancers — stem cells are going to revolutionize many things that we do in medicine. It took 40 years from the discovery of stem cells for us to see the impact of that discovery, but what a dramatic impact it is going to be.

The road to success is long and costly, no doubt about it: from discovery research (identification of new targets and approaches, the basis of innovation) to applied research; screening systems; cell cultures; animal systems; pharmacokinetics; toxicology studies; product selection; phase I, II, and III trials; and, finally, new products. It is a 12- to 14-year process, and it costs literally hundreds of millions of dollars. In the lifecycle of technology, there is discovery and invention, development, clinical testing, early adoption into clinical practice, and then a phase of widespread adoption of the new and effective therapy, followed by its gradual replacement by the next therapy.

The market in the healthcare system differs from other markets in that there are many intermediaries assessing value. In most markets, the end user is the consumer, the decision maker is the consumer, and the payer is the consumer. In the healthcare system in Canada, the end user is the patient, the decision maker is often the prescriber, or physician, and the payer is a third party. What is the value of innovation in this market? Who determines value, and who determines the benefits to society and to the patient? Not many patients come to the doctor and say, "I want treatment, but which is the best for the sake of society?" Patients usually think about what they need themselves. But sometimes we are faced with making decisions that are best for society, and all of these discussions lead to questions of cost effectiveness.

Drug research and development costs have increased rapidly, but the rate of new drug approvals has not increased very much [see slide 8. "New drug approvals are not keeping pace with rising R&D spending"]. In other words, the new drugs coming to the market are expensive. A frequently asked question is whether we can keep up with these escalating costs. Yet if we look carefully, we see that some pharmaceutical costs are levelling out. The reason for this is that innovative drugs are coming off patents as quickly as new ones are coming on. Generic drugs cost considerably more in Canada than in the United States, however; and if we want to control drug costs, there needs to be discussion about the pricing of generic as well as non-generic drugs. Some of our studies show that early introduction of new therapies often saves rather than costs money, in addition to improving outcomes.



TYRRELL: SLIDE 8

I would like to leave you with four messages about innovation. First, we need to be careful in our early assessments of the future of technologies. Lord Kelvin, president of the Royal Society of London (1890-95) said, "The radio has no future," when it was first discovered. He also said, "Heavier-than-air flying machines are impossible" and "X-rays will prove to be a hoax." With predictive success like this, one wonders how he got into the Royal Society.

The second message is that innovation is often incremental, with occasional unplanned leaps of improvement. Innovation rarely involves a single giant step, and you cannot simply order up breakthroughs. Innovation is a fragile, slow, and unpredictable process. There are many false starts and dead ends. The first product that reaches the market often has inadequacies and is gradually replaced by superior products. Our assessment process must recognize some of these difficulties.

The third message is that we are living in a period of interdisciplinary science and of collaboration across the public–private sectors. This multidisciplinary and intersectoral cooperation is going to be key in addressing many issues in the future. We strongly believe in this in the Institute of Health Economics, and this is why we have brought together government, academia, and industry.

Systems biology and engineering are critical; and the convergence of mathematics, chemistry, biology, physics, computing science, medicine, and engineering has never before provided us with the opportunities that we have today. I believe that some of the research now being applied to disease models will help us to find solutions that we did not dream possible just a few years ago. Image analysis, predictive modelling of biological systems, interdisciplinary approaches, and clinical and regulatory support are all key factors.

The fourth message is that implementing new technologies is not a simple yes/no decision. Certain technologies are beneficial for some populations, but not for others. The new field of pharmaco-genomics will have a major influence in this area. We need to examine the use of technology along the entire care pathway and to develop ways to ensure early coverage with evidence, development, and what we are talking about today — analysis of the comparative effectiveness of different approaches.

Here are some recommendations for getting started:

- Support opportunities for dialogue between funders, innovators, and practitioners.
- Fund and support discovery research. Over 70 percent of the new patents are based on discoveries in university research.
- Provide resources for knowledge and evidence transfer.
- Develop a capacity to review the evidence for new as well as established technologies.
- Focus attention on conditions that account for most of the healthcare costs: cancer, heart disease, infectious diseases, neurodegenerative diseases, and mental illness.
- Focus attention on high-quality research across the full spectrum of research.

Finally, I would never give a talk without at least once mentioning hepatitis. As we talk about focusing our research, I want to point out that Health Canada doubled the HIV budget — and I am not arguing against that — at the same time that they cancelled the hepatitis C budget. There happen to be about 250,000 hepatitis C patients in Canada. There are 45,000 HIV patients. There are 24 drugs now for HIV. There are still only ribavirin and interferon for hepatitis C. Hepatitis B is also a major disease. Worldwide, there are 33 million people with HIV and 600 million with hepatitis B and C. I simply wish that we could get the same attention for these other diseases that we have for HIV. It shows how important the lobby is.

Dreamers, innovators, and entrepreneurs are needed. As George Bernard Shaw said, "The future belongs to the unreasonable ones, the ones who look forward, not backward, who are certain only of uncertainty, and who have the ability and confidence to think completely differently." Some of our solutions will come from Shaw's philosophy.

I thank you very much and look forward to hearing the talks of today.

ANDY GREENSHAW: Thank you very much, Lorne. As usual, some thought-provoking comments and a good stimulus to begin this forum.

I now have the great pleasure of introducing one of our guest keynote speakers. We are deeply grateful that Dr. Gail Wilensky has taken time out of her very busy schedule to come to Edmonton and participate in this event. Dr. Wilensky is one of the pioneers behind the movement of comparative effectiveness research and the use of evidence to help guide decision making in health care. She is an economist and a senior fellow at the International Health Education Foundation, Project HOPE, and a commissioner on the World Health Organization's Commission on the Social Determinants of Health. Gail is an elected member of the Institute of Medicine and has served on its governing council. She is the former chair of Academy Health and has had numerous very prestigious appointments, including Deputy Assistant for Policy Development to the President of the United States of America.

Gail, welcome to Edmonton and to Alberta. We look forward very much to your remarks about comparative effectiveness in health care.

Comparative Effectiveness in Health Care: Operationalizing Evidence-based Medicine

Gail Wilensky, Economist and Senior Fellow, Project HOPE

Good afternoon. It's a pleasure to be here. When Egon Jonsson called to ask if I would consider speaking at this forum, I hesitated only to say that I had to make sure that the airlines would be cooperative so that I could keep my previous commitments in London and in Washington, but that otherwise I would very much like to do so. There are two reasons for that. First, comparative effectiveness research is a topic for which I have enormous passion because I think it is one of the means by which we can make real progress in improving our healthcare system; and, second, I was fascinated by the concept of this institute, which not only convenes a meeting with representatives from industry, academia, and government, but has representatives from each of these three areas on its governing board. This is quite unusual in my experience. I am delighted to be here and to have an opportunity to talk about operationalizing evidence-based medicine by using comparative clinical effectiveness.

In the United States and most other parts of the developed world, we have a common set of problems in health care. Of course, in the United States we have one problem that most of you do not have, which is that 15 percent of our population is without health insurance. But aside from that, we have a set of problems in common with the rest of the developed world. One of these is unsustainable growth in healthcare spending. In the US, spending begins at a much higher base level than in other countries, but our rate of growth has not been very different from that in other parts of the Western world. We also have problems with patient safety, although this is not very well documented, as you have heard. The Institute of Medicine has estimated that as many as

100,000 patients in the United States die each year from medical errors. This is not a precise number, but the problem is well recognized. And we have many problems with quality and clinical appropriateness. Some of the studies in this area — and I am thinking in particular of a study by Elizabeth McGlynn from the RAND Corporation — indicate that the likelihood of receiving all clinically appropriate interventions when you have an encounter with the healthcare system in the United States is about 53 percent. This puts it only slightly above a crap shoot, so this is a serious issue. This is not unique to the United States, but it is particularly frustrating — galling even, given how much we spend on health care — that we still find ourselves with these problems.

As important as it is to find a way to bring in the 15 percent of our population without insurance coverage, it is even more important to find ways to slow the growth of spending and to improve the value we receive for what we spend on health care. There are several reasons for that. In the United States, the single biggest predictor of an increase in the number of people without insurance coverage is increased healthcare spending. We need to find a way to slow the growth of spending. I am not talking about reducing absolute spending, but rather slowing what has historically in the United States been a growth rate in per capita spending on health care that is 2 to 2.5 percent faster than growth in the rest of the economy. If we can find a way to slow that and to improve the value we receive for the money we are spending, it will be much easier to expand insurance coverage.

Increases in healthcare spending put huge pressure on employers, who are the primary sponsors of health insurance. As healthcare spending increases, there is less money available for increasing wages. Occasionally, you will hear people say that premiums are growing even faster than wages, and economists like me want to say, "Well, of course." Healthcare insurance is the other part of the compensation package. If one part is growing faster, then by definition the other part will grow more slowly. Healthcare spending is also putting growing pressure on the federal budget, and this limits investment in other areas. We will continue to experience this if we cannot find a way to slow spending and, in particular, spending on our medical entitlements, Medicare and Medicaid. Because they tend to track each other, it is very hard to imagine slowing the spending on Medicare without slowing healthcare spending in general, especially as the number of people on Medicare increases in the next 20 years.

We already know some things that give us reason to hope that we can find ways to "spend smarter" on health care. For example, we know that there are huge variations in the way that care is provided and in the amount spent on health care in various parts of the United States. This holds true in every country that has been examined, irrespective of the manner in which health care is financed. It is true in the UK. It is true in Norway. I haven't seen the statistics, but I am quite confident that it is true in Canada as well.

The second thing we know is that spending more on health care is not the same as increasing quality. We have had, for some 30 years, work by Jack Wennberg and others at Dartmouth that showed very substantial variations in per capita spending on Medicare across the country; but until recently all we knew was that spending varied, which is not the same as knowing what is the most appropriate amount to spend. It was not clear whether the mean or the median was the right amount, or whether the high spending or the low spending was the appropriate amount. We knew only that there were very substantial differences. In more recent work by Elliott Fisher and others in the Dartmouth group, as well as by other institutions, there do not appear to be comparable health gains, either clinically or even in terms of patient preferences, in the high spending areas. This is hard to chip away at. You have to look for people who have specific medical diagnoses or who are in the last six months of life in order to feel confident that you are looking at comparably ill individuals. But as work continues in this area, it does add some fuel to the notion that clinical outcomes and responsiveness to patients in the high-spending areas do not in any way justify that high spending.

We also know that spending growth, historically, has been related in significant part to technology growth. I do not want to suggest that I am against the introduction of new technology. The older I get, the more frequently I am inclined to find reasons to applaud new technology. For instance, I recently managed to tear my ACL [anterior cruciate ligament] playing tennis, and had it repaired arthroscopically, with only two very small incisions in my left knee. I was on an international trip about six or seven weeks later. My daughter managed, unfortunately for her, to tear her ACL in the 1990s and had hers repaired with an arthrotomy, an open-knee procedure. It was about six months before she was able to get around even quasi-normally. So I don't want to imply that I am against technology, but we clearly need to learn how to use technology appropriately, in ways that help individuals who can be helped, and to spend smarter as we do so. And because so much of our spending growth has to do with the increase in chronic diseases, such as diabetes, congestive heart failure, and obesity-related conditions, we need to learn not only how to spend smarter, but how to treat smarter as well.

If we are going to do that, there are many things that we have to do differently. We need to be able to measure better. We need a scorecard. We need to have measurements for quality, efficiency, and patient-centeredness. Professor Ara Darzi, as part of the 60th anniversary of the National Health Service in the UK, recently completed a parliamentary study looking at improving quality. One of their activities in the area of quality accounting will be to publish, for each of the provider groups, examples of how to improve measurements, to make information public, and to ensure that the quality and efficiency of care provided is clear.

We also need better information, and, of course, that is in large part the issue in comparative clinical effectiveness. In the US, we are starting to see a new interest, or renewed interest, in making better information available. You can go online to Hospital Compare and check the hospital to which you are likely to be admitted, in order to find out how well they do on major procedures, how often they follow clinical appropriateness guidelines, their mortality rates, and a variety of other measures. Some information is starting to be collected even for physicians. I say 'even,' because that is a group, particularly those practicing outside of hospitals, which has quite successfully resisted data collection for a long time. Information on some of the HMOs [health maintenance organizations] and hospitals has been available, but there has been very little on physicians. That started to change in July 2007 when an add-on payment was provided to physicians who reported on a set of measures that had been agreed to. The Joint Commission on Accreditation of Healthcare Organizations (JCAHO) also has a publicly reported quality-check program.

