Cost-effectiveness analysis and its application for policy evaluation for medicine or public health

Yasushi Ohkusa

Senior Researcher, National Institution of Infectious Disease

Tamie Sugawara

Graduate School of Comprehensive Human Science, University of Tsukuba

Abstract

In comparison to the policy for other field, the policy for medicine and public health is to consider the value of life or the value of the quality of life. Quality of life is very well known as a concept of QOL. Also, Quality Adjusted Life of Years (QALY) which integrates QOL over life of years is widely used as a measure of the value of life. Cost-effectiveness analysis for medicine and public health adopts two approaches to incorporate value of QOL or QALY. We summarize those advantage and disadvantage briefly at first. Unfortunately, cost-effectiveness analysis has not been committed and operated as an official rule for the method of policy evaluation for medicine or public health in Japan, yet. Thus we show some researches about it which examines *ex post* or *ex ante* policy evaluation using cost-effectiveness analysis. In other countries, some political decision making in medicine or public health is based on cost-effectiveness analysis. However, the pressure of the financial deficit will require more accountability about evidence. Therefore, cost-effectiveness analysis must be more important even in political decision making in medicine or public health in Japan.

I. Some Features of Cost-Effectiveness Analysis in Policy for Medicine or Public Health

It is a well known fact that cost-effectiveness analysis is a basic tool for policy evaluation even in Japan as in this issue. This paper focuses on policy for medicine or public health and summarize and discuss about cost-effectiveness analysis.

Nevertheless, the policy for medicine or public health share the same grounds with other

public policy such as road construction, agriculture or defense. It needs an enormous cost which is financed mainly by tax or premium other than copayment. On the other hand, the policy for medicine or public health has inherent features. Namely the policy for medicine or public health is closely related with life and its quality. Of course, other policies may take life and its quality into consideration. For instance, the policy for highway construction must evaluate death or disability due to traffic accidents as its negative benefit or cost, or the policy for the airport construction must cost health problem by noise pollution. However, life and its quality is definitely the most important issue for evaluation of medicine or public health.

Then how should we evaluate life and its quality and how do we take it into cost-effectiveness analysis. This is the important feature in cost-effectiveness analysis of the policy for medicine or public health. At first, we have to evaluate life and its quality numerically. Unless we evaluate it numerically, we cannot compare the quality of life of appendicitis patients and patients suffering from heart attacks. It is very important and it is a big issue in health economics, but it is not well known in other fields of economics, and thus we discuss it at first.

Then we have to use this numerical evaluation of life and its quality for political decision making. There are mainly two types of methods. One method is direct monetary evaluation of life and its quality. This method ignores the features of life and its quality and treats them as other goods. Thus we can apply the cost benefit analysis as other fields in economics other than medical or public health do, and we can compare medical or public health to highway construction, environmental policy and so on.

Another method is to avoid direct monetary evaluation for life or its quality, and use another unit other than monetary unit. Thus we can not compare the evaluation by this method with other fields since it emphasizes the specialty in medical and public health. Below, we explain briefly three features of cost-effectiveness analysis in the policy for medicine or public health, and then show its applications.

II. Numeral Evaluation for Life and Its Quality

Cost-effectiveness analysis in policy for medicine or public health sometimes evaluates life and its quality numerically (QOL). QOL measures severity of illness or disability which defines death as zero and perfect health as one. There are several ways to measure QOL such as EuroQoL(Brooks(1996)), SF36(Ware and Sherbourne(1992)), and HUI(Le, Buron, Costet, Rosman and Slama(2002)). These methods measure the total situation of health and apply to general health people in addition to all kinds of illness or disability in principle. On the other hand, disease specific measures are also well developed. For example,Q-tility(Weeks et al(1994)), FLIC(Functional Living Index-Cancer), CARES(Cancer Rehabilitation Evaluation System) measures the QOL of cancer patients, and DQOL(Diabetes Quality of Life Measure) is for diabetes.

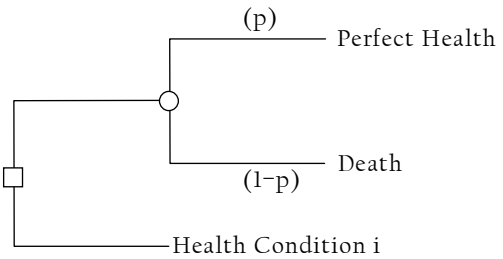
QOL is measured by question are and it asks the individual's physical, psychological and/or social function as several attributes. For example, Euro QoL has five attributes such as moving, usual activity, self care, anxiety/depression, and pain. On the other hand, SF36 has 36 attributes. Each attributes has three to five levels and the respondents select one level.

These measurements aim to evaluate QOL in total. Conversely, QOL in specific condition is measured by mainly Standard Gamble (SG), Time-Trade Off (TTO), or Rating Scale (RS). SG asks the probability of being in perfect health (QOL=1) and otherwise die (QOL=0) if intervention is implemented which intervention and no-intervention are indifferent as Figure 1 and following trivial equation

$$QOL\ in\ a\ situation = pQOL\ in\ perfect\ health + (1-p)QOL\ in\ death = p \tag{1}$$

based on the expected utility theorem which is very well known in economics. SG has a firm foundation in economics and it is the best way to measure QOL, but it seems to be very difficult to understand for the respondents.

