

The Value of Health Conference Summary

RAND Headquarters – February 3, 2007

Opening Remarks and Introduction

Dr. Dana Goldman began the conference by noting that the schedule of talks would revert to the originally planned schedule. In addition, Darius Lakdawalla would be presenting Dr. Kevin Murphy's paper, since Dr. Murphy himself would be unable to attend. He then asked Richard Suzman for some opening remarks. Dr. Suzman stated that NIH was interested in having a conference on assessing the value of health, because NIH wishes to incorporate economic valuations of disease in deciding how to allocate research funding. While policymakers and interest groups find the Murphy/Topel method of health valuation interesting, they are not entirely convinced by it. NIH is therefore interested in finding ways to strengthen research in the field. Moreover, this research could also be useful for WHO and national health accounts research.

Dr. Suzman then turned the time over to Dr. Goldman, who presented the results of a RAND study on health improvements that are currently in the R&D pipeline. As an example the case of a pill which could reduce caloric intake for humans. Such a pill would reduce the instantaneous hazard rate by 63% and increase life expectancy by 15 years, leading to a 13% increase in the elderly population and a 70% increase in health expenditures by 2030. The cost for an additional life year would range from \$9,000 to \$30,000 per year, which is quite cost-effective by present standards. The purpose of this example was to show that while anticipated medical advances may prove very cost-effective by current standards, they will nonetheless have large effects on health care spending.

There were two points raised in response to this example. First, Dr. David Cutler pointed out that the results of this study may be overly pessimistic, because the vast majority of health care spending (2/3) occurs at the end of life. Therefore, the additional health care spending attributable to life-extending technologies should only be the cost of the intervention itself. Second, Dr. Richard Epstein questions whether life extending technologies may affect the labor market, for example, by affecting retirement decisions. However, these labor effects may be mitigated by existing social structures, such as Social Security and Medicare.

The Economics of Fighting Disease: Balancing Costs and Benefits

(Paper by Drs. Kevin Murphy and Robert Topel, presented by Dr. Darius Lakdawalla)

The purpose of this paper was to investigate the gains of advances in health between 1970 and 2000. Using existing studies based on labor market decisions, the authors calibrate the value of a statistical life (VSL) at \$6.3 million. This value is then used to estimate the gains in life expectancy between 1970 and 2000. Overall, the author find that the value of these gains to be roughly \$95 trillion total, or, in annual terms, \$3 trillion per year (25% of annual GDP). Indeed, the value of the gains in health is roughly equal

to the value of material wealth gains over the same time period. An important implication of the paper is that controlling medical costs is important. For example, while the gross value of health gains is \$95 trillion, increases in medical expenditures are equal to roughly 1/3 of these gains. Moreover, consider a research initiative which, if successful, leads to a treatment that reduces the chance of dying from cancer by 10%. From a social perspective, the most costly outcome is not failure of the initiative to discover a treatment; rather, it is an initiative which discovers a treatment whose costs are higher than its benefits. Accordingly, medical R&D will be cost-effective only if accompanied by systems that ensure that only cost-effective treatments are implemented.

Several comments focused on the VSL the authors used, which is based on labor market studies of younger persons' willingness to accept compensation in exchange for riskier employment. A number of persons thought using alternative values for the VSL, such as the willingness to pay for improved mortality risk, could be useful. In addition, it was noted more generally that the VSL used in the paper is an "average" VSL, and may not reflect the VSL for persons of different SES status or race. As pointed out by Gary Becker, ultimately, the VSL itself is not the driving factor for the large values in health gains, rather, it is the fact that the total gains increase in proportion to population, and the US has a very large population. Therefore, the gross value of improvements in health is likely to be less sensitive to changes in the VSL. However, others pointed out that when the costs of medical care and the costs of behavioral changes in response to better knowledge about health risks are factored in, the *net* gains of increases in health may be sensitive to changes in the VSL. Moreover, as pointed out by Dr. Kip Viscusi, the value of net gains in health implicitly assigns all of the gross gains in health to medical care, which may overestimate the value of medical care. However, Dr. Cutler pointed out that, on average, medical advances have been cost effective.