As important as it is to have better information, it will not be enough only to have better information. We also need to change the incentives. As someone who ran the Medicare and Medicaid programs in the first Bush administration and has advised Congress on two different Medicare commissions, I feel that I am as entitled as anyone to say that Medicare has gotten it exactly wrong for the last 25 years, taking great pride in paying exactly the same amount for the best in class and the worst in class. It is hardly a way to improve quality. In some areas, with the bundling of payment, such as the use of a DRG [diagnosis-related group] for a hospital discharge or an episode of home care, it is possible for efficient providers to gain efficiency. There are other areas — for example, in the physicians' fee schedule, which has 7,000 individual billing codes — where it is not possible for efficient physicians to gain efficiency, either because they are of high calibre or because they practice in a conservative style. In fact, it is quite the opposite. Those that are not increasing the volume and mix of the services they provide are seeing their fees held constant while their costs are increasing. At the same time, spending on physician services under Medicare has been increasing at 12 or 13 percent per year, which is untenable. The private sector, frankly, has not been much better: some interesting experimentation has been done in the area of pay for performance or reward for improving quality, but this is still very much at the margin.

But while it is important to have better information and to improve incentives, these measures really do not get to what I think is needed to operationalize evidence-based medicine. This is why I find the renewed interest — and, in the US, the new interest — in comparative clinical effectiveness so important. Comparative clinical effectiveness research is not likely, in and of itself, to drive the changes that we need; but it is a building block. It is an enabler and, without it, we are unlikely to fix these problems.

When I speak of comparative clinical effectiveness research, I am talking about gathering information on what works, when, and for whom, among the interventions provided in specialized facilities or other duly licensed institutions. It is rare that an intervention is always effective or never effective. The issue is recognizing that what may be a very effective intervention for one subset of the population with a certain combination of morbidities will not be effective with another group that looks somewhat like them, but is not exactly the same. This is a particularly important issue in the US, where we are so aggressive in our interventions, but it is also important in other countries where there are substantial variations in spending.

If we are going to compare the clinical effectiveness of interventions, we need to think about what we should be looking at. While I am impressed by and a little envious of the progress of other countries engaged in this area, I do not think that they have quite gotten the focus right. In general — although this is not as true now as in the past — the focus is mostly on comparing one pharmaceutical with another pharmaceutical. It may on occasion be on comparing a device with a pharmaceutical. But it misses what we need to look at. In the US, pharmaceuticals account for only about 10 percent of the healthcare dollar. If we are to learn how to spend smarter, we need look at where the money is being spent. We need to look at ways to treat a particular medical condition rather than focusing on a particular therapy or intervention, and to look at medical procedures rather than drugs and devices. I don't want to spend whatever political capital I have trying to drive change for something that accounts for 10 or 20 percent of the healthcare dollar. If we don't look at medical procedures, we miss where the money is.

Looking at medical procedures is important in learning how to spend smarter, but it is also obviously important in learning how to treat better. We need to invest not only in better understanding of what is already known — better systematic reviews — but in what is not yet known. And, yes, that will mean new types of data collection. It is not enough to make use of the data we have; it is a dynamic process. We cannot make an investment once and then assume that we are forever done with that.

The question of what kind of data we should use is very interesting. There is a tendency to place data in a hierarchy in which the randomized clinical trial is regarded as the gold standard and everything else is deemed not nearly as acceptable. But the fact is that all information has flaws. Even in randomized controlled trials, double blinded or not, decisions about the mix of patients studied may not reflect the co-morbidities of those patients. Most information has some strengths, and we need to include data from a variety of sources and to recognize the strengths of each. We need to find more economical ways of conducting prospective trials. There is much interest lately in using Bayesian statistics as a means of shortening the time and allowing for mid-course corrections, but there are other data sources that can be considered. We can look at epidemiological studies, at medical records, and at registries, recognizing that it is important to correct for statistical bias in these various sources, but also that all data have limitations. The question is how large the differences in data are, and how confident we are that we know how to adjust for the biases that are likely present. If this is going to work, the data produced must meet certain criteria. The most important of these, I believe, are that data be objective, credible, timely, transparent, and understandable.

Balancing these is difficult and leads to different ideas about where an agency or centre focusing on comparative clinical effectiveness should be placed. I think it will doom an entity associated with comparative clinical effectiveness if we are not very careful and politically circumspect about how we establish it. The first thing we need to do is

generate the kind of information that can be used ultimately by patients and their clinicians as well as by payers; and that means having a safe and protected place where objective, credible, timely information is generated and made available for use. I believe that what we will see in the US is a comparative clinical effectiveness (CCE) research institute or centre, either governmental or quasi-governmental, which will provide information so that payers can make better decisions, but will not make decisions about coverage or reimbursement. In the US, it would be a big mistake for this be a decision-making centre, although that is much less of a concern in jurisdictions where there is a single payer who has a clear role in decision making.

In the US, it will also be very important not to bring cost-effectiveness into the CCE centre itself. This is where I part company with many of my fellow advocates. It is not that I don't think cost-effectiveness is relevant. I do. But having the information on comparative clinical effectiveness generated as pristinely as possible is the most important step, and it will require more time and money than studying cost-effectiveness. Payers can fund cost-effectiveness studies on their own, hopefully in a transparent way; but they ought to be done separately, at least for an initial period while the CCE centre is being established.

As I have indicated, I do not think that cost-effectiveness is irrelevant. However, if you are thinking about whether or not to reimburse a new therapy, pharmaceutical, or device, the first question is whether it does more. If it does not do more, I can't imagine why you would pay more. And if it does do more, how much more does it do, and how much more does it cost? It is appropriate that these discussions be part of the decision making of the payer, whether that is Medicare, Medicaid, or private payers. However, without basic comparative clinical effectiveness information, cost-effectiveness cannot be determined. As it happens, US Medicare does not have the authority to consider cost as an element, either in coverage or in reimbursement decisions. This is something that needs to change if we are going to make better use of our spending.

Spending smarter also means having better incentives. If we want to encourage institutions, clinicians, and physicians to behave differently, we need to reward those who provide effective and efficient care that is clinically appropriate. We also need to think about value-based purchasing and value-based insurance, whereby the co-payment varies not with the cost of the pharmaceutical, but with its clinical appropriateness. The lowest co-payments would be for those pharmaceuticals that are most clinically appropriate. In other words, we would not say no to coverage, but would make treatments that are of questionable clinical appropriateness more expensive. We should also find ways to reward people who have healthy lifestyles. Small steps have been taken in this direction in the US, whereby some county governments have talked about having higher premiums for people who smoke but making smoking cessation aids available to those who want to stop.

What does this mean for industry? It potentially raises the bar for increasing reimbursement. That is, if you want to be paid more, your products are going to have to do more; it is not enough that they be new. For people in industry, this is an area that, along with transparency, raises concerns about delays in getting pharmaceuticals or devices to market. Comparative clinical effectiveness research need not delay entry to market. If the industry is willing to risk a portion of an increase in payment while information on comparative clinical effectiveness is being collected, the payer and industry then share the risk that the new product may not offer the improvement that is claimed. For some small biotech firms, there might be a clear second level of decision making after three years, at the point when the information is in for most of these small companies. While the medical community is only beginning to think about whether or not to support this, there is a great deal of interest among political parties in the United States.

Industry support has been somewhat mixed. In general, big pharmaceutical companies have been quite supportive of comparative clinical effectiveness, perhaps because it is not so different from what goes on in Europe and Australia. Some of the small pharmaceutical and biotech companies are very worried about any additional delay. Some of the

device companies have been extremely nervous about whether their small incremental improvements will be incorporated in CCE. This is a fair enough concern. The answer is to figure out how to accommodate those types of improvements. Some of the physician groups to whom I have been speaking about this issue are beginning to be supportive. It will be very important if the academic community and some of the major associations come out in support.

Just before the US Congress recessed in August, a bill was introduced by two significant players in the senate: Max Baucus, Chairman of the Senate Finance Committee, which is where healthcare reform at large is likely to be, and Kent Conrad, head of the Senate Budget Committee. It is a pretty good bill and likely to be a starting point for whatever happens next. Both candidates were supportive of comparative clinical effectiveness, Obama more explicitly so than McCain; and, since Obama is the President-elect, there is a reasonable chance that comparative clinical effectiveness will be in whatever legislation emerges in 2009. There will be some legislation, and whether there will be changes in Medicare, such as expansion in coverage or reimbursement, I am not sure, but very likely comparative clinical effectiveness will be part of the package. This issue seems to draw a great deal of support across the chambers and across the parties as a recognition that we need to do things that will allow us to slow spending and improve our value for health care. This has got to be the basis for any sensible way of doing that, so I think it just could happen. At least I am in that cautiously optimistic stage as we end 2008.

Questions and Answers — Gail Wilensky

Stirling Bryan: I agree with much of what you are said, but have to take issue with the manner in which cost is to be dealt with. Much of the challenge that we have in relation to new technologies is that while we have some gains in comparative effectiveness, these gains are potentially quite marginal in many cases, but the increase in cost is enormous. If we focus on comparative effectiveness, we will see some positive messages coming through; but if we focus on cost-effectiveness, then we get a much more balanced picture of the gains.

I understand that in the US it may, politically, make more sense to go with comparative effectiveness first and then to tackle the cost-effectiveness issue. But I wonder whether, elsewhere, we should use the example of NICE in the UK, which has been very explicit about cost-effectiveness as the criterion — not effectiveness, but cost-effectiveness; and I think there has been a receptivity among the public to acknowledge that there are constraints and that we have to make difficult decisions. Possibly the US may not be the place that we should look if we are taking this issue forward in a Canadian context.

Gail Wilensky: There are two reasons that I think they should be kept separate — but separate is not the same as not making use of comparative effectiveness information. Partly, it is a political issue. Without the generation of comparative data on clinical effectiveness, which requires a time-consuming, expensive analysis, you don't have any information on which to base your cost-effectiveness analysis. That data therefore needs to be protected. Now, to be very clear, I am not talking only about comparing the use of one statin versus another or a fourth-generation antibiotic. I am talking about much more fundamental questions, such as what are the comparative gains from the medical treatment of cardiovascular disease versus angioplasty with or without drug-eluting stents versus bypass surgery. Only a payer, certainly in the US, would be willing to invest the millions of dollars it would take to answer that kind of big question.

As an example, about a year ago, there was an article in the New York Times about ablation therapy for atrial fibrillation. This therapy is very popular among cardiologists, and the article quoted a number of cardiologists from the west coast who talked about their success rates with patients. But about three-quarters of the way through the article was a comment that we have little indication of how patients three or four years out fare relative to those who

stay on warfarin instead of having vessels ablated. An academic cardiologist said that, actually, the real question is how we even define success, given that the ablated vessels are no longer a problem but problems frequently develop in other vessels. And, by the way, the patients have to stay on warfarin anyway, so what exactly is it that you have accomplished? This is the kind of issue that I want to see tackled. In the US, we tend to choose very aggressive interventions without comparable clinical outcome data to support that. Trying to understand who is actually likely to gain clinically and who is not is the first step. This does not dispute the importance of cost-effectiveness.

A second reason why I, as an economist, would like to keep clinical effectiveness separate from cost-effectiveness is that there is too much about cost-effectiveness that is a little vague. What one is willing to define as a cost is very important, and there is a tendency to talk about cost to society. Cost to society is an excuse for demanding a very low discounted rate, and the rate of reimbursement can determine the feasibility of developing anything that takes more than a year or two to bring to market. There are enough areas of legitimate dispute in cost effectiveness that I would rather it be the responsibility of the payer. In places that have single payers, doing what is both clinically effective and cost-effective (of the greatest value) has a very definite payoff. It is much easier in the UK for NICE to make the argument that if we do things that are questionable, we are not doing things that are of more value, unless — as they have for the last decade — we get the government to kick in more money. In the US, we have a very open system. It is not that we don't have the same kind of pressures, but they are much less obvious.

I don't think cost-effectiveness is a fair and legitimate measure to use in deciding how much to reimburse for new technology. Comparative clinical effectiveness research is critical, and without it, you don't have information with which to determine cost effectiveness. In the United States, we are seeing a threefold difference in spending. Something is going on. There is some dispute or uncertainty about how to treat people with the same medical condition, and there are no obvious health outcomes associated with those differences. If you are spending a lot of money on treating a particular medical condition and there is huge variation in expenditure, there is clearly a lot of dispute about how to appropriately treat patients with that medical condition. That is an area just screaming for comparative clinical effectiveness analysis.