Figure 1 : Standard Gamble



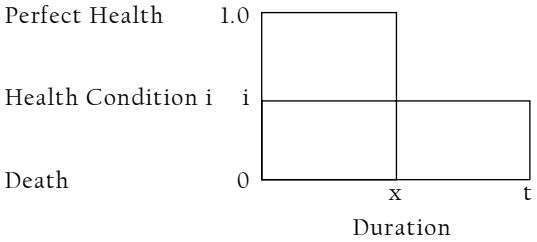
On the other hand, TTO asks years of perfect health which indifferent with a fixed years (say, 10 years) under a situation as shown in Figure 2. Since the equation

$$QOL\ under\ a\ situation \times t = QOL\ under\ perfect\ health \times x = x \tag{2}$$

is held and thus QOL under a situation is measured by x/t. This method also seems to be difficult as SG.

SG and TTO focus on one aspect of utility but ignore another aspect. Namely, SG focuses on the avoidance of uncertainty and thus it measures risk aversion. However, it lacks duration or time and thus it ignore time discount. Conversely, TTO compares two durations including time discount and ignores risk aversion. Therefore these two methods are not complete and we have to recognize that they have limitations.

Figure 2 : Time-Trade Off



RS measures QOL in a very simple way. Namely, it asks QOL directly as a point more than 0 which corresponds to death, and less than 100, which corresponds to perfect health. Figure 3 indicates an example of RS which uses a thermometer. This method is easy to understand and widely used.

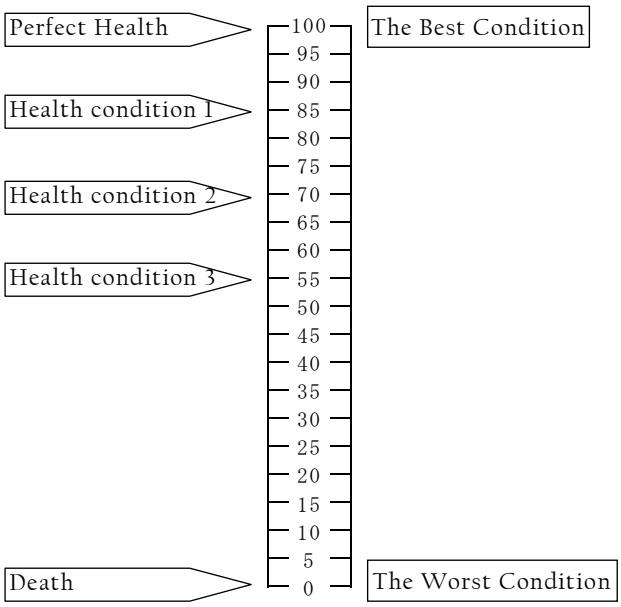
The most widely used method to measure QOL is self-evaluation. This method is adopted internationally and has been used for a long time. In Japan, Comprehensive Survey Living Standard, which is a national survey and is used every three years. This survey asks "What is your current health condition ?" and asks the person to choose one choice of very good, good, fair, bad, and very bad. However it does not specify duration and reference group to compare it with. In other words, the meaning of "current" seems to be ambiguous. It may be this moment, today, or these days. Since it is also unclear whom to compare with, some one may compare their condition with others in the same age group or other people may compare with their own health in the past. Internationally, the question are clearly defines the duration to refer to as one year, half a year or one month, and asks them to compare their health with other people in the same age group(Van Doorslaer, Wagstaff, et al (1997)).

Since QOL measures instantaneously quality of life, it lacks the concept of duration. Thus we have to integrate it over time so as to evaluate the individual's total quality of life such as

$$QALY = \sum_{s=0}^{n} \beta^{s} QOL_{t+s} \tag{3}$$

where β the discount factor for discounting future conditions. It is well known and applied as Quality Adjusted Life of Years (QALY).There is some controversy surrounding the discount factor. Namely it should be set as interest rate so far, but the insistence that health condition should not be discounted and so β=1 is emerging (Marthe, Joanna, Louise, and Milton(1996)).

Figure 3 : Rating Scale



Sometimes QALY is called as utility in health economic literature. If the source of utility is only health, and utility is described as a linear function of health, QALY coincides with lifetime utility. However, if not. QALY and lifetime utility may diverge. We have to remind ourselves that term utility may be different from other fields of economics other than health economics.

III. Monetary Evaluation of Life and QOL

So as to achieve a more efficient distribution of medical resources and then to constrain unlimited increasing medical costs, cost effectiveness analyses has been introduced, especially in the approval or price setting of new drugs. This idea can be described by the following simple equation:

$$ICER = \frac{C_1 - C_0}{E_1 - E_0} < \mu \tag{4}$$

where ICER is the incremental cost effectiveness ratio, C is cost and E is effectiveness represented by QALY. The subscript zero in the equation means traditional or currently standard drugs or treatments so far and new indicates a new drug or treatment.