In addition, there were several questions about why the VSL peaked when earnings peaked, particularly since under full annuity markets, consumption should be smoothed over time. While there was no clear answer, the general consensus seemed to be that this was related to the fact that while consumption can be smoothed, health (morbidity) is not smoothed over time. Dr. Emmett Keeler pointed out that the time series of the health weights used in the paper, which are analogous to Quality Adjusted Life Years, does not match the time series observed from surveys. In particular, morbidity seems to rise too quickly (health seems to fall) too quickly with age, when compared to survey data. In accordance with this interest in health weights, Dr. Suzman was interested in using the Murphy/Topel methodology to estimate the value of delaying/treating diseases that affect morbidity, such as Alzheimer's disease.

Social Determinants of Health

(Sir Michael Marmot)

Sir Michael began by stating that he believes that GDP is an inadequate measure of welfare, since there are several countries with high GDP and low life expectancy, and vice versa. There were several questions about this point. First, Dr. Goldman and Dr. Epstein suggested that while cross-country correlation between health and GDP may be

low, within-country correlation is likely to be very strong. Moreover, Dr. Epstein suggested (and Sir Michael concurred) that, in general, it would be difficult to imagine a situation where GDP increased but welfare fell.

Sir Michael then proceeded to show that health care spending and health are generally unrelated, particularly since the US spends a large share of GDP on health, and yet, has lower life expectancy than many countries which spend less. The difference cannot be solely explained by differences in wage distributions, as wealthy Americans are less healthy than wealthy British. It was pointed out within the United States, when one examines state-level data; there is also lack of correlation between health care spending and health. Given that health in the US is quite poor, Sir Michael believes it is inappropriate to develop “single-score” indices, such as the UN HDI, that rank countries based on health and GDP. In particular, the HDI does not seem to be reliable, since the US ranks 10th on this scale. While Drs. Tomas Philipson and Gary Becker suggested that perhaps it would do to re-weight the HDI, Sir Michael said that he prefers separate measures of health and economic prosperity.

More generally, Sir Michael stated that neither he nor most non-economists believe that prices can reflect a person’s valuation of a good. Along these lines, he also noted his problems with the Murphy/Topel paper. In particular, in his view, the paper implies that the value of health improvements increase with (a) population, (b) current health, (c) wealth, and (d), the closer in age a population is to the onset of disease. Sir Michael disagrees with (b) through (d), because they imply that resources to allocate health should be focused among the wealthy, elderly, and healthy. Therefore, the Murphy/Topel approach would exacerbate health inequalities. In response, Dr. Epstein asked whether Murphy/Topel really suggests that resources should be devoted towards improving health in already healthy countries, especially since the marginal impact of a dollar may be larger among countries with poor health.

Sir Michael also showed evidence from the Whitehall study that wealth is highly correlated with income, and again suggested that the Murphy/Topel approach would exacerbate health inequality by allocating resources towards improving health among the wealthy. He stated that, if there were only kidney available for a kidney transplant and there were two patients, one poor and one wealthy, the allocation of that kidney should not be based on income. Dr. Becker asked whether it would be a good idea to allocate the kidney to the rich man, if it were possible to properly compensate the poorer man. Sir Michael answered in the negative, stating that it would be difficult to properly compensate the poorer man, and that he was generally uncomfortable with people accepting money in order to avoid treatment. Dr. Victor Fuchs suggested that Sir Michael was uncomfortable with allowing people to make their own decisions, and Dr. Kahneman suggested that in general, it is very difficult people to determine the amount they are willing to accept to avoid treatment.

Sir Michael also presented some data suggesting that life expectancy has been diverging across countries. However, Drs. Becker and Philipson argued that their research shows that, when countries are weighted by population, life expectancy has generally been

converging. Sir Michael concluded by discussing his work at the Commission for the Social Determinants of Health, where he made two points. First, none of the commissioners believed that it was necessary to place a dollar value of health; rather they were concerned with a system of social justice. Second, he pointed out that many of the initiatives that the Commission believes could improve health relate to internal political reforms within a country, as opposed to transfers from wealth countries to poorer countries. Dr. Cutler suggested that one point to take from Sir Michael's talk may be that, unlike most goods, for health, the inputs themselves may be valued apart from output.