GREG SABLE, MERC: One of your slides spoke about doing analysis of "what works when, for whom, provided by…." One thing that I have seen increasingly in the last couple of years, at least in Canada, is what I will call a narrower and narrower focus in the application of evidence-based medicine. I am wondering if you have any advice about how narrow we ought to be. My fear is that we are almost at a place where we are looking for something that is cost-effective for left-handed redheads between the ages of 64 and 68. Of course, we don't have that kind of data, but the dialogue has gotten almost that minute in some circumstances.

GAIL WILENSKY: I have two responses. First, we need to stop focusing on the cost of drugs and devices. That ought to be music to the ears of pharmaceutical and device companies, but I am not an apologist for the industry. I just want to go to where the money is, because that is how you moderate spending; and that means you cannot ignore procedures. Physician and the hospital costs are where the money is, and when you see huge variations in spending, that's the reason. So the first thing is to stop focusing on the little ants around you and start looking at the big picture.

The second is to ask if there is any reason to believe that these various subgroups actually differ. Now, they may. Randomized clinical trials tend to be very pristine in narrowing the group of participants they will consider, but people typically don't present themselves that way. You can open a trial to white males between the ages of 21 and 54 who have, for example, stage-3 congestive heart failure. But if most of the people who actually show up have congestive heart failure and diabetes or congestive heart failure and COPD; and if there is some reason to think that co-morbidities may affect the outcome — which doesn't seem a huge stretch — then there is no need to focus so narrowly on gender, age, or race.

Another question that frequently gets raised has to do with whether or not differences in outcome are likely to be major if treatment is given in specialized clinics. For example, in an academic health centre that does many of whatever procedure you are looking at, are you likely to see significantly different results than in a community hospital that has appropriately accredited practitioners? You need to know whether or not the results you are seeing are highly sensitive to where the treatment is provided, because it will tell you the likely outcome.

The aim is not to stop people from receiving medical care that is likely to offer them clinical gains. It is to have a better idea of whether or not it is possible to differentiate, among the large numbers of people who are receiving the care, some who are more likely than others to have a significant clinical improvement. Given that we see very different spending rates across the world and within countries, without any underlying clinical factors, we ought to be finding out what is going on in areas of low spending or low intervention and what can we learn from them.

Panel Discussion—Investing in Innovation: Impact of Comparative Effectiveness

ANDY GREENSHAW: The next part of the program is a series of panel discussions. Each of the three speakers will make a presentation for around 20 minutes, and then we will then have a general panel discussion.

Our first speaker is Dr. Newell McElwee, Vice President for Evidence-based Strategies for Pfizer Pharmaceuticals, where he leads a group of scientists focused on evidence-based medicine, health technology assessment, and the application of evidence to coverage decisions. He is a faculty member at Tufts University School of Medicine in Boston, serves on many advisory boards, including the Agency For Health Care Research and Quality Effective Health Care Steering Committee, and Duke University Health Sector Advisory Council. Please join me in welcoming Newell.

Paying for What Works: Comparative Effectiveness in Healthcare

Newell McElwee, Vice President, Evidence-Based Strategies, Medical Division, Pfizer Inc.

Thank you for inviting me. It is a privilege for me to be here in Edmonton, and before I get started, I would just like to say that Gail Wilensky is actually a national treasure for us in the US It has been about two years almost to the date since she published her article on comparative effectiveness and health affairs. From my perspective, the importance of her paper was not only its content, because I think many people had thought it and some people had said it before, but also who said it. Because it was Gail, I think a lot of people took notice that this was something important.

Why all the fuss about comparative effectiveness? Effectiveness is already a part of the health technology assessments (HTAs) that we do. However, it is well recognized that evidence to inform adoption and treatment decisions is sometimes lacking. Often, the evidence that we have is insufficient, whether it comes from the evidence-based practice centres (EPCs) in the US and Canada or from the Oregon Drug Effectiveness Review Project (DERP), or from wherever. A second situation that we often encounter is that the evidence we have is the wrong evidence: it is evidence that was developed for a different research question and for a different decision. Therefore, a key driver is recognition of the need for better evidence and for evidence that is tailored to the

decision. That is mostly what I am going to talk about. Perhaps coincidentally, a recent Institute of Medicine report called "What Works" addresses this very topic.

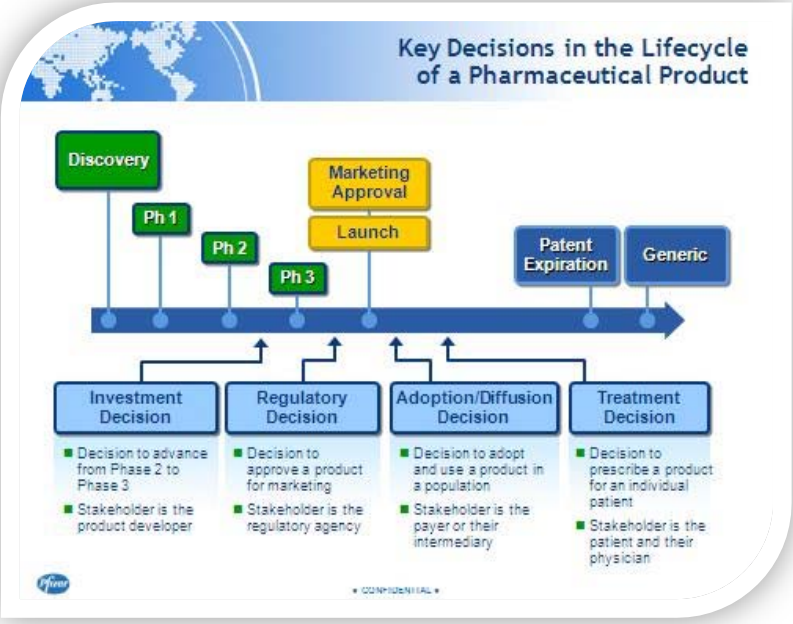
From my perspective, strategies to develop evidence must do two things. They must ensure that evidence generation is linked to the decision, and they must balance the need for cost control with society's desire for new technologies. If we had to do only one of those things — control costs or develop new technologies — this would be easy, but there is a tension when you have to do both. Everybody talks about evidence-based decision making, but what we really should be discussing is decision-based evidence making.

There are some additional questions about evidence that I want to highlight. First, what if we had perfect evidence? Would we have perfect decisions? And if we didn't have perfect decisions, would the decisions at least be predictable? Gail referred to the work that Beth McGlynn has done showing that we do all of the right things only about 55 percent of the time when we have the evidence. Clearly, therefore, having the evidence is not the sole answer. Another question is whether or not better evidence can improve quality and save money. I would say that it is possible, but my concern is that people put too much faith in comparative effectiveness research as a cost-saving strategy. Estimates of cost savings from the Congressional Budget Office (CBO) and the Commonwealth Fund in the US were quite divergent, with the CBO estimating $10 billion in savings over 10 years and the Commonwealth Fund estimating $368 billion over 10 years. The reason for the difference, which Gail referred to, is that comparative effectiveness research is necessary but probably not sufficient. We also need payment reform. The difference in these two estimates is that the Commonwealth Fund includes payment reform in their methods.

My goals for today are to propose a framework for evidence generation, assessment, and appraisal based on key decisions made during the lifecycle of a healthcare technology, and to highlight some areas for improving the transparency and predictability of decisions. Before we get started, I want to refer back to some work that David Eddy did almost 20 years ago in which he talked about the anatomy of a decision. There are two components of a decision, a scientific-analysis component and a value-judgment component. One is left brain, one is right brain; or, if you live in the HTA world, one is assessment and one is appraisal. Analysis of the evidence is what you hear when you go to scientific meetings. An example is what I refer to as the "quality folly," which is debate about whether quality is useful in cost-effectiveness analysis. A former Nobel Prize winner is now arguing that the utility assessment takes only a cross-sectional look at utilities and does not account for patients that adapt or accommodate to their disease. There is huge debate about that right now. Another subject of much debate is the method used for analyzing observational data, whether it is propensity scores or instrument variables or whatever.

The analysis is the easy part, however. The hard part is the appraisal, or the making of value judgments that inform decisions. Little work has been done on this. There is no training or framework for people on appraisal committees and there are no standards, so this is an area ripe for additional work. We have recently finished a study in the UK looking at societal preferences, which should play a role in appraisal. All things being equal, UK residents cannot distinguish between treating a 40-year-old and a 60-year-old, but they do distinguish between treating a child and an adult. They also distinguish between causes of disease, so that a genetic disease or a disease that they consider self-inflicted, such as obesity, they prioritize lower than a disease that is inflicted by others, such as iatrogenic MRSA [Methicillin-resistant Staphylococcus aureus] infection. There is much room for additional work in appraisal.

The lifecycle of a pharmaceutical product starts with discovery and animal studies, and then phase I, phase II, phase III, marketing approval and launch, adoption and diffusion, patent expiration, and then going generic [see slide 6, "Key decisions in the lifecycle of a pharmaceutical product"]. I want to talk about four key decisions among the many decisions that occur during this lifecycle. The first is the investment decision to advance from phase II to phase III, which is probably the most critical decision that pharmaceutical companies make for investing in innovation. The stakeholder in this decision is the product developer. The next decision is the regulatory decision, or the decision to approve a product for marketing. The stakeholder is the regulatory agency that makes that approval decision. The third decision is the adoption/diffusion decision, or the decision to adopt and use a product in a population. The stakeholder is the payer or, in the US, their intermediary. Last are the treatment decisions by patients and their physicians, the decisions to prescribe products for individual patients. The stakeholders here are patients and their physicians. We will go through each of these in a little more depth, because each requires thinking about evidence a little bit differently and asks different questions.

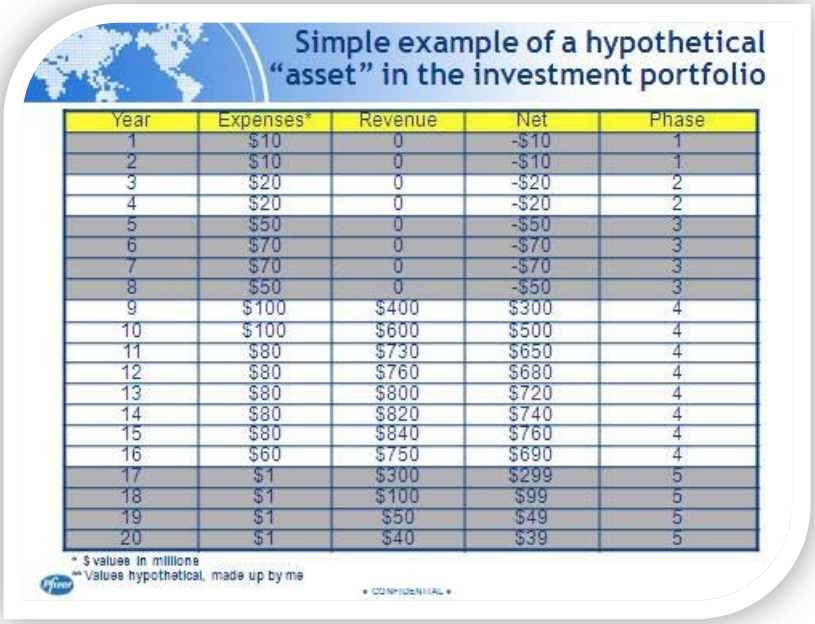


McElwee: Slide 6

Investment decision: Investment decisions are, I think, widely misunderstood outside the industry. The techniques used for making investment decisions are largely from the financial community. I would use the analogy of stock investments. You are looking at five stocks, but you can invest in only two of them. Which two do you pick? Depending on events in the marketplace, what gets ranked first and second may change. Decisions are based on the opportunity cost for the portfolio and on several different metrics, the most common one being the expected net present value. Historically, going back five or ten years, the expected net present value was based mostly on the probability of technical and regulatory success. The industry did not omit information about adoption and treatment use, but these were very crude estimates. In the past few years, there has been an increasing emphasis on more granular input for prediction of adoption and treatment decisions. In fact, what my department does is help to inform these phase-II to phase-III decisions.

We have done simulation modeling to evaluate the impact of coverage with evidence development (CED), which was talked about a little bit earlier. What we are trying to do is to minimize the risk of either a false positive or a false negative. A false positive would be developing a product that we cannot sell. An example is an insulin product called Exubera that was taken off the market about a year ago for commercial reasons. It is very rare that a product gets pulled off the market for commercial reasons. A false negative would be to stop the development of a beneficial treatment unnecessarily.