It is well known, that this criterion has been already applied in Australia and the Canadian province of Ontario. Though there is no such movement for this in Japan, it needs to be introduced in the near future. Thus, basic research about it is necessary for its introduction; or moreover for the argument in its favor and preparation for its introduction.

First, it is very important to know the number of μ. For instance, it is 50 thousand US dollars

in the USA (Goldman et al(1992)), 20 thousand Canadian dollars in Canada (Laupacis et al(1992), 30 thousand pounds in the UK [1)], 36 thousand Australian dollars in Australia(George, Harris and Mitchell(2001)) and 30 thousand Euros in the Netherlands (Welte, Postma, Bos, van Alphen(2002)). However, the number itself is not so important; in fact, in Australia where this criterion is used in actual decision making for the approval of new drugs, an application is not accepted if the ICER is more than 69 thousand Australian dollars. Conversely, they never reject it if it is less than 18 thousand Australian dollars. Similarly in the UK, NICE accepted riluzole, a drug to treat people with motor neuron disease, even though its ICER was more than 40 thousand pounds. However even though it has some ranges, there is not any concrete foundation to decide the criterion and its range; it is usually based on "expert opinion."

On the other hand, there is an alternative idea for deciding the number μ from the point of view of human capital or foregone income. This idea is typically used in compensation for loss in court, or as a premium in wages for risky jobs. Tolley, Kenkel and Fabian(1994) survey the empirical research regarding this and find that it ranges from 70 to 175 thousand US dollars. In particular, Cutler and Richardson(1997) employ 100 thousand US dollars.

In principle, however, this number must be decided by the preference of the general population, who are sovereign in a democratic society. The "expert opinion" of a few experts like Goldman might not be appropriate to decide it. Hence, we need a social survey to discover the number in a given society.

This paper undertakes such a survey and analyses its distribution and characteristics. Especially, we consider whether a method based on human capital or foregone income is appropriate. As mentioned before, there is no previous research to decide this critical value from a survey of the general population survey. Therefore this paper should contribute to this field.

III.1. Data

The data set used in this paper is based on a survey carried out in March 2002 in Japan. This survey was implemented through a research company with which the sample households contracted. The households were randomly selected. However, the research contract might yield a sample selection bias and thus some caution is needed in estimations and interpretations. The number of surveys distributed was 900, of which 783 were returned. We use 1119 individuals for our analysis.

The questionnaire about μ was as follows;

Let's assume that there is a patient who in their current situation will die today. However, a new treatment can extend her/his life for one year in perfect health. What amount should society pay for this new treatment? Please write a particular number.

[1)] The number in UK is according to http://bmj.com/cgi/reprint/323/7325/1324/a.pdf.

The response field is constructed to represent from 0 to 100 million yen.

III.2. Summary Statistics

Summary statistics are summarized in Table 1. Though the average μ is 6.29 million yen (50.32 thousand US dollars), the standard deviation is so huge (18.12 million yen) and the median is only one million yen (8 thousand US dollars).

Table 1: Summary Statistics

	Average	Std. Dev.	Minimum	Maximum
μ	628.9751	1812.207	0	9999
Age	44.37182	19.49136	13	99
Female Dummy	.5294118	.4993011	0	11
Household Income	700.3255	410.6517	0	3250
Financial Net Asset	−136.0333	1468.617	−4500	4450
Real Asset	.8254805	.3796145	0	1

Note:Real asset is dummy variable for whether they have own their house or live in rental house. Unit for μ, household income and financial net asset is ten thousand yen (eighty dollars). Sample size is 1119.

Since μ is heavily skewed, some results may be dependent on whether we use its average or median as the representative number. If we more or less suppose a democratic decision making process, we have to use the median. In the following analysis, we are reminded of such a skewed ness in the distribution of μ.

Figures 1 to 6 illustrate the distribution according to some characteristics of the respondents. In these figures, the horizontal axis measures the log of μ and the vertical axis indicates the frequency by Kernel density estimation. Figure 1 shows the whole distribution and we can see its peak is around five, which is about 1.5 million yen (12 thousand US dollars). The distribution by age (less than 65 years old, or more than), gender, income (less than average, or greater than), net financial assets (less than average or greater than), and real assets (whether they have their own residence, or they live in a rented accommodation) are illustrated in Figures 2 to 6. It can be seen that the elderly evaluate μ more than the younger, males have a greater concentration around peaks, and the peak in higher income or asset households is higher or moves to an upper level more than lower income or asset households.

III.3. Estimation

To confirm the intuition from these figures, some estimations are performed as follows. The dependent is the log of μ, which is the same in the figures. The explanatory variables are age,

gender, household income, and net financial and real assets.