The New Economics of Terminal Care

(Drs. Gary Becker, Kevin Murphy, and Tomas Philipson; presented by Dr. Philipson)

This paper why the amounts that are spent for terminal care imply much larger values of a statistical life (VSL) than more common estimates of VSL using data from job market decisions or demand for safe products. The authors propose four reasons for the discrepancy. First, common estimates of VSLs examine small changes in risk, while terminal care involves large changes in mortality. Indeed, when death is the only other option, then, assuming no bequest motives, a person's willingness to pay for even marginal improvements in mortality should equal his entire wealth. Second, in making terminal care decisions, there is value to "hope"; the value of future survival impacts today's utility. This may particularly be the case because living longer increases the possibility that a cure may be found. Third, the quality of life for the terminally ill is lower, which decreases the costs of terminal care, since the value of the current consumption used to finance current care is lower. Of course, this may be mitigated by the fact that the benefits of terminal care are also lower, since the value of the increased longevity will also be lower. Finally, while standard VSL estimates examine the private VSL, terminal care decisions involve the social VSL. The social VSL differs from the private value because the social VSL incorporates the value family members and society in general place on the individual's life. Because terminally ill patients are able to use other person's resources to finance their care, the social VSL is therefore higher than the private VSL.

Dr. Neeraj Sood asked whether "hope" might be endogenous (and there its effects on VSL hard to ascertain), because family members who desire the terminally ill to live longer may pressure them into having hope. Dr. Keeler pointed out that empirically, the reverse seems to be true, as the terminally ill seem to live longer than what their family members would desire. Dr. Keeler also suggested that it would be worthwhile to examine whether the type of death may affect terminal choices. Dr. Cutler suggested that it would be useful to use cancer therapies to examine quantity-quality of life tradeoffs, since most cancer treatments, particularly chemotherapy, increase the length of life at the cost of reducing quality. Dr. Goldman also pointed out that standard VSL studies examine younger people, who have few assets, while terminally ill patients are older and usually have acquired more assets. Dr. Becker agreed that to the extent that people have unannuitized assets that will be lost if they die, then the value of these assets will affect their valuation of terminal care.

Dr. Fuchs said that the paper implies that, much like other goods, the willingness to pay for survival shows diminishing utility, so that the willingness to pay for survival is decreasing with current survival probabilities. He wondered whether this effect is solely due to incomplete markets. In particular, if people could ex ante arrange to be compensated when they are terminally ill, then the willingness to pay for increased survival might not decrease with current survival. In response, several people noted that it is unlikely that people can correctly estimate their willingness to pay for increased survival, or the amount they should be compensated if they enter adverse health states. Dr. Suzman stated it was unlikely that a persons' valuation for a disease 20 years prior to

diagnosis would also reflect his valuation when he actually contracted the disease. Similarly, Dr. Paul Brest noted that preferences are likely to change over time, and that people are unlikely to correctly anticipate the probability of getting a disease; thus, they are likely to under-invest in safety and over-invest in terminal care. Dr. Kahneman pointed that people are unlikely to think about events far in the future (focusing illusion). He also stated that the concept of willingness to pay is very context dependent, since a person's willingness to pay can depend on how he asked. Since it is unlikely that a person has a fixed willingness to pay, Dr. Kahneman asked whether there are other measures that might be used to value health. In response, Dr. Becker noted that the willingness to pay for health, or any other good, is always context dependent—the crucial question is whether people are able to anticipate changes in the preferences, and whether economists can understand the context under which people determine their willingness to pay.

Cost Effectiveness and Benefit Cost Approaches to Health Decisions

(Dr. Emmett Keeler)

Dr. Keeler began by telling of his role on a recent IOM Committee that issued a report on incorporating cost effectiveness analysis (CEA) into regulatory analysis. As a result of this work, Dr. Keeler now believes that most of regulatory studies should measure health gains by life years gained rather than lives saved (which he terms “deaths postponed” with the explanation that no one lives forever). Likewise, he promotes using a constant value of a statistical life-year over a constant value of statistical death postponed. With that as his introduction, Dr. Keeler outlined his presentations three central topics: measuring health in quality-adjusted life-years (QALYs); arguments against using willingness to pay (WTP) for deaths postponed in health decisions; and statistical values for QALYs.