In order to make better, smarter decisions earlier, industry needs more accurate estimates of expected net present value. I am going to give you a simple example of what we call an asset [see slide 8, "Simple example of a hypothetical "asset" in the investment portfolio"]. Think of it like a stock in an investment portfolio. I am calling it 'simple' because this example does not include taxes, overhead, and all sorts of other things; and normally we would extend the simulation to 32 years, but in this example, we are taking it only to 20 years. In the first three phases of development, there is no revenue and there are increasing expenses. The reason that I earlier identified the transition from phase II to phase III as being a key decision point is that between years four and five there is a dramatic increase in expenses. Once you start phase III studies, you are starting to spend real money. Later, in years 9 through 16, you have expenses but you have patent coverage for a product that is on the market, so you are generating revenue. Once the product goes off the market, revenue drops and expenses drop as well. Using this sort of spreadsheet, you can calculate expected net present value.



Year	Expenses*	Revenue	Net	Phase
1	$10	0	-$10	1
2	$10	0	-$10	1
3	$20	0	-$20	2
4	$20	0	-$20	2
5	$50	0	-$50	3
6	$70	0	-$70	3
7	$70	0	-$70	3
8	$50	0	-$50	3
9	$100	$400	$300	4
10	$100	$600	$500	4
11	$80	$730	$650	4
12	$80	$760	$680	4
13	$80	$800	$720	4
14	$80	$820	$740	4
15	$80	$840	$760	4
16	$60	$750	$690	4
17	$1	$300	$299	5
18	$1	$100	$99	5
19	$1	$50	$49	5
20	$1	$40	$39	5

McElwee: Slide 8

We used Monte Carlo simulation modeling to look at the impact of CED if a product has to go into a clinical trial for three years following launch. The model that we used for this was lung-volume reduction surgery, which was one of the programs that CMS [Centers for Medicare and Medicaid Services] did with CED. We looked at three different scenarios: one with a probability of trial success of 0.5 (we don't know whether or not this product works any better than the alternative), one with a probability of trial success at 0.7 (we think it works, but we are not sure), and the last one with a probability of trial success at 0.9 (we are pretty sure that it is better than the alternative).

In these scenarios, there are winners and losers, and there is much more downside for the losers than upside for the winners. The winners get a 5 percent prize, and the losers get a 75 percent penalty in the subsequent years. In the first scenario, with the probability of trial success at 0.5, the base case shows an expected net present value of about 1.9 billion dollars. With CED, the expected net present value substantially drops. We asked the question, if we could cut back our production costs in development by 20 percent, would it matter? And it made little difference. Skipping to the last scenario, note that this is simulation modelling, so the expected net present value is about the same as in the base case, about $1.9 billion. With a higher probability of trial success, CED still reduces the net present value, but not as much. Again, decreasing production costs has little effect. The conclusions from this are that the loss of revenue during the period of coverage with evidence development decreases the expected net present value of the base case. It is partially, but not completely, restored by better predicting winners. Finally, improved production efficiency has little impact on expected net present value.

Regulatory decision: The next key decision in the lifecycle is market approval by a regulatory agency. These are pretty predictable decisions. In the US, if we go to the FDA with the label that we want approved and the

development plan for generating evidence to support that label, it is not 100 percent certain, but it is close. It is a very predictable process. We meet with the regulatory agencies at the end of phase II. If the data are not clear, the agency will have an external advisory board to help them deal with uncertainty in the interpretation. Regulatory agencies have historically focused more on efficacy and safety than on cost, so it is clinical value rather than economic value that they are determining. There is a trend in some markets, such as in Europe, towards requiring active comparators in the regulatory trials. The regulatory agencies would like randomized trials and are not very interested in observational studies. They are willing to trade external validity for internal validity. You may have an active comparator in a trial, but it is still not a real-world trial that reflects real-world conditions.

Adoption and diffusion decision: Adoption and diffusion decisions by payers are not as predictable as regulatory decisions. There are no rules of the road or defined evidence requirements. Historically, the input from payers in phase II has been very informal. More recently, NICE has developed a consultation process based on the end of phase II regulatory process. I think the industry would rather go in this direction than have coverage with evidence development after launch. Their attitude is, tell us what we need in phase II, tell us the evidence requirements and make the process predictable rather than waiting until the product is on the market and then putting us into coverage with evidence development.

For a number of reasons, the evidence requirements of payers are different and much more challenging than the requirements of regulatory decision makers. Payers are more interested in external validity than in internal validity, and therefore are more interested in the use of observational studies than in randomized trials. It is very context-specific. Another challenge is that when you launch a product you do not yet have real-world data, and you therefore need a period of time in which to generate that data. Finally, there are no methodological standards for observational studies of benefit. As was previously mentioned, I sit on an AHRQ [Agency for Healthcare Research and Quality] steering committee for the Effective Health Care Program. In April, we had a weeklong methods conference at which statisticians from Harvard, Vanderbilt, and UCLA presented results of simulation modelling on when to use propensity scores, instrument variables, and risk stratification. They all had different recommendations. We need a standard way of going about this.

Finally, in the US, clinical-value and budget-impact assessments are much more common than economic-value assessments. Many of the healthcare plans do not conduct cost-effectiveness analysis, or have not done so in the past, although some of the large plans, such as United Health Care and WellPoint, have the capability of doing cost-effectiveness analysis and are now doing it. So I think that situation is changing. I would have answered a little differently the earlier question about the use of cost-effectiveness analysis, in that a national estimate of cost-effectiveness is not relevant for most plans. They all pay different prices for the products that they buy, and they want a cost-effectiveness analysis that is relevant to them. I predict that the large plans will take the evidence generated in the centre for comparative effectiveness and use it to do their own cost-effectiveness analysis.

I want to point out some data from an ongoing project of ours, the EBM Best Practices Study [see slide 20], in which we are using RAND appropriateness criteria to evaluate the impact of various aspects of the appraisal process. We gathered a group of experts from academia and the payer community and asked them to evaluate 42 hypothetical scenarios involving 6 drugs and 2 devices. The scenarios varied in the quality of evidence on efficacy, the quality of evidence on safety, whether or not the intervention is lifesaving, whether or not it improves quality of life, and whether it costs more than other treatment options. There was agreement in exactly half of the 42 scenarios, disagreement in 5, and the other 16 were indeterminate. The RAND approach uses a scale of 1 to 9, and the criteria for disagreement are pretty strict: there must be three or more ratings in the 1 to 3 range and three or more in the 7

to 9 range. In the summary scores for all 42 scenarios, the likelihood of reimbursement ranged from a median of 5.3 to a median of 2.1. So we see significant variability in experts' evaluation of the evidence.

McElwee: Slide 20

Treatment decision: Treatment decisions by patients and their physicians are ideally informed by best evidence, physician expertise, and patient preferences. As we all know, formal incorporation of patient preferences into treatment decisions is rarely done, but a number of recent studies have shown the importance of this. A study that was reported two months ago at the meeting of the Society for Medical Decision Making showed a significant difference in preferences between white and African-American patients being treated for rheumatoid arthritis. Whites were more driven by the benefits of treatment, whereas African Americans were more driven by safety concerns, particularly catastrophic safety concerns.

Getting from population-based "best evidence" to the best choice for a given patient is not easy, especially when heterogeneity exists in the data. There are two ways to go about this. One is through genotyping, which reduces uncertainty at the individual patient level. We do that to some extent now, but it is a long way from being in the mainstream. The second way is through what I call actuarial diagnostics, where the patient phenotype or other data is used to reduce uncertainty at the subgroup level. When physicians make treatment decisions, they base them on the average effect shown in a study of many individuals. Whether or not an individual patient should get a drug really depends on how much that patient is like the average. It is very difficult to determine.

In summary, the value of pharmaceuticals may be assessed and appraised differently, depending on the type of decision and the preferences of the decision maker. We need thoughtful policies that balance cost control with broad access and continued innovation. This requires more predictability and transparency in adoption and diffusion decisions, and will require that developers partner with stakeholders throughout the lifecycle of the product. The type of programs that NICE is starting to develop should, I think, be expanded, particularly in order to ensure input into early-phase development decisions. Individual treatment decisions could be improved by better incorporating patient preferences and heterogeneity of treatment effect into the treatment decision. The final point I would like to make is that we need first to do decision-based evidence making and then do evidence-based decision making. Thank you.

Andy Greenshaw: It is my pleasure to introduce Stuart MacLeod, the Executive Director of the Child and Family Research Institute and Associate Dean at the Faculty of Medicine at the University of British Columbia. He is also Vice-President of Academic Liaison and Research Coordination for the Provincial Health Services Authority. Stuart has had a tremendous career track, which includes tenure as Dean at the Faculty of Health Sciences at McMaster University and founding Director of Clinical Pharmacology at the Hospital for Sick Children.

Decision-Maker Perspectives on Valuing Health Outcomes for Policy Decisions

Stuart MacLeod, Vice-President, Research Coordination and Academic Development, Provincial Health Services Authority, Vancouver

I am going to give you a very quick overview based on my 40 years of experience in trying to make some sense of this business of valuing health outcomes. I should warn you that I am a clinical pharmacologist, not a health economist, so the naïvety of my views will probably be obvious to all of you. I am going to try to build on the things that Newell has been saying, but I am going to do so with much less erudition and elegance. I have been around for a long time, as you can see, and in my lifetime there has been a huge paradigm shift in medicine. When I started out a long time ago, when dinosaurs roamed Alberta, the people in the major teaching hospitals, the sub-specialists, specialists, and research experts, had major sway in decision making. But as everybody in the room knows, this has shifted entirely so that most of the power now rests with patients, families, consumers, and, most particularly, with the payers, including — in Canada — government.

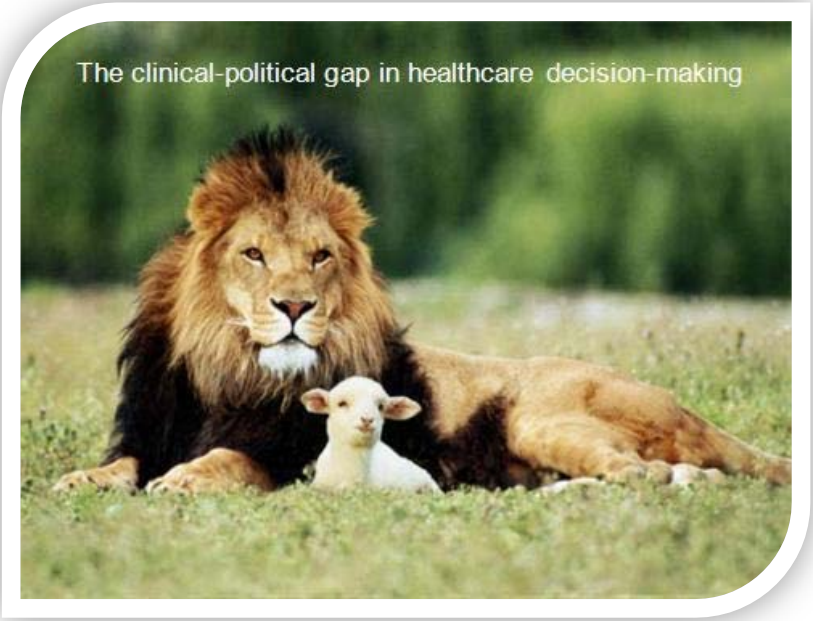
If we are going to look at how we value health outcomes, we have to ask whether the values of decision makers [are] different from societal values. One would think not, but as a sometimes clinician, I frequently see clashes between what appear to be the values of the political decision makers and the values that are more commonly held in our society. And I think that both are sometimes divergent from the values held by the health professionals who are looking after patients.

The challenge, and it is a very, very complex one, is to find a balanced view. Clearly, we know what questions to ask about health interventions: Does it work? Does it work in real life? For whom does it work? Is it safe? Compared to what? How much are we willing to pay for it? But there are different emphases placed those questions, depending on whether you are a public-health decision maker, a physician, or a patient. Obviously, public-health decision makers are interested in knowing what is best for a population of patients; physicians are interested in what is going to work for their particular patients; and the patients just want to know whether their disease is going to be cured or mitigated in some way.

I don't have a complex message today. I want to suggest that decision making should be the result of a deliberative process that considers the context in which the treatment is given or the intervention is used. In deference to our hosts today, I start by saying that there should be a central focus on comparative effectiveness. I think we have been doing this for most of my lifetime, but perhaps calling it by different names. You heard in the introduction that I was for a time the Dean of Health Sciences at McMaster University. Not everybody in the room may know that that is the homeland of evidence-based medicine. I was there when they came down from the mountain with the stone tablets; and certainly far be it for me to be skeptical about evidence-based medicine, but I think it can be used in too reductionist a fashion. We do not always have the evidence on which to base decisions, and we need to start to accept the idea that our decisions should be evidence-informed rather than based on the kind of evidence that comes from randomized controlled trials (RCTs). We need to improve our ability to use observational studies that go beyond RCTs. We need to look at composite outcomes and at colloquial evidence. For those of you who have not read what Jonathan Lomas has written about colloquial evidence, it is very compelling. And, of course, I am not diverging

completely from the dogma of the Cochrane Collaboration: I do think that we should follow their Effective Practice and Organization of Care Group approach.