We adopt two estimation methods, the interval method and the quintile method, to take into account these features of the characteristics on the distribution. The former is basically a linear estimation that represents the average of the distribution in some situations. However, it treats that a zero or a more than 100 billion yen response indicate a negative or more than 100 million yen, because the response field is not constructed to represent negative or more than 100 million yen. Never-the-less, this estimation is linear for the responses which are more than zero and less than 100 million yen. Conversely, the latter estimation procedure focuses on their median in some situations. As emphasized before, the median is the important characteristic in a democratic society, and it reduces extremely high responses to a level for evaluation, which is confirmed in the figures; a comparison with the linear model including the former estimation.

The estimation results are summarized in Table 2. The estimation results are similar in each of the tables. Namely, for every one year older, μ declines by 1.3 to 2.2%. This age effect seems to be large. If we add the quadratic term of age as an explanatory variable, it is not significant, which is not shown in the tables. Hence, μ decreases monotonically due to age.

Table 2: Estimation Results

	Interval regression		Quantile regression	
	Coefficients	p-value	Coefficients	p-value
Age	-0.0222687	0.057	-0.0130291	0.000
Female Dummy	-0.1083792	0.629	-0.1064494	0.168
Household Income	0.2793647	0.010	0.2367467	0.000
Financial Net Asset	-0.0000312	0.684	0.0000149	0.574
Real Asset	-0.4822878	0.120	-0.1319622	0.217
Constant	3.231015	0.000	3.726981	0.000
log likelihood	-2371.7455			
pseudo R^2			0.0021	
p-value for Wald test	0.0033			

Note:The null hypothesis in Wald test is that all coefficients except for constant term are zero.

The coefficients of household income are significantly positive and the elasticity is 0.24 to 0.28. In other words, increasing income causes μ to rise, but it is not responsive. If μ is based on human capital or foregone income like Tolley, Kenkel and Fabian(1994) or Cutler and Richardson(1997), since these are proportional to current income after controlling age, the elasticity should be one. However, χ^2 statistics on the interval estimation is 44.36, and F statistics on the quartile estimation is 395.19 for the null hypothesis that the elasticity is one and that these probabilities are far less than 0.01%. Therefore, we can reject the null hypothesis and so not calculate μ based on human capital or foregone income.

III.4. Discussion

μ : the critical value of expenditure per QALY which a society can tolerate or accommodate, gives some standard for effectiveness regarding medical services. Cutler and Richardson (1997) use 100 thousand US dollars as μ and argue its effectiveness for medical services.

However, there is some confusion involved in this idea because 100 thousand US dollars is calculated based on foregone income, and thus it may be the value of life, but it is not an upper limit which the society tolerate for medical services [2)]

In reality, medical services must be allocated a value less than that of a life in their optimal resource allocation. It should not be optimal that all resources are allocated to medical services and nothing to other activities such as consumption and leisure.

Moreover these differences between the critical value of expenditure per QALY which the society can tolerate and the value of a life could reflect the difference between WTP and WTA, which is well known in empirical research. In general, WTA tends to be larger than WTP in a similar situation (Horowitz and McConnell(2002), Walton, Graves, Mueser and Johnson, Fries, and Banzhaf(2002), Salkeld, Randy, and Short(2000), Hanemann(1991)). In this context, the value of a life is WTA, which amounts to accepting death, and the expenditure per QALY which a society can tolerate, of course, is WTP. Therefore, it is somewhat natural that the latter is far less than the former.

In addition to these rather logical arguments, we add some evidence of differences from empirical investigations from a general population. As emphasized before, since the income elasticity of μ is significantly less than one, such a difference is very important even in a theme of economic sense. Therefore, we have to use one million yen as the median or 600 million yen as an average as for μ in Japan.

Conversely, since we confirm that μ has quite a skewed distribution, the characteristics of the distribution, such as an average or even a median, seem to be meaningless. Hence the criterion of cost effectiveness analysis in eq.(1) should change to

$$Prob\left[\frac{C_1-C_0}{E_1-E_0} < \mu(i)\right] > \alpha \tag{5}$$

Where $\mu(i)$ is μ for the i th person.

III.5. Further Research

This paper considers that the critical value of expenditure per QALY that a society can tolerate is the most important parameter in the cost effectiveness analysis. As a result, its

[2)] Another problem in Cutler and Richardson (1997) is whether the extension of life expectancy is due to medical services. This dubious assumption is imposed implicitly, but it might be too simple. At least, some contribution from improvements in nutrition and/or a great development in the public health services needs to be taken into account.

average is six million yen (48 thousand US dollars) but its median is just one million yen (eight thousand US dollars), and thus it has a very skewed distribution. Additionally, age and household income affect it significantly.

However, there are many unsolved problems remaining in this paper. First, the questioning procedure has some faults. As explained, the amount was asked directly, but this procedure is frequently misleading, resulting in some extreme responses and a high correlation with income (Donaldson, Brich and Gafini(2002)). In fact, an improved procedure has already been suggested; recently conjoint analyses have been widely applied to this field. If applied here, the significant correlation between μ and income could disappear.

Moreover, there may be a problem in that a question in the survey includes reference to any other social cost other than medical costs. To confirm that possible effect, we need to perform a comparison survey limited to medical costs. It should be noted that such costs are sometimes included in the analyses of the cost–effectiveness from a societal view point. For instance, the cost for nursing care, long term care or cost of lost opportunity by the family care giver. Hence, incorporation of these costs is appropriate from a societal point of view.