Benefits-Cost analysis (BCA) and CEA use different methods to solve different problems. They consider allocative efficiency and technical efficiency, respectively: BCA asks the worth of a particular purchase whereas CEA addresses the best way to spend an amount of money. The relevant CEA term is the value of a statistical life-year (VSLY), the cost divided by life-years gained. Dr. Keeler argues against the use of BCA because of its reliance on WTP and the unreliability of using hypothetical survey questions to set such a value. For example, such questions may vary by audience: would a disabled population acknowledge the value of a preventive measure for a condition they are able to happily live with? Would a healthy population overstate their WTP to avoid such a condition? Dr. Keeler also disfavors the use of deaths postponed, particularly in person weights, but further states that the alternative – CEA – can not answer whether health gains are worthwhile.

The consensus in CEA is that all medical and public health interventions are rated in terms of how much they improve QALYs. After measuring health status consequences in a situation, QALYs are a way to evaluate them. But what are QALYs? Dr. Keeler defines them as a standard way of combining morbidity and mortality effects into a single measure. Each future year is weighted by the HRQL adjustment factor in addition to

discounting. The scale is from zero (death) to one (health), with sliding states of health between. Dr. Dennis Fryback added that he views QALYs as deliberative measures that represent capacity rather than a condition.

The use of QALYs is not without debate. Psychologists have discovered many contradictions in this methodology, following to Kahneman and Tversky's finding that people don't always follow expected utility theory. People have a hard time with questions that elicit the health adjustments needed to combine health and longevity. There are many methods for eliciting the quality adjustments (standard gambles, time trade-offs, visual analog scales, regression of health on illness, person tradeoffs, and WTP/total value of life) but each one has problems.

Most of the theory used in CEA deals with how to convert a stream of future health consequences into a simple utility function, typically one based on quality and quantity (remaining years) of life. Regarding quality, CEA usually utilizes average values for the probability of good health. In order for QALYs to be an accurate representation of preferences using utility functions, one must assume risk neutrality for years of life. Dr. Keeler argues that this is not realistic, as many people have diminishing returns in years of life. If the utility curve incorporates discounted years rather than simple quantity of years remaining, he states the results are closer to most people's preferences on length of life, and he thus endorses using discounted life expectancy in health decisions.

Turning to his second point, Dr. Keeler then argued against WTP for deaths postponed in making health decisions in three arguments. First, for interventions affecting chronic disease, deaths postponed are not well-defined. One can obtain survival curves for different populations, but the people within them are different. That is, the lack of a counterfactual complicates interpretation of findings.

Life extension as the main criterion is in better accord with Rawls' 'veil of ignorance' argument for utility maximization rather than WTP maximization. Because WTP is proportional to individual wealth, it gives more weight to the wealthy and short-lived than the poorer and longer-living population. This is a result of maximizing utility because both the poor and the short-lived have low marginal utility of money.

Dr. Keeler then presented the divergence between popular opinion and analytical reasoning. CEA findings often violate social judgments, as illustrated by the work of Erik Nord. For example, if a certain treatment has any substantial effect, social judgment research indicates that the sickness of the recipient is more important than how productive a treatment is. Likewise, people generally respond that the poor should get treatment rather than the rich.

Finally, Dr. Keeler addressed the difficulty is quantifying the VSLY versus the VSL. This finer calculation is necessary to value life expectancy gains for BCA. Some researchers ask patients their HRQL and their WTP for a lifetime cure, but using these values to calculate WTP for a QALY gives low values across studies. Possible reasons for the low values include income effects and the possibility that the respondent may not

live a normal lifespan as assumed by researcher. Using discounted life years gained is not perfect, but it improves calculations otherwise based only on quantity.

Agenda Discussion, Part I

At this time, Dr. Goldman paused the progression of speakers in order to ask for outstanding research questions before some participants excused themselves early from the conference. Dr. Richard Epstein raised the discounting issue: whether people should take a future stream of life years at a discount. He suggests researching a method to break up future time into shorter, more manageable bins, as he thinks the holistic approach is not conceivable. Sir Michael Marmot followed up to this comment with the suggestion of following a young person's value of a year of life through panel data.

Dr. Tom Rice suggested an agenda to refine how to put a value on lives. Most means are dependent on income, he noted, but it may be preferable to value lives without including a wealth effect.

Dr. Kahneman raised a concern that economists are too wedded to ex ante approaches. An ex post view may be beneficial. Researchers must consider multiple ways of asking questions; there may be alternatives to public welfare than aggregating individuals' views.