This is my metaphorical depiction of the clinical–political gap in healthcare decision making [see slide 6: "The clinical–political gap…"]. It won't surprise you, perhaps, to know that I associate the politicians with the lion, although I have to say that most of the world's politicians haven't looked particularly leonine in recent times, and that's nowhere more true than in Canada at the moment. The lamb, of course, the sweet little lamb, represents to my mind the clinical decision makers, the medical fraternity, and the researchers who struggle to produce the evidence that we need to inform our decisions. For those of you with a biblical bent, this reference comes from Isaiah in a passage about the calf and the young lion and the fatling lying down together. I am not quite sure what a fatling is, but we can discuss this later. I go to Africa often, and my friends in Africa like to say, "Well, the lamb may lie down with the lion, but the lamb isn't going to sleep very soundly." I think that is the reality in health care.



MacLeod: Slide 6

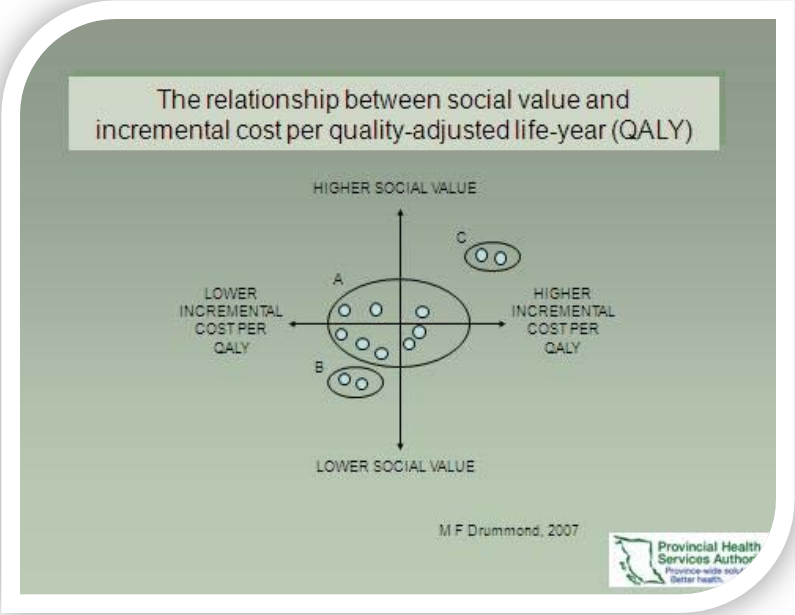
I am intrigued by the relationship between the scientific community and our political decision makers. Recently, I came across Jacob Bronowski's famous book, The Ascent of Man, in which he talks about the reasons that Einstein refused the presidency of Israel. Einstein was apparently offered the opportunity to be the first president of Israel, and he declined because he did not believe that he could reconcile his life as a scientist with the demands of being a politician. This is Bronowski's take on it: "…science is also a source of power that walks close to government and that the state wants to harness. But if science allows itself to go that way, the beliefs of the twentieth century will fall to pieces in cynicism. We shall be left without belief, because no beliefs can be built up in this century that are not based on science as the recognition of the uniqueness of man." I think we are still facing this dilemma, and this is why we are looking for tools like comparative effectiveness research to help with our decisions.

I have already talked about alternate perspectives on outcomes. I think that public and private views on this must be aligned in some kind of true partnership, and we are starting to see that with organizations like the IHE and similar institutes in other parts of Canada.

When we try to put a value on health outcomes, there are a number of things that we need to consider alongside our cost-effectiveness calculations. (And I agree absolutely that clinical- effectiveness and cost-effectiveness have to be approached separately and then joined together.) We need to consider whether alternative treatments are lacking or inadequate. Obviously, we are going to approach it differently if the treatment is completely novel. We need to look at how serious the illness is. Is it fatal? Patients do have a right of rescue, but that is not always acknowledged by our decision makers. We need to look at affordability, at the budgetary implications for government. And I think that, at

some level, we have to consider the overall social impact of the innovation and the environment in which innovation occurs.

I stole this graph from Mike Drummond (2007) [see slide 11, "The relationship between social value and incremental cost per quality-adjusted life-year"].



MacLeod: Slide 11

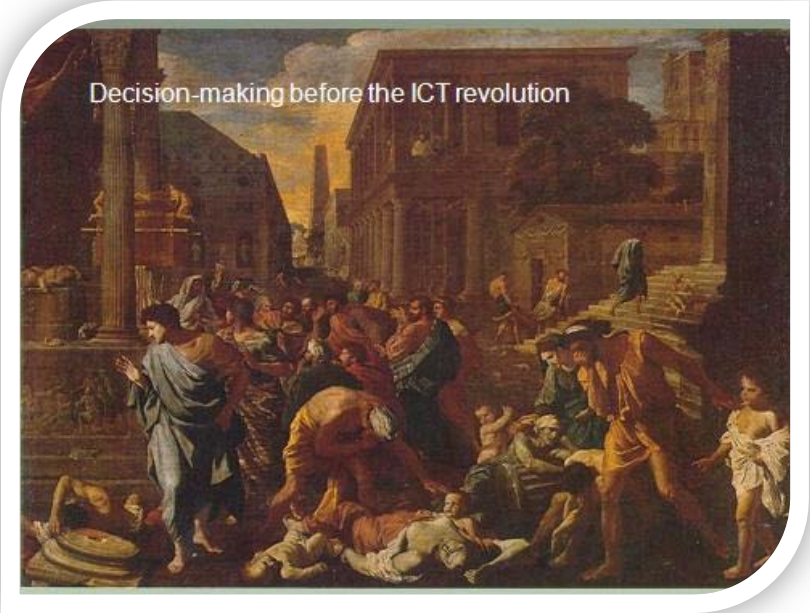
In the last two years, we have had a number of seminars on drugs for rare disorders, and Mike talks very effectively about the need to look at social value as we make these decisions. Most decisions sit in the middle ground where we can clearly see the incremental cost per quality-adjusted life-year (QALY) and make a decision on that basis. But sometimes we are willing to pay a higher incremental cost for something of higher social value; and when the social value is trivial, sometimes we are willing to say we don't want this innovation at any price.

I will skip over the question of research design. It is self-evident that if we want to get to the heart of these questions, we have to have the strongest possible research designs.

The late Allan Williams gave a lecture at the Office of Health Economics in 2004 that was entitled "What could be nicer than NICE?" He stated, "NICE is the closest anyone has yet come to fulfilling the economist's dream of how priority setting in health care should be conducted. It is transparent, evidence-based, seeks to balance efficacy and equity, and uses a cost-per-QALY benchmark as the focus for its decision making." That is the quality folly, and here is the zinger: "Experience has taught me that it is not uncommon for an economist's-dream-come-true to be seen as a nightmare by everyone else." I am not sure that many of us would say that it's a nightmare, but it does have some limitations.

Williams also talked about the turbulent political, intellectual, and cultural environment in which NICE is called upon to conduct its work. It is no surprise to any of us that there are clashes across this spectrum. In the end, we are left trying to reconcile an individual perspective and a population perspective and to decide whether or not we are pursuing optimization or some kind of bounded rationality. Voltaire told us 200 years ago that 'best' is the enemy of 'better.' We are too often held back from pursuing an improvement in treatment because we hope that we are going to find the optimal treatment, the ideal treatment. Often this is reinforced by the political decision-making process.

This is Poussin, the plague coming to Ashdod, just to remind you that no matter how difficult are the times we live in, they were more difficult in the past [see slide 15, The Plague of Ashdod, painting by Nicolas Poussin, 1630]. In fact, we have many advantages, because the revolution in information and communications technology allows us to do the kind of population health research that I think will lead us to better decisions.



MacLeod: Slide 15

Another position that is worth considering is that expressed in the Daniels and Sabin paper of 1997 that looks at the four elements of legitimacy and fairness in public decision making: stakeholder involvement, transparency and dissemination to the public, revision or appeal, and leadership. The latter includes — and I think this is the most important point — accountability for reasonableness. I think that we do have to hold our political decision makers accountable for reasonableness; but to be realistic, we need to recognize the impediments to their appropriate valuation of outcomes. People who make decisions about public expenditure may well feel that inertia is the safer course. They recognize that innovation has a high cost in many cases. They have to consider the interplay of values and evidence, as Newell talked about, and to appreciate that there are divergences in social values. Decision makers also face ambiguity about effectiveness and lack of consensus on methods. I think we need to pursue better methods — new methods appropriate to the times. Finally, I believe that if we are going to make the right decisions in these contexts, we need to have free access to data across the provinces and allow people with appropriate expertise to work with the data. If we do not keep evolving and considering new ways of doing this and developing appropriate methodologies, we will fall behind. As they say, if you are not the lead dog, the view doesn't change very much. Thank you very much.

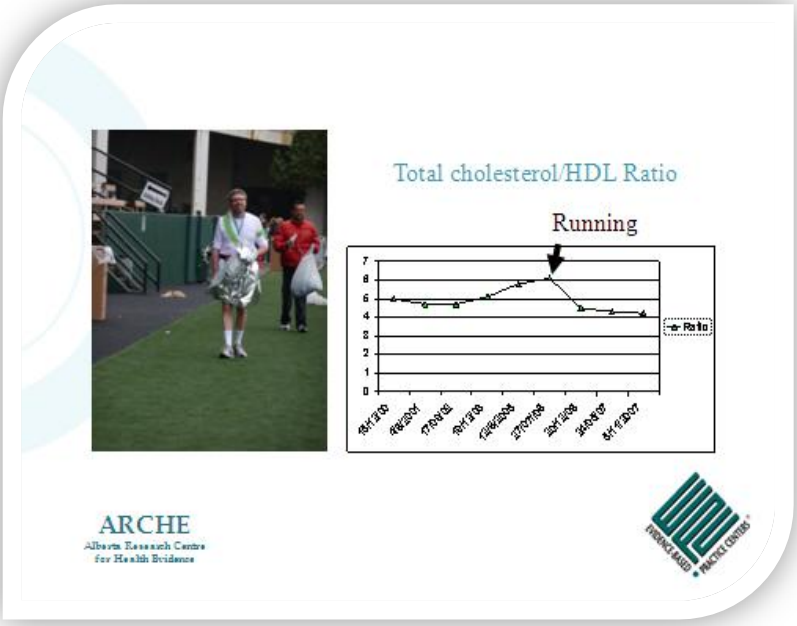
Andy Greenshaw: It is now my pleasure to introduce Dr. Terry Klassen, the founder of the Alberta Research Centre for Health Evidence and the champion of evidence-based practice. He is a very successful researcher, widely known and widely published in the major journals in a variety of areas, including acute respiratory disease and injury prevention and control. He is currently the director at the University of Alberta Evidence-based Practice Centre, which is funded by the Agency for Health Research and Quality in the United States.

Cross-Border Shopping for Innovation: The EPC Model

Terry Klassen, Director, Alberta Research Centre for Health Evidence (ARCHE)

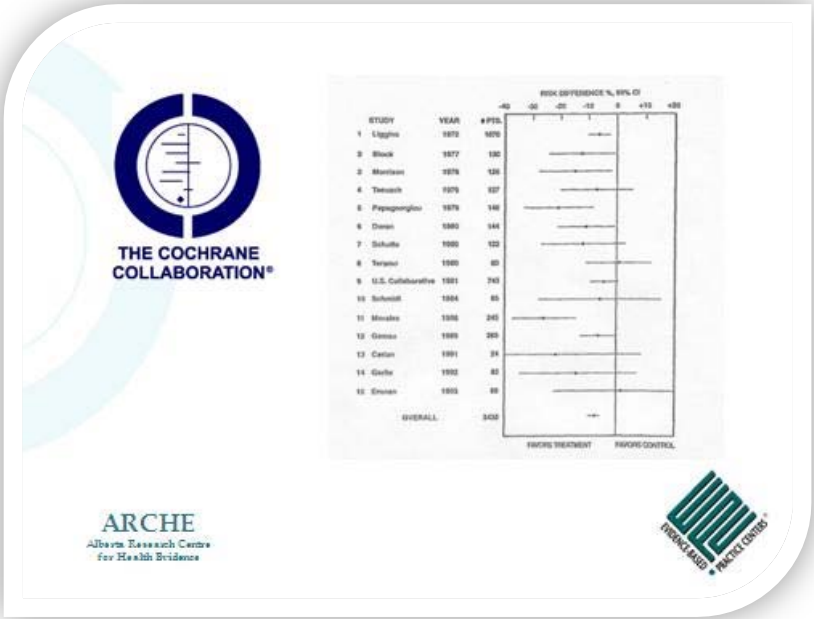
I want to talk about our experience at the Alberta Research Centre for Health Evidence, but first, I would like to acknowledge my co-director, Lisa Hartling, who has done much of the work. You can run a research centre of this nature only if you have people like Lisa working alongside you.