Finally, it appears to be inappropriate that the question is asked only about two reference points, i.e. perfect health status and death. If we used some intermediate status, we could define μ at a marginal rather than in an average like eq. (1), and thus we could define it based on QOL levels. However, such an intermediate status needs a QOL evaluation for it. This implies adding more unknown parameters into the analysis, and it would make the survey far more complex. Though this paper avoided this complexity by using just two polar cases; a future survey might be able to solve this problem.

Research regarding a critical value of expenditure per QALY which a society can tolerate, or accommodate, has just started; more following such research is necessary. Following research projects should find more reliable parameters for the distribution, and by this it would possibly confirm the robustness of our results.

IV. Non–Monetary Evaluation of Life and QOL

We sometimes face an emotional contradiction to a monetary evaluation of life or QALY. Moreover there is no strong evidence in its value except for Japan. Hence it has not yet reached an agreement in academic and consensus in public sentiment. So as to avoid such a situation, life or QALY are often evaluated by non–monetary units in the medical and public health field. For instance, cost–effectiveness analysis evaluates life as a time unit or outcome as a physical unit other than QOL, and cost–utility analysis use QALY itself as an outcome measure.

IV.1. Cost–Effectiveness Analysis in Narrow Sense

The physical units which measure the outcome in cost–effectiveness analysis are test result such as blood pressure, blood sugar and so on, or survival rate and life expectancy. Especially, it is so called cost–effectiveness analysis in a narrow sense which means the effectiveness is defined by only physical units. Meanwhile cost–effectiveness analysis in a broad sense includes any definition of effectiveness such as monetary or utility units. Physical units are widely used in the studies in medicine. For example, clinical trial usually measures its outcome as a physical unit. In this sense, its accumulation of facts is enormous. We can use such results to investigate in cost–effectiveness analysis, and thus it is the most advantageous of it. If we can get the information about cost, we can combine them and obtain the consequence of cost–effectiveness. This method, which means to get the data of outcome and cost separately, is so called scenario study, and it is the most widely used in cost–effectiveness analysis in a broad sense. However, we have to remind ourselves that usually cost and effectiveness are correlated, particularly negative correlated, and thus the synergic study has bias.

IV.2. Cost–Utility Analysis

QALY is used as outcome measure in cost–utility analysis which has advantage in integration with QOL and life of years. The most important shortcoming in cost–effectiveness analysis is that a little improvement in the test result may not affect QOL or life of years. For example, antihypertensive drug can reduce blood pressure, but patients cannot feel to improve heartthrob comparison with the past, QOL must not be developed. In other words, it must be meaningless if such a drug or treatment cannot raise QOL. In this sense, cost–utility analysis is based on the patients' perspective, instead of the doctors' perspective in cost–effectiveness analysis in a narrow sense.

On the other hand, the most important weakness in cost–utility analysis is difficulty in the estimation of QOL and even expected life of year. As explained before, it is usual that the measurement of QOL itself is difficult and typically it has a wide variety in its distribution. Moreover, concerning life expectancy, studies which have a long observation period, say ten years or more, are rare and thus survival rate is typically is used as a proxy of it. Due to these problems, cost–utility analysis is not applied as often as cost benefit analysis or cost–effectiveness analysis in a narrow sense.

IV.3. Statistical Problem

As explained, cost and outcome have different units in cost–effectiveness analysis in a narrow sense and cost–utility analysis and thus we have to evaluate them as a ratio instead of

net subtraction as we would a net benefit in cost benefit analysis. Ratio and subtraction seem to be similar, but there is a big difference statistically. Namely, since cost and outcome are assumed to follow certain types of distribution, typically normal distribution and thus the probability of that cost or outcome being zero may be more than zero, the distribution of ratio may not exist, even though the distribution of subtraction is always defined. It is not a rare case that the outcome is zero in the field of medicine or public health (Lothgren and Zethraus(2004),Willan and O'Brien(1999), Gardiner, Huebner, Jetton and Bradley(2000))

So as to fix this problem, some methods are proposed. The most simple way is the delta method. It is well known in economics which Taylor expansion to ratio and linearized it, and then evaluate it approximately. However, nevertheless, linear approximation for ratio seems to be inaccurate. Thus this method usually is not applied to cost-effectiveness analysis.

Another method is the Fieller method (Willan and O'Brian(1996), Fieller(1954)) and it is a traditional method to test for ratio. It uses χ^2 distribution to test and thus it does not approximate the distribution, but it assumes a normal distribution for cost and outcome. These assumptions may not hold if cost or outcome does not follow the normal distribution. In particular, it is well known that cost has fatter tail than normal distribution. In that case, this method does not seem to be appropriate.

Recently, bootstrapping is often used so as to fix this problem. This method does not assume cost or outcome will follow certain types of distribution and thus it has advantages over the above two methods. Moreover, though the above two methods assume two distributions for cost and outcome and do not consider the distribution of its ration directly, bootstrapping can consider the distribution of a ratio directly. Therefore it is more accurate and appropriate for cost-effectiveness analysis in this sense.