Comparing Estimates of the Measurement of Non-Fatal Health

(Dr. David Cutler)

Dr. Cutler's two goals for his presentation were to discuss measurement of health over time and to attribute changes to specific causes and diseases. Health is typically quantified as quality-adjusted life expectancy (QALE), where the quality measure is of course the complicating factor. There is no perfect measure for quality, but he suggests there are some good measures. A theoretically sound approach is to ask populations for preferences with a utility function survey. Such a question can be phrased in terms of dollars (i.e., How much would you pay to avoid spending the rest of your life with the functional ability of a stroke victim?) or trade-offs (i.e., How many years of life would you trade to avoid this condition?). The limitations to such a theoretical approach include the impact of medical advances, the use of counterfactuals, etc.

The measure Dr. Cutler prefers is based on a conception of health as the product of symptoms and impairments that occur with different chronic conditions. He notes that symptoms and impairments affect health in a relatively stable way over time, with changes infrequently (every 10 to 15 years). Diseases, on the other hand, could have variable relations with symptoms and impairments, suggesting a much more frequent update of these linkages (yearly). Dr. Cutler suggests inferring the disutility of symptoms and impairments to combined domains of health from surveys.

He critiques the use of overall health questions, whether on a Likert scale from poor to excellent or a rating scale. Self assessments of health are not reliable, with patients who

have overcome life-threatening illness often reporting the highest levels of health. Moreover, there is very little correlation between different measures. Dr. Kahneman critiqued that utility questions trading off life extension for quality of life don't apply to real life because they're not substitutes. Dr. Cutler responded that in chronic pain, the literature suggests that the two are about equally important.

Finally, Dr. Cutler suggests further exploration into the psychometrics of these questions to assess preferences and differences across groups. Dr. Suzman added that improvements would be useful to guide government spending across health conditions: cancer, arthritis, etc. Dr. Cutler noted in response that surveys that ask about domains don't generally ask about disease. Dr. Kahneman noted that well-being and health surveys should include a few questions of the other's emphasis.

Age Differences in the Value of a Statistical Life

Dr. W Kip Viscusi

Dr. Viscusi's presentation focuses on age differences in the VSLY through labor market data. Specifically, he asks how willingness to pay for risk reduction in the labor market varies with age. Valuing older citizens' life years at a discounted rate is politically unpopular as illustrated by the EPA's Clear Skies Initiative experience. Dr. Cutler noted and Dr. Viscusi agreed that income constraints could be entering into life extension preferences, as middle-aged individuals have more financial responsibilities. Sir Michael Marmot added that if analyses incorporate social differences then we would standardize preferences regardless of class.

Dr. Viscusi then compared estimates of VSL from varying model types in the literature. Using a life-cycle hedonic model, implicit discount rates have ranged in the literature from two to 17%, with the VSLY as high as \$950,000. Age-fatality risk interaction studies ignore variation in fatality risks and imposed a rigid linear structure, resulting at times in negative VSL for middle-aged workers. The literature utilizing this model type shows an indirect relationship between age and VSL. Life-cycle consumption models equate utility to consumption. The addition of consumption measures flattens the peak of VSL in middle-age. Age group VSL estimates (Dr. Viscusi's 2006 work) yield an inverted-U pattern of VSL with respect to age. Lessons from the literature translate to suggest the utilization of VSLs that are annuitized and age-specific. It is incorrect to assume a constant VSLY, but more research is needed.

From questions regarding the wealth impact on VSL, Dr. Viscusi posed the question, "Do airline passengers' lives have higher values than car passengers' lives?" The former tend to be 150% as wealthy as the latter, implying to economists that they do, but it can be sensitive if government money is involved. Political salience matters.

Dr. Epstein suggested similar work be taken post-retirement. If we look at the whole population – not just the workforce – would the VSL-age curve continue to decline? Dr. Epstein describes the possibility as a rational judgment.

Agenda Discussion, Part II

The remaining time of the conference was spent taking suggestions on future research questions on the value of health. Dr. Dennis Fryback noted that most of these models are based on preferences revealed by people in low-information circumstances in random moments. He proposes finding a methodology to prompt more deliberative decisions. This may relate to Dr. Epstein's previous suggestion of using retrospective information. Much of the future research agenda should be how people make decisions when they do have necessary tools. Not unaided. As to how researchers would approach this topic, we need to think hard about what people want to consider.