Evidence is very personal. This photo is me, 48 hours ago, having finished the Seattle Marathon [see slide 2]. Now, you might ask, why did I get into this activity and what's the use of it? My running career started about two years ago, after I had a wake-up call from my doctor. My total cholesterol to HDL [high-density lipoprotein] ratio was out of whack, meaning that my HDL had gone down and my total cholesterol had gone up, so I started on a course of running. On the graph, you can see the gradual downward trend in the total cholesterol to HDL ratio. Later, evidence of reduced disability and mortality among aging runners came out ["Reduced disability and mortality among aging runners," *Arch Intern Med* 2008;168:1638-1646], and, being over 50, I do qualify as an aging runner. I must say, though — and this is the personal dimension to evidence — in reference to reduced mortality, in the second half of the marathon, I thought I would die. I can also say that my quality of life has been quite reduced since the marathon: I barely could make my way up to the stage. So, as you can see, you don't always experience the evidence. You instead have to look at the cold hard numbers and the fact that maybe one day you will benefit.



KLASSEN: SLIDE 2

This *New York Times* article from October 24 [2008], caught my eye. In ***How to Take American Health Care from Worst to First***, authors Billy Beane, Newt Gingrich, and John Kerry asked why the New York Yankees, the Detroit Tigers, and the New York Mets, with the three highest payrolls in baseball (a total of $494 million per year), were all, by the time the playoffs arrived, sitting in the bleachers and enjoying the summer break. Meanwhile, the Tampa Bay Rays, with the second lowest payroll in baseball, at $44 million, were playing in the World Series. The authors tried to make the case that baseball has gone evidence-based, that using the evidence of players' statistics to form high-quality teams results in a tremendous output. Making the transition to health care, they pointed out that a key organization in evidenced-based health care is The Cochrane Collaboration. On the left [see slide 5] is a logo of the Cochrane Collaboration. Each line in the logo represents one of the trials (shown on the right) that contributed evidence about the use of corticosteroids in reducing morbidity and mortality in premature infants. You can see that,



KLASSEN: SLIDE 5

in 1972, Liggins made a basic science observation. He noticed that premature lambs treated with corticosteroids had air in their lungs and a higher survival rate than those who were not treated. Based on the reduced mortality in his baby lambs, he then conducted a randomized trial in humans. In 1972, the first trial, which was probably one of the largest, showed the effectiveness of corticosteroids in reducing morbidity and mortality. However, more randomized trials followed, and, of course, the outcomes fluctuated with the sizes of the studies, and so more research was needed. Between 1972 and 1993, 14 more randomized trials were conducted. Probably the answer was there in 1972, and many more babies might have lived or had reduced morbidity, which would have reduced healthcare costs. The take-home message in this is the centrality of evidence in providing health care that has an impact on patients' lives.

Another example has to do with croup. Croup is a pretty simple disease. I am a pediatric emergency doctor, and, to me, croup is about a windpipe that narrows. The patient cannot get air into the lungs, so you give something to reduce the swelling, and the air flows in and out quickly. The evidence in Cochrane that supported this, in 1994, showed that glucocorticoids, or steroids, could reduce the probability of being admitted to hospital. If you are running a healthcare system or a hospital, making sure that such evidence is used has an impact. Hospitalization data from Ontario between 1988 and 2001 shows evidence of a dramatic drop in the hospitalization of children. However, David Jonsson from the Alberta Children's Hospital and I looked at hospitalization in Alberta, and we noticed a high variation. The probability that a child with croup will be admitted to hospital depends on where in the province the child presents. In particular, we noticed a greater use of steroids in high-volume hospitals. We are doing a randomized trial to assess various strategies for ensuring that evidence is utilized, from just mailing out guidelines to other interventions. We will know the results of that soon. A little anecdote in this is that David, who had the pleasure of driving to many small Alberta communities, visited one little rural town where they still had my 1994 New England Journal trial pasted up on the wall of their emergency department. Now, the trouble was that we had done a later trial, in 1998, that showed that the steroid budesonide maybe isn't optimal; but at least they had evidence up on the wall for the treating doctor or nurse to consult.

Scope of ARCHE

ARCHE, or the Alberta Research Centre for Health Evidence, was initiated in 2000 by an establishment grant of half a million dollars from the Alberta Heritage Foundation for Medical Research (AHFMR). It was probably a shrewd investment by AHFMR, as we have now parlayed that into about $8 million in funding. Key components of ARCHE are The Cochrane Collaboration and the University of Alberta Evidence-based Practice Centre. We see ourselves as an important partner in the whole healthcare scenario, and I will outline some of our projects in order to give you an idea of the scope of ARCHE:

- We worked with the Canadian Agency for Drugs and Technologies in Health (CADTH) in sampling some of the evidence on emergency department overcrowding. Brian Row of CADTH received a great deal of attention for this particular report, and the evidence that was developed became very critical when the Health Quality Council of Alberta addressed the issue of emergency overcrowding in this province. We haven't eliminated emergency overcrowding, so we know that evidence does take a while to solve the problem.
- We did a study for Health Canada on volume–outcome relationships in pediatric care, especially pediatric cardiac care, showing that consolidation into high-volume centres produces better outcomes. This supports some of the service-delivery changes we have seen in western Canada, such as the consolidation of pediatric cardiac surgery into two centres, in Edmonton and Vancouver.
- We did some work for the Alberta Centre for Child, Family and Community Research (ACCFCR), looking at the risk factors and treatment interventions for crystal meth use and at behavioural interventions for autism. That was an interesting experience of engaging decision makers in the process of doing systematic reviews. Afterwards, the ARCHE researchers said they did not know if that interaction had been important; but when we did an analysis of the knowledge transfer, the decision makers said that they had found the process to be very useful.
- More recently, our national funding agency, Canadian Institutes of Health Research (CIHR), has seen the value of our work and has funded work on evidence synthesis and practice guidelines for bronchiolitis.

We think that the creation of this centre has had several benefits. As I mentioned, we have generated funding, much of it from the US, but we have also been able to employ Albertans and generate intellectual capital right here in the province.

Evidence-Based Practice Centre Program

A key component of ARCHE is the University of Alberta Evidence-based Practice Centre program, one of 14 evidence-based practice centres across North America that conduct systematic reviews for, and are funded by, the US Agency for Healthcare Research and Quality (AHRQ). The mandate of AHRQ is to review all scientific literature on clinical, behavioural, and organization and financing topics to produce evidence reports and technology assessments. These reports are used for informing and developing coverage decisions, quality measures, educational materials and tools, guidelines, and research agendas. A third part of the mandate is to improve the way we do these things by conducting research on the methodology of systematic reviews. The 14 centres operate on five-year contracts as arm's-length, objective entities. We are required to declare any conflict of interest on evidentiary report.

The topics for systematic reviews are assigned by AHRQ through a formal topic-nomination and selection process. Topics are nominated by external agencies, such as Medicare, Medicaid, the Office of Medical Applications of Research, and the US Preventative Services Task Force, and then there is a process of selection. AHRQ is attempting to make this process more transparent and to provide opportunity for external input, including public comments (e.g., Is this the right topic? Should it be this?) They are also trying to engage the evidence-based practice centres in refining the questions.

We started in 2002 and now have 20 staff with a diverse set of skills in epidemiology, statistics, information sciences, and basic and medical sciences. We also have collaborators and partners across the university and the province, people with knowledge in key clinical areas. I will give some examples of the work we have done.

We conducted a systematic review on occupational asthma, looking at the optimal diagnostic approach and the effectiveness of alternative management and therapeutic options. The topic nominator and sponsor was the American College of Chest Physicians. The result of our review was the publication in 2008 of a consensus statement to serve as

a guideline for physicians. A key advantage in working with such a partner is that there is a greater chance of having an impact.

Another of our systematic reviews was a report on secondary prevention programs for patients with coronary artery disease. The key questions were: "Do secondary programs improve health outcomes?" and "What program-related factors influence effectiveness?" This review had political implications, in that someone had developed a rehab program and had been very active in pushing to get his program to the top, and here was the evidence report giving a cooler assessment of his program. The topic nominator was the Centers for Medicare & Medicaid Services (CMS), and the review was used as a report to guide CMS in its coverage decisions. We presented the report at a meeting of the Medicare Coverage Advisory Committee in Washington. It was then discussed and a decision was made that cardiac rehab is reasonable and necessary for specific indications.

Our latest review looked at the evidence for the use of positron emission tomography (PET) as a diagnostic test for nine cancers. It was very difficult, very complex evidence, moving into the areas of diagnosis as part of the management strategy and of cost-effectiveness. The review was commissioned by the Centers for Medicare & Medicaid Services (CMS) and used to guide decisions about reimbursement for PET scans. The following quotes from the news release show how evidence gets played out in the public arena: "Medicare recently commissioned a separate analysis that found the quality of available studies on PET's use in cancer was poor to moderate….Some panellists said it was hard to reconcile that finding with the industry's conclusion from the registry data…. 'I don't see how you make the leap,' said panel member and consumer representative Linda Bergthold. 'The quality of the data… was stunningly poor.'" The news release noted that Medicare is in the process of drafting a decision before it makes its final ruling.

I hope that I have shared some of our experience of the use of evidence reviews in decision making, whether it be to develop clinical-practice guidelines, to compare an array of different interventions in a network analysis, or to try to identify uncertainty. Despite the example that I gave of the potentially unnecessary trials on the use of corticosteroids in treating asthma, sometimes future research is needed and important. But sometimes it is a copout. I don't think there is a formula for deciding what we should fund, but we definitely need to engage in a decision-making process that is transparent and that has input from the various interest groups that have a stake in it. Thank you very much.

ANDY GREENSHAW: Thank you very much, Terry. I would now like to introduce Dr. Tom Feasby, one of my colleagues on the board of the Institute of Health Economics, and the Dean of Medicine at the University of Calgary. Tom is going to lead the question and answer session.

Tom Feasby, Dean of Medicine, University of Calgary

It might be useful to shed a different light on the issue of comparative effectiveness by looking at therapeutic interventions through the lens of appropriateness. Appropriateness in health care refers to making the right healthcare intervention. In other words, there should be a net positive effect. It is measured by a process that takes into account best evidence, balanced expert opinion (a panel), and a chart review of real cases. This assessment method was developed at RAND/UCLA in the 1980s and has been used to assess the effectiveness of many diagnostic and therapeutic interventions. I will give you

just two examples, both common procedures and both quite expensive and having significant related complications.

The first, carotid endarterectomy, a procedure that has been around for 50 years, is an operation on the carotid artery in the neck to prevent stroke. We conducted a study using the RAND/UCLA method to look at the appropriateness of carotid endarterectomy. A Canadian expert panel reviewed over 3,000 cases in the four western provinces of Canada in 2000 and 2001. In 10 percent of the cases, carotid endarterectomy was found to be inappropriate; in other words, it clearly should not have been done. In 37 percent of cases, mostly cases in which patients had no symptoms whatsoever, the procedure was found to be of uncertain appropriateness. So, all told, almost 50 percent of the cases were of uncertain appropriateness, and yet we continue to do these procedures right here in western Canada.

My second example is the infusion of intravenous immunoglobulin or IVIG, a blood product used to treat immune and infectious diseases. This procedure has been common practice for about 25 years and is very expensive. Again, a RAND study of over 3,000 cases in Alberta and British Columbia in 2001 and 2003 showed the procedure to be clearly inappropriate in almost 30 percent of the cases, and of uncertain appropriateness in another 17 percent. So, again, almost 50 percent of the cases were either inappropriate or of uncertain appropriateness.

The questions to be asked about an intervention are can it work and does it work? The question "Can it work?" refers to efficacy, and the method for determining that is randomized controlled trials. The question "Does it work?" refers to effectiveness, and the measures of effectiveness are comparative effectiveness, in particular as determined in practical or open clinical trials, and appropriateness. Comparative effectiveness and appropriateness are related, but they have different qualities. Comparative-effectiveness studies examine whether interventions that have been shown to be efficacious in trials actually work when they are used appropriately in a diverse population, or in the so-called real world. Appropriateness studies examine whether interventions that we know are efficacious are actually applied appropriately in a diverse population. Our studies have shown that healthcare interventions in the real world are used inappropriately in up to 50 percent of cases. This clearly reduces the overall effectiveness of those interventions. The question of appropriateness needs to be taken into account in comparing the overall effectiveness of treatments, and it offers an avenue for improvement.