IV.4. Fundamental Problem

Let we get data cost and QOL or outcome which is measured by physical unit fortunately. Moreover, we can the get average or distribution of cost effectiveness ratio. For example, Wiiliam(1985) shows that so as to gain one more year of life, smoking cessation campaign needs 170 ponds, pace maker require 699 ponds, femoral head and neck replacement surgery(752 ponds),bypass surgery for non-severe angina pectoris(1042 ponds), control of blood-cholesterol level(1695 ponds), kidney transplantation(3030 ponds), screening of breast cancer (3448 ponds), heart transplantation(5000 ponds), bypass surgery for severe angina pectoris(12658 ponds), extracorporeal dialysis(14706 ponds). This list of cost-effectiveness ratio in each intervention is called a league table. Then how should we use this league table to actual decision making in medicine or public health. For example, which intervention should be implemented and which intervention should be rejected ?

In principle, it appears to be rational decision making to accept the intervention in order of

small cost per one year of life gained or low cost-effectiveness ratio until the budget is exhausted. In the above example, we have to approve or cover by health insurance in the following order; advertisement for smoking cessation, pace-maker, femoral head and neck replacement surgery, bypass surgery for non-severe angina pectoris, control of blood-cholesterol level, kidney transplantation, screening of breast cancer, heart transplantation, bypass surgery for severe angina pectoris, extracorporeal dialysis. In this case, dialysis in the hospital may not be covered by health insurance, and thus the total cost should be paid by the patient, even though the patient will die within a few days if they cannot use dialysis.

However, this idea seems to be too simple to implement. Namely, can we compare two treatments among different diseases? In the above example, CABG is certainly a higher cost per one year of life gained than a smoking cessation campaign, but the campaign cannot treat heart attacks. In other words, CABG and smoking cessation campaign cannot be substituted for each other, but we have to think of them independently even though they are both things in medicine or public health. Therefore, we must limit comparing the its cost effectiveness to substitutable drugs, treatments, or policies.

We have to remind ourselves of one more thing about policies for medicine or public health. That is that we always have status quo for drugs, treatments or policies even if it means we in fact "do nothing". Obviously, status quo and new drugs, treatments or policies are substitutable. Thus we have to compare these two alternatives and discuss which is better from the view point of cost effectiveness. ICER as mentioned in the previous section embodies this idea. If we do not have any particular drug, treatment or policy, we set C_0=0, E_0=0 and thus ICER is

$$ICER = \frac{C_1 - C_0}{E_1 - E_0} = \frac{C_1}{E_1} \tag{6}$$

It goes back to being a simple cost-effectiveness ratio in the league table.

If a new drug, treatment or policy is more effective than $E_1 > E_0$, but it is cheaper than the status quo $C_1 < C_0$, we, of course, always have to adopt the new one. However, if $E_1 > E_0$ but it is more expensive than $C_1 > C_0$, then we need some criterion. In fact, we face this situation in almost all of cost-effectiveness analysis. If we adopt new alternative, ICER is less than criterion, and reject it otherwise. In the case of cost-utility analysis, since effectiveness is defined by QALY, this criterion is surely μ which we discussed in the previous section. Therefore, we can conclude that cost-utility analysis becomes cost-benefit analysis at the time of decision making. Even in the case of cost effectiveness analysis in narrow sense and effectiveness is defined by expected years of life or survival rate, it also goes back to cost-benefit analysis at this stage.

V. Application of Cost–Effectiveness Analysis to Policy for Medicine and Public Health in Japan

Unfortunately, cost–effectiveness analysis has not been adopted in practice even in the case of approval of new drugs in Japan as mentioned before. Moreover, it also has not been declared formally for use it as a formula in other policy evaluation for medicine or public health. However, it does not mean that cost–effectiveness analysis is useless in policy evaluation for medicine or public health. In the near future, financial pressure, especially pressure for effective medicine, will play a key role in its decision making. If the new policy lacks a foundation in cost effectiveness, it will face strong opposition. Conversely, if a policy has strong evidence of cost–effectiveness but the government does not adopt it, the government will be condemned harshly due to their inaction.

As in other policies, cost–effectiveness analysis was applied to some policies for medicine or public health so far, and some examples are shown in the following. The first example is application for a vaccination program and the second one is for a smoking cessation program.

Vaccination program is one of the most important issues in which cost–effectiveness analysis should be applied. Since it is expensive and it has some side effects, the scientific evaluation and comparison between its cost and benefit is key for both those for and against it. In other countries, cost–effectiveness analysis has played a key role in the decision making for vaccinations (Stratton, Durch, and Lawrence(2000)), but in Japan, it has just started.

We show cost–effectiveness analysis for subsidy of influenza vaccination for the elderly. In November 7, 2001, the vaccination law was reformed and it started to subsidize the influenza vaccination for the elderly. This policy should be confirmed by the cost–effectiveness perspectives because it costs very much. It was published originally as Ohkusa(2005).