Dr. Epstein observed that the aggregate trends (save for obesity) are positive. We could learn lessons from our past experiences to replicate our previous successes.

Dr. Tomas Philipson suggested that we bring social estimates to par with estimates on private willingness to pay. The construction social estimates would help guide policy on elder care, etc. On a smaller scale within-family models could also be useful to help understand how they view and pay for healthcare for an elderly family member. He acknowledges this would be a daunting task, especially in health care. Dr. Cutler noted that the public seems to want to offset the social-wealth gradient. He is skeptical on social preferences because people often act without being informed. Dr. Philipson responded that behavior could be emphasized rather than survey information.

Dr. Goldman noted that the field is getting better at asking various questions. We might be able to compare behavioral and subjective analyses. As Dr. Fryback noted earlier, we could explore ways to teach people to answer questions better.

Dr. Epstein brought up intergenerational work in the area. He notes that all models seem to be single generation, and suggests we should look forward even though the tradeoff may not be one-to-one.

Sir Michael turned the conversation to social choices, suggesting that the compensatory approach to health allocation is a "rigged choice", meaning that offering poor people money does not provide them a choice. He questions whether people really believe health should be a privilege of the wealthy. This view prompted much debate amongst attendees. Dr. Jim Smith suggests that Sir Michael is focusing on willingness to pay rather than income.

Dr. Paul Brest observed that the primary issue of discussion isn't income but age difference. He believes identifying the value of life for an elderly person is challenging, and he has some skepticism on Dr. Viscusi's views. He notes that informed choice is terrific but we always frame it somehow. Dr. Epstein has done work in people's state of well-being but Dr. Brest maintains that this is different. He finished with the thought that if WTP is not the only way to go, then we need to look from the behavioral side.

Dr. Fryback responded that behavioral decision theory has a large literature, but the problem is measuring success. How do you know you helped someone? Researchers don't have a standard or a counterfactual to know if they've improved a situation. People need tools to think about problems.

Dr. John Graham added that it would be useful to have preference information between side effects of drugs and surgical risks versus alternative therapies. He noted that this is not all rational; there is systematic body of data on what people think when make decisions. He thinks there's an analog to Dr. Viscusi's body of evidence. Moreover, he'd like to rely on health data rather than extrapolating from labor.

Dr. Epstein moved the discussion back to the macro-level, saying he believes equal outcomes are desirable but that we can only do it through taxes. There is a long literature of impact of taxes on productivity, which he describes as more dramatic than expected. He notes that the American health system is not a market system. He suggests doing general equilibrium analysis on the health and the wealth sides.

Sir Michael responded that life is valuable and trying to figure out what the policy implications are doesn't make you ignore equity or focus only on medical care. He says we should, however, look at the broad implications for society. We lack a methodology for that.

Dr. Cutler named eliciting average values of health states as an area in need of progress. The place where people consider this is in the context of treatment decisions. Supposing a researcher informs people on decisions, we ought to compare more and less intensive things. What do we need in order to get a good decision? Perhaps having good information is as important as having a condition is on responses. We could ask experts what they would if they faced certain health states. He finished with the thought that we may not know what the right answer is but we can compare and see what is "out there".

At this point, Dr. Goldman thanked all for their attendance and input on this issue. The conference was adjourned.

A G E N D A
The Value of Health Conference
RAND Headquarters
February 3, 2007

- 8:00 am *Breakfast*
- 8:30 Opening Remarks
- 9:00 Darius Lakdawalla**
The Economics of Fighting Disease: Balancing the Costs and Benefits
- 10:00 *Break*
- 10:15 Sir Michael Marmot**
Social Determinants of Health
- 11:15 Tomas Philipson**
New Economics of End of Life Care
- 12:15 *Lunch*
- 1:00 Emmett Keeler**
Cost Effectiveness and Benefit Cost Approaches to Health Decisions
- 2:00 *Break*
- 2:15 David Cutler**
Comparing Estimates of the Measurement of Non-Fatal Health
- 3:15 *Break*
- 3:30 Kip Viscusi**
Labor market evidence:
Age differences in the value of statistical life
- 4:30 *Break*
- 4:45 Discussion of NIA Value of Health Agenda**
- 5:30 *Adjourn*