General Discussion

Tom Feasby, Moderator: We have an expert panel ready to discuss a number of issues, and I invite you all to join this discussion. We have plenty of food for thought. I will start with a question that I think might be tackled by a number of our speakers. It is one thing for us to decide what we should do or not do, based on comparative effectiveness. It's an entirely different matter to stop doing the things that we shouldn't be doing, the things that are less effective. You talked a little bit about incentives. I wonder if you could give us a sense of how you might go about doing this.

Gail Wilensky: I am going to do that in exchange for asking you a question, which is, why do you think that in Canada, where financial incentives to physicians are very different from those in the US, 50 percent of the endarterectomies being done are either inappropriate or of uncertain appropriateness?

Tom Feasby: That's a good question, and I am not sure I know the answer. The first study on this, done in the late 1980s and published in the New England Journal of Medicine, had results not so different from what we found 15 or

20 years later. We have come a little bit further and have better evidence now, but the incentives are rather the same. It's mainly fee for service in our system, so you could perhaps think of that as an incentive to perform endarterectomies. With intravenous immunoglobulin, I think that physicians sense that it is readily available and easy to use, so why not? That's the attitude.

GAIL WILENSKY: In response to the question that you posed about how to effect change, I don't think there is a good understanding of why it has been so difficult to change clinical behaviour in the face of what appears to be pretty convincing evidence. I think the strategy will require bringing together disparate pieces of information to provide better clinical guidelines based on evidence, and financial incentives to back them up — and I do think that you need both. One of the sobering experiences that I had earlier this year was meeting and speaking to the group that Newell mentioned — a committee led by Hal Saxon and Barbara McNeil that produced a report which the Institute of Medicine put out in January. One of the things that they found is that the systematic reviews that have been done have raised many questions with regard to the data included in the various meta-analyses. What I had assumed was not an issue, which is identifying what we know based on existing research, is much more problematic than I had understood. It is not nearly as difficult a problem as generating new data, but it was sobering to see how lacking we are in rules of the road for doing systematic reviews. That really caught me by surprise.

I don't know whether this explains why it is so hard to change behaviour or whether the explanation lies in the huge number of guidelines that exist — I think they cite 200-plus for the treatment of hypertension. There is an effort in the UK to establish clinical pathways — to get the various groups working together, the royal colleges, the royal societies, maybe NICE as well, and to come to a consensus opinion and then distribute it to all clinicians treating chronic disease, acute care, or birth or end-of-life matters. Rewards would be based on compliance — not 100 percent compliance, but majority compliance, on the grounds that interventions will not be appropriate in all cases. This is a way both to firm up the information and to provide incentives to use it.

The proliferation of guidelines and information is a problem, and the lack of agreement about what constitutes evidence is probably more of an issue than I thought. However, even if we resolve those issues, change will, I think, require positive financial incentives. Even in areas where individuals are not financially benefiting, we see aberrant behaviour. For example, in the UK, high levels of tonsillectomies and hysterectomies were being performed contrary to clinical evidence, and they were going through GPs who were not gaining by it. Only by giving incentives to behave differently, consistent with clinical guidelines, have they been able to change that behaviour. So it clearly remains an issue of having both credible information that is accepted by the professions and financial incentives that move individuals toward the behaviour you want.

TOM FEASBY: Thanks, Gail. I invite others to comment, particularly Terry. Terry, you are in the business of sometimes trying to get your colleagues to change what they do on the basis of evidence. It's not an easy job. Any comments on this?

TERRY KLASSEN: It is tough. First of all, on the evidence-synthesis side, I agree that historically there have been some challenges and that the quality has not been as good as it should be. But The Cochrane Collaboration and other groups have been improving the science of systematic reviews, and I think now the challenge is behaviour change: that is, getting those who do systematic reviews to adhere to the highest standards. The methodology is well developed, and we know that Cochrane reviews generally have a higher quality of methodology than those published in the paper journals.

As for changing the behaviour of healthcare providers, I think the science of knowledge translation has been evolving rapidly. In Canada, we are particularly fortunate to have leaders in what we now call KT, such as Jeremy Grimshaw,

Sharon Straus, Carol Estabrooks here at the University of Alberta, and others. They recently were given a large chunk of money to advance the science of knowledge translation. We are learning a lot about successful strategies, maybe incentives at times, for bringing about behavioural change. In our little study in Alberta, we mailed out guidelines to the physicians in our control group and to the two intervention groups. One of these was an opinion-leader intervention, in which you identify a physician-leader in a hospital or community and use the Jonathan Lomas opinion leader approach. The other intervention was a champion approach, in which you identify a nurse or respiratory therapist leader who champions the change process within a hospital. What we have seen is that the hospitals that had either an opinion leader or a champion leader increased their use of steroids, and this had an impact on what happened to the kids. There is a strong temporal trend in the use of steroids, but the evidence was used about two or three years earlier in the intervention sites. Combined with that cluster randomized trial, we also included a qualitative part in this project. We noticed that in the sites that had intervention, some responded and some didn't. It was sort of like opening a black box to try to uncover why one hospital started to change and grab the evidence and another one did not. So those are some of the strategies, and that's just an example of where the science is going. As we understand better how to incorporate change, there will be strategies at the organizational and systems levels. Our new funding models in Canada, the Academic Alternate Relationship Plans (AARP), have no volume-based incentives, and it should be interesting to see the impact of those on care provision. In paediatrics, we have been on an AARP for, I think, four years. We still have to look at outcome and performance and we don't know quite how to evaluate that, but at least we are not driven by "this pays us more." It has been a very innovative change in the funding model.

GAIL WILENSKY: Let me clarify some the comments I made about incentives. I don't necessary mean paying people to do the right thing, although I think that is certainly one strategy. I do mean making sure that incentives are not either prohibiting or making it difficult for clinicians in institutions to do the right thing, as is the case in some parts of Medicare, where the structure is such that if you follow the rules, you will be financially penalized. That is a very bad position to be in. You can also develop incentives where the individual may not be directly rewarded, but the group that he or she is involved with, either the institution or the group practice, gains in a positive way. In other words, I define 'incentives' broadly, rather than using the most narrow interpretation of paying somebody to behave in one way versus another.

NEWELL MCELWEE: A lot of the discussion in the US has been about payment reform, and Gail herself has published on gainsharing, which is what she was just talking about. From my perspective, it is not all about financial incentives, which is what both Terry and Gail were hinting at.

I will just give you an example of a study that we did, which is still ongoing, with the Evidence-Based Medicine Centre at Oxford in the UK. They did a point estimate of the proportion of time that physicians adhere to treatment guidelines — it was around 19 or 20 percent of the time — and then they talked to physicians about why they didn't adhere to the guidelines. What they heard was, "Well, you know, the guidelines are developed in ivory towers, and they don't really apply to our patients." So they developed a program called a "journal club," but it's not a journal club as we know it. They brought together physicians within a 50-mile radius of Oxford on a weekly basis. Every week they were asked to write down the clinically important questions that they encountered in their practice. At the journal club meeting, they vetted those questions, developed a priority list, and then looked at the evidence to answer those questions. They had one bucket of questions for which they didn't have enough evidence to answer, and another of questions for which there was enough evidence to come up with some guidelines, which they developed themselves. The adherence rate for the guidelines that they developed in this bottom-up approach was almost 100 percent. So I think there are ways to approach this other than through financial incentives. I don't know whether this is scalable, but there are all sorts of ways that you can look at doing this.

TOM FEASBY: This is very timely since, in Alberta, we are now undergoing a major reorganization of our healthcare delivery system, and we have all of the problems, I am sure, that Gail alluded to in her opening slides: large variation of practice, quality issues, inappropriate use, and so forth. , and in addition, we have funding concerns. We have a single pair system, but there are substantial funding concerns. I wonder if any of the healthcare leaders in Alberta would like to comment on how they see using evidence of the kind we have talked about to improve the delivery of services and to respond to these issues to which we have been referring. Anyone want to take on that one?

KEN HUGHES: My name is Ken Hughes, Chair of the Alberta Health Services Board. This has been very insightful and a bit sobering. From a governance perspective, we would like to deliver health care to all Albertans based on an evidence-based, science-based model. I think that if we are able to move in that direction, we will also make much better use of the resources we are already deploying. One of our objectives is to create a sustainable healthcare delivery model, one that we can continue to fund in spite of almost daily fluctuations in revenues. I think we have a fabulous opportunity here in Alberta to do things in a more disciplined, focused, and scientific fashion, because we have the capacity to execute on a province-wide basis. From my perspective, and I think from our board's perspective, this is the direction we should go. This kind of forum, bringing the family together for this kind of discussion, is extremely useful. The challenge we face is creating a continuously improving organization in an industry that frankly is relatively resistant to change. I welcome the engagement of everybody in this room in figuring out how we can facilitate continuous improvement and scientific improvement in the delivery of health care.

GAIL WILENSKY: I think there are two places to start. The first is to ask what kind of data exists in Alberta about clinical outcomes or variations in the rates of major procedures, and whether that something that you need to start working on while you are proceeding with whatever else you do. And the second question is whether or not there is a general comfort level with the kind of decision making that goes on now regarding coverage or reimbursement. I assume that most of that is provincial rather than federal.

TOM FEASBY: Certainly, the decisions are provincial. We have a tremendous need to use the data that we have, as well as to collect more data, and to pool this information in a way that makes it accessible so that we can make informed decisions. We cannot implement the changes that are required, which I think are vital, without first getting the right data and then doing the analysis. Many of us in this room are eager to help with this. We think it is very important, and we need to work together. I am glad you brought that up.

TERRY KLASSEN: I think we have very high-quality evidence in Alberta. The challenge has been the accessibility of evidence and the ability to use it to shed light on the issue. What we have found in producing evidence reports is that we need to find the receptors within the system. It's not always clear to me how decisions about new interventions are being made. We need to focus decision making within the system. We also need to look at the way in which we keep adding things but never get rid of other things. We would love to be part of this effort. As an evidence centre, we go to the US to engage in some of that, but we would find it very refreshing to do that right in our own backyard.

TOM FEASBY: Any other comments on that from the audience?

LORNE TYRRELL: One of the messages that may go out from here is that half of the things we do are inappropriate, and I am just a little concerned about that. I want to ask you a question on the appropriateness studies. Are they done prospectively?

TOM FEASBY: No. They are done retrospectively, reviewing charts.

LORNE TYRRELL: Is that fair?

TOM FEASBY: It's the rear-view mirror, but you know, the future is often best predicted by the past.

GAIL WILENSKY: As much as people complain about overuse of cardiovascular procedures, we have high numbers of individuals who first present themselves when they have sudden cardiac arrest. So when you do these studies, you don't know who you are missing. You have to be careful. It is one thing to say that 15 or 17 percent of interventions are inappropriate, however you define inappropriate. But the "uncertain" category, which is quite large in the examples, is a whole different world, because "uncertain effects" are likely to be viewed differently from the patient's perspective and the clinicians' perspectives. I think you have to treat "uncertain" as a third category. Half of the interventions are appropriate, but the bigger part of the other half are of uncertain appropriateness, not inappropriate.

TOM FEASBY: I think that's a very good point. Appropriateness doesn't tell you the whole picture. It's only one side of the coin. It deals with overuse. It doesn't deal with underuse. Underuse was referred to a couple of times in the study by Beth McGlynn in The New England Journal of Medicine of 1974, which showed that only 50 percent of recognized appropriate treatments were actually given to US patients in hospitals. This is 50 percent underuse.

LORNE TYRRELL: I always like to use an example. When I went through medical school, they used to say that a very good surgeon doing appendectomies might find that 70 percent of cases are truly appendicitis, and in 30 percent, the appendix is normal; and if only 5 percent of the appendixes you are taking out are normal, you are missing some and causing major problems. Maybe there should be an evolving concept of appropriateness. For example, I imagine that today there are many more ways of imaging the appendix, and the rate of unnecessary appendectomies should not be 30 percent. It might be 5 or 10 percent. There should be an evolution of appropriateness in the system.

TOM FEASBY: These studies have to be contemporary.