V.1. Example 1:Policy Evaluation for the Subsidy of Influenza Vaccination for the Elderly

V.1.1. Material

The data about copayment of influenza vaccination, population, and shot rate of the elderly in the '01/'02 and '02/'03 seasons was surveyed by telephone interview to correspondents in the local governments of metropolitan and 12 other big cities in Japan. This survey was performed by the author.

The cost of the copayment is determined by these local governments every year, and the excess cost more than the copayment is subsidized by the central and local governments directly to the medical institutions. The total cost of the vaccination, which is charged by the medical institution to the elderly and local governments, is decided through negotiation between the local governments and physicians' association in each city. Unfortunately, it is not announced

publicly. In other words, we can only know the copayment for each year and each city, while the total cost and thus the amount of the subsidy are unknown. In this sense, the total cost includes all components of items for the vaccination and the profit of medical institutions.

The mortality rate due to pneumonia or influenza was obtained from Vital Statistics of Population in 2002 and 2003. The data of total population was obtained from the National Population Census in 2000.

V.1.2. Method

(1) Estimation

Estimation is performed with the following two parts. At first, we examine the impact of the variation of copayment on shot rate. Let $R_{i,t}$, $C_{i,t}$, and T_t respectively denote shot rate and copayment in i area and t year, and year variable for '02/'03 season that captures the difference between sample seasons keeping constant all other aspects. The estimation equation is

$$R_{i,t} = \alpha_i + \alpha_c C_{i,t} + \alpha_t T_t + \varepsilon_{i,t} \tag{7}$$

The second part is to estimate the relationship between shot rate and mortality rate due to pneumonia or influenza. The estimation equation is

$$D_{i,t} = \beta_i + \beta_R R_{i,t} + \beta_t T_t + \upsilon_{i,t} \tag{8}$$

where $D_{i,t}$ is pneumonia and influenza mortality rate. Unfortunately, since pneumonia and influenza mortality rate of the elderly by area and season is not reported, we use the mortality rate of the total population irrelevant to the age.

The estimation method is weighted least square with the elderly population and the total population as a weight respectively in the first and the second estimation.

Note that we have to remark if $\varepsilon_{i,t}$ and $\upsilon_{i,t}$ are correlated, the estimated coefficient β_R certainly has bias. Moreover, the direction of bias may be positive or negative depending on $E[\varepsilon_{i,t},\upsilon_{i,t}]$. For example, an increase in the number of weak elderly persons residing in institutions, shot rate for these people usually is higher than elderly people living at home, and the mortality rate may still be higher due to their weakness even if the shot rate is the same. This correlation may lead to the upper bias in the coefficient. Conversely, the shot rate may represent the overall welfare spending or situation of the elderly in that area controlled out copayment. If this spending or situation improve the elderly health condition and reduce the mortality rate, this relationship make the lower bias in the coefficient.

In both cases, these are very well known as the simultaneous bias and we have to adopt a method that corrects such biases. The method, called instrumental variable method, uses the fitted variables of $R_{i,t}$ in the first estimation as an explanatory variable in the second estimation

rather than the observed raw $R_{i,t}$[1].

(2) Benefit–Cost Ratio

Using these estimation results, we can evaluate the policy by net benefit (NB) and benefit–cost ratio(BCR). NB is defined simply by the difference of benefit and cost due to the policy, and BCR is defined by its ratio.

NB can be calculated as follows. The perspective is of the society, and the time horizon is set to be one year because the effect of the vaccination is less than one year and the vaccination can extend their life one year at maximum. Due to the limited data, the effectiveness of the vaccination is limited to the prevention of death. The vaccination cost is defined as the sum of the copayment and subsidy, but the opportunity cost for the shot is not taken into consideration because they are typically retired. Moreover, the side effects of vaccination are also ignored for simplicity.

Vaccination cost is assumed to be 4.5 thousand yen (36 dollars), and the benefit of one year increase in life expectancy is assumed to be 6 million yen (48 thousand dollars).These numbers are widely used number in the US[2] and it is confirmed to be plausible even in Japan[3].

Then NB is

$$\begin{aligned}&\text{Monetary value of avoidance in mortality by rising shot rate} - \text{Additional cost by rising shot rate}\\ &= \text{Rising shot rate due to subsidy} \times \text{Reduction in mortality rate due to rising shot rate}\\ &\quad \times 6 \text{ million yen} - \text{Rising shot rate due to subsidy} \times 4500 \text{ yen}\\ &= 4000/3 \times \text{Reduction in mortality rate due to rising shot rate}\end{aligned} \tag{9}$$

Similarly, BCR is

$$\begin{aligned}&\frac{\textit{Monetary value of avoidance in mortality by rising shot rate}}{\textit{Additional cost by rising shot rate}}\\ &= \frac{\textit{rising shot rate due to subsidy} \times \textit{Reduction in mortality rate due to rising shot rate}}{\textit{Rising shot rate due to subsidy}}\\ &\quad \times \frac{6 \textit{ million yen}}{4500 \textit{ yen}}\\ &= \frac{4000 \times \textit{Reduction in mortality rate due to rising shot rate}}{3}\end{aligned} \tag{10}$$

V.1.3. Result

(1) Estimation Result

Summary statistics are shown in Table 3. Estimation results are summarized in Table 4.