STUART MACLEOD: Just before you leave the room too pessimistic, Lorne, I remember when David Sackett left McMaster and went to Oxford in the 1990s. Having brought the benefits of evidence-based medicine to Canada, he arrived in Oxford quite certain that he would find that most of the practice in the Oxford hospitals was not evidence-based. They did a prospective study on the wards, greatly to the distaste of the British, who were pretty sure that evidence-based medicine was the work of the devil. To David's surprise, they found that 80 percent or more of the routine interventions on a general medical ward were, in fact, evidence-based. We get into trouble with words like 'appropriate' because there are degrees of appropriateness and there are degrees of certainty in what we do. I think if you use a reasonable lens, you find that most medical practice, at least in a tertiary-care hospital, is on the appropriate end of the spectrum.

GAIL WILENSKY: US statistics do not suggest that. I am not commenting on what is going on in Oxford, but in the US the appropriateness, although somewhat higher in the academic health centres, is about 60 to 62 percent.

CHARLOTTE ROBB: My name is Charlotte Robb. I am the interim CEO of Alberta Health Services and probably the least qualified person in this room to speak on clinical matters, but let me say that I strongly believe that we have a very committed and dedicated clinical team in this province. We are trying very hard to do the right things. I have come to know the healthcare delivery system in this province, and there is nothing simple about it. I propose that some of the behaviours we find may not be driven by clinical activities or knowledge, or lack thereof. They may be driven by the systems that we have. We had nine health regions in a fairly large geographical area with a fairly small population. This organizational structure did not lend itself to cooperation, collaboration, and sharing of knowledge and standards; and, in fact, we do not have common standards, so data is extremely important. We are looking at

clinical pathways. We are looking at patient pathways. We are trying to find ways to assist our very passionate, highly-qualified clinical team to do the right things, and I truly believe they want to do the right things.

TOM FEASBY: Thank you. Stuart.

STUART MACLEOD: I am going to follow up on what Gail said. The Sackett study at Oxford is a particular example that looked at an internal medicine ward, which is probably quite different from a surgical ward or a diagnostic imaging suite. The decisions are sometimes a little more clear.

TOM FEASBY: I thought those were the more complicated cases, with more problems.

STUART MACLEOD: Well, of course, you find a different outcome in an intensive care unit than in a general ward as well.

TOM FEASBY: These are challenging issues. Now, time is short, and I want to ask the panel if they have any final comments they would like to make, any issues we haven't dealt with that we shouldn't let go.

GAIL WILENSKY: I hope you don't become too discouraged by the challenges that we have all raised. The fact that you are here, you are asking these questions, and you have taken on the issue as one that you want to grapple with is a good part of getting a solution in place. Much of the task is to mobilize the forces. I very much agree with the comment that most clinicians want to do the right thing. You need to make sure that there are no artificial barriers preventing them from doing the right thing. In the US, financial incentives are often extremely perverse, and, fortunately, professionals most often do the right thing anyway; but it's really dumb to make it against an individual's own financial interest to be doing the right thing. Many times, there has been a lack of good information or easily accessible information. These are things that you can change. Consider whether or not there are better ways you can make use of information and better information you can be collecting. But a lot of it is just deciding you want to take it on.

TOM FEASBY: Thank you, Gail. I will wrap up the panel discussion by saying thank you to our panellists, Stuart, Gail, Terry, and Newell, and thank you to Andy Greenshaw, John Sproule, and Egon Jonsson for organizing a wonderful discussion this afternoon. I really appreciate the attention and active participation of our audience. Thanks very much.

ANDY GREENSHAW: Thank you very much for moderating that session so well. I would like to thank everybody for their participation. It has been well attended and a there has been very lively set of presentations, very relevant to what is going on in Alberta now. Thank you very much.

JOHN SPROULE: On behalf of Dr. Tyrrell and the IHE board and Dr. Jonsson, thank you very much for coming.

Appendix I – Program

INSTITUTE OF HEALTH ECONOMICS
ALBERTA CANADA

IHE INNOVATION FORUM I: PAYING FOR WHAT WORKS

Comparative Effectiveness of Health Technologies and Programs

December 2, 2008
Edmonton, Alberta, Canada
Sutton Place Hotel

Program

Master of Ceremonies: Andy Greenshaw
Associate Vice President (Research), University of Alberta

15:00 Greetings

Alberta Health and Wellness

15:15 Introduction to the IHE Forum

Lorne Tyrrell
IHE Board Chair and CIHR/GSK Chair in Virology

15:30 Comparative Effectiveness in Health Care

Gail Wilensky
Economist and Senior Fellow, Project HOPE

16:30 Investing in Innovation - Impact of Comparative Effectiveness

Newell McElwee
Vice President, Evidence-Based Strategies, Medical Division, Pfizer Inc.

Stuart McLeod
Vice President, Research Coordination and Academic Development, BC Provincial Health Services Authority

Terry Klassen
Director, Alberta Research Centre for Health Evidence (ARCHE)

17:30 General Discussion

Moderated by Tom Feasby
Dean of Medicine, University of Calgary

18:30 Reception and Dinner

Address by the Hon. Ron Liepert, Minister of Health and Wellness

Speaker Biographies

Dr. Andrew Greenshaw

Andy Greenshaw was appointed a faculty member in the Department of Psychiatry in 1986. Since that time, he has played a significant role in facilitating research at the University of Alberta as a faculty member in Psychiatry and a former Associate Dean in the Faculty of Graduate Studies & Research. He has been a Full Professor of Psychiatry since 1996 and Associate Vice-President (Research) at the university since 2004. Dr. Greenshaw is currently working with the Canadian Circumpolar Institute and the International Polar Year Secretariat to facilitate the Northern Research Strategy at the University of Alberta. He is also the Vice-Chair of the Board of Governors of the University of the Arctic.

Dr. Lorne Tyrrell

Lorne Tyrrell is the Chair of the Board, Institute of Health Economics. He is the CIHR/GlaxoSmithKline Chair in Virology at the University of Alberta. Dr. Tyrrell is also the Chair of the Board of the Alberta Health Quality Council and Chair of the Gairdner Foundation and a member of the Research Council of the Canadian Institute of Academic Research. In 2004, Dr. Tyrrell completed 10 *years as the Dean of the Faculty of Medicine and Dentistry at the University of* Alberta.

Dr. Tyrrell has won numerous awards at the University of Alberta (Rutherford Undergraduate Teaching Award, J. Gordin Kaplan Research Awards, and the University Cup). He won the ASTech Award for Research in 1993 and the Gold Medal of the Canadian Liver Foundation in 2000.

Dr. Tyrrell was appointed to the Alberta Order of Excellence in 2000, an Officer of the Order of Canada in 2002, and a Fellow of the Royal Society of Canada in 2004. He was awarded the F.N.G. Starr Award from the Canadian Medical Association in 2004 and the Principal Award of the Manning Foundation in 2005 for his work on the development of oral antivirals for the treatment of HBV.

Dr. Gail Wilensky

G ail Wilensky is an economist, and a Senior Fellow at Project HOPE, an international health education foundation. She is a Commissioner on the WHO's Commission on the Social Determinants of Health, co-chaired the recently-completed Department of Defense task force on the Future of Military Health Care, is Vice Chair of the Maryland Health Care Commission and serves as a trustee of the Combined Benefits Fund of the United Mineworkers of America and the National Opinion Research Center. She has recently been appointed to the Defense Health Board, advising the Department of Defense on health matters and to the Board of Regents of the Uniformed Services University of the Health sciences.

From 1990 to 1992, she was Administrator of the Health Care Financing Administration. She also served as Deputy Assistant to President (GHW) Bush for Policy Development, advising him on health and welfare issues from 1992 to

1993. From 1997 to 2001, she chaired the Medicare Payment Advisory Commission, and from 1995 to 1997, she chaired the Physician Payment Review Commission. From 2001 to 2003, she co-chaired the President's Task Force to Improve Health Care Delivery for Our Nation's Veterans. In 2007, she served as a Commissioner on the President's Commission on Care for America's Returning Wounded Warriors. Dr. Wilensky testifies frequently before Congressional committees and speaks before professional, business and consumer groups. She is an elected member of the Institute of Medicine and served two terms on its governing council. She earned her Ph.D. in economics at the University of Michigan.

Dr. Newell McElwee

Newell McElwee is pharmacist and epidemiologist with 20+ years of experience in Outcomes Research. He has been with Pfizer Inc. since 1998, and is currently the company's Vice-President of Outcomes Research, where he leads a group of scientists focused on health economics, patient-reported outcomes, quality of care, and health disparities. He has worked in the pharmaceutical industry since 1992. His educational background is in pharmacy (BS, University of Louisiana; PharmD, Mercer University) and epidemiology (MSPH, University of Utah). He also completed clinical residency and a research fellowship. One of his key interests is fostering the use of decision sciences to inform real world decisions, including internal investment decisions in industry, technology assessment and adoption, and health policy. Dr. McElwee has been a member of the AHRQ Effective Health Care Stakeholder Group, AHRQ study sections and has had leadership roles various professional societies related to Outcomes Research, including the International Society for Pharmacoeconomics and Outcomes Research, the Society for Medical Decision Making, and International Society for Pharmacoepidemiology.

Dr. Stuart MacLeod

Stuart MacLeod is Executive Director, Child & Family Research Institute, Professor, Pediatrics and Associate Dean (Research), Faculty of Medicine, University of British Columbia, and Vice President Academic Liaison & Research Coordination for the Provincial Health Services Authority. Before moving to Vancouver, he was Professor, Clinical Epidemiology and Biostatistics, Pediatrics, and Medicine at McMaster University, and a member of the Centre for Evaluation of Medicines at St. Joseph's Healthcare in Hamilton.

Dr. MacLeod received his MD from the University of Toronto in 1967 and completed postgraduate training in Internal Medicine (clinical pharmacology) in 1973 at McGill University and the Montreal General Hospital. He obtained a PhD in Pharmacology from McGill in 1972. From 1973 until 1986, he held various positions at the University of Toronto and its teaching hospitals. At the time of his departure from Toronto, Dr. MacLeod was Professor of Pharmacology, Clinical Biochemistry, Medicine, Pharmacy, and Pediatrics, and cross-appointed to the Faculty of Pharmacy. He was the Founding Director of the Division of Clinical Pharmacology at The Hospital for Sick Children. In the period January, 1987 through March 1992, Dr. MacLeod served as Dean of the Faculty of Health Sciences at McMaster University.

Internationally, Dr. MacLeod has coordinated several projects and taught in Africa. He has worked with international agencies and institutions including CIDA, IDRC, WHO and the Rockefeller Foundation. A sabbatical year in 1993 was spent at the London School of Hygiene and Tropical Medicine, Department of Clinical Sciences.

Dr. MacLeod's scientific interests are in improved understanding of the determinants of drug disposition and action, particularly in children and women. His concerns embrace the multitude of factors that influence optimal therapeutic drug use and the use of research findings to inform clinical and public policy.

Dr. Terry Klassen

Terry Klassen is Chair of the Department of Pediatrics at the University of Alberta and Director of the Alberta Research Centre for Health Evidence and Director of the Evidence-based Practice Center at the University of Alberta. He is a clinician scientist whose clinical base is Pediatric Emergency Medicine and has been active in Pediatric Emergency Research of Canada collaborating on a national research program involving randomized controlled trials, systematic reviews and knowledge translation. He is the leader of the Cochrane Child Health Field. Dr. Klassen's research has transformed the practice of Pediatric Emergency Medicine with his work in croup and bronchiolitis, having had randomized controlled trials published in the *New England Journal of Medicine*. He has a consistent record of national and international peer reviewed funding, along with over 150 publications, with a significant number being in top-tier medical journals.

Dr. Tom Feasby

Dr. Tom Feasby has been Dean of the Faculty of Medicine at the University of Calgary since 2007. Previously he was Vice-President of Academic Affairs at Capital Health and Associate Dean in the Faculty of Medicine at the University of Alberta. He is a practising neurologist and a health services researcher, who studies the appropriateness of health care interventions. Dr. Feasby has been a long time member of the Institute of Health Economics Board and was recently appointed to serve on the Minister's Advisory Committee on Health. He completed his BSc and MD at the University of Manitoba, followed by a research fellowship at Institute of Neurology in London, England and professorships in neurology at the University of Western Ontario in London, Ont. From 1991 to 2003, while leading the U of C Faculty of Medicine's Department of Clinical Neurosciences, he was also head of the Calgary Health Region's regional Department of Clinical Neurosciences, where he assembled an internationally-recognized clinical neurosciences group.

His record of research excellence is reflected in more than 100 research publications in areas such as neurologic diseases and the appropriateness of health care interventions, his supervision of numerous graduate students, and his involvement in professional societies and organizations including the CIHR Institute of Health Services and Policy Research, the Canadian Academy of Health Sciences and the American Academy of Neurology.