The first and second columns in Table 4 show that the increase in copayments significantly reduces the shot rate. As its coefficient is −.007, since it means the shot rate would rise by .007 percentage point in every 1 yen subsidy, if copayment is subsidized by 1000 yen (8), then the shot rate will rise by 7 percent points. Since the coefficient for '02/'03 season is significantly positive, the shot rate raised by 8.8 percent points in '02/'03 season compared with the '01/'02 season where the other situations are completely the same. All area dummies, which indicate the difference from Sapporo, are insignificant. Since the degree of freedom adjusted R^2 is high, it fits quite well.

The third and fourth columns in Table 4 summarize the estimation results of crude weighted least square about mortality rate and they indicate that the shot rate negatively effects mortality rate but it is not significant. On the other hand, the fifth and sixth columns in Table 4 show the results for the instrument variable method. They show significant effect of shot rate on mortality rate and its estimated coefficient is −0.003, i.e. if the shot rate is raised by 10 percentage points, the mortality rate of pneumonia and influenza would decrease by .03 percentage point.

(2) Net Benefit and Benefit Cost Ratio

Suppose the calculation of the net benefit and BCR of the policy change which raise 1000 yen (8) in subsidy. At first, this policy would increase the shot rate by 7 percentage points as mentioned above and this reduces the mortality rate of the would population by .0196(=7×0.0028) percentage points. It means to avoid 23520 (=.000196×120 million) death. This benefit can be evaluated as 141.2 billion yen (1.13 billion dollars) (=23520×6 million yen) if the value of life is assumed to be 6 million yen or 50 thousand dollars.

On the other hand, the additional cost of this policy change must be the product of the 7 percent points rise in the shot rate, 4500 yen (cost of vaccination in social per one elderly) and 20 million (population of the elderly). It expends 6.3 billion yen or 50.4 million dollars. Therefore the net benefit must be the difference of benefit and cost and it is 134.9 million yen or 1.08 billion dollars.

Table 3:Summary Statistics

	Average	Std. Dev.	Minimum	Maximum
Shot Rate(%)	29.6695	6.067872	18.4074	45
Copayment (yen)	1171.429	427.618	1000	2200
Mortality Rate(%)	0.0409995	0.0315513	0.0033683	0.1753567

Table 4 : Estimation Result

Explanatory Variable	Estimator	*p*−value	Estimator	*p*−value	Estimator	*p*−value
Copayment	−0.0066561	0.002				
Shot Rate(Instrument)			−0.0006669	0.304	−0.0027877	0.034
'02/'03 Season	8.757308	0.000	0.0112177	0.088	0.0295542	0.015
Sendai	1.208579	0.727	−0.0047122	0.500	−0.0021133	0.780
Chiba	5.458579	0.153	0.0057438	0.475	0.0173561	0.141
Tokyo	−1.674325	0.300	0.0023957	0.727	−0.0144918	0.090
Yokohama	−0.6914208	0.781	−0.0021031	0.682	−0.0035337	0.464
Kawasaki	−5.184099	0.123	0.0056449	0.442	−0.0053089	0.455
Nagoya	−3.341421	0.222	0.0034983	0.561	−0.0035525	0.502
Kyoto	−4.723365	0.113	0.0073267	0.297	−0.0026762	0.680
Osaka	−4.441422	0.095	0.012161	0.065	0.0027774	0.618
Kobe	−4.691421	0.117	0.0034541	0.614	−0.0064597	0.325
Hiroshima	4.058578	0.225	0.0061583	0.394	0.0148015	0.141
Kitakyuushuu	−5.79142	0.076	0.0122467	0.137	0.0263081	0.548
Fukuoka	−2.991421	0.350	0.0009681	0.883	−0.0053403	0.376
Constant	34.46885	0.000	0.0437944	0.030	0.1028622	0.005
Sample Size	28		26		26	
F statistics	10.81		2.53		2.60	
p−value for *F* statistics	0.0001		0.0639		0.0537	
$\overline{R}^2$	0.8357		0.4622		0.4548	

Note: Coefficients for '02/'03 Season indicate the structural difference of it from '01/'02 Season keeping constant all other aspects. Positive coefficient means that the average is larger in '02/'03 season than in '01/'02 season if the situation which is represented by figure of explanatory variables are the same in both season.

Following the similar way, we can calculate its BCR easily, i.e.

$$\frac{.0028(-1000)(-.007)\ 6\ million\ yen\ /\ (2000/12000)}{-1000(-.007)4500yen} = 22.4 \qquad (11)$$

where 2000/12000 in the numerator is the adjustment factor for the elderly because the potential population of the numerator is the whole population but the counterpart in the denominator is the population of the elderly. Moreover its 95% confidence interval is calculated as [2.2,43.7] and we can confirm that this BCR is significantly greater than 1.

V.1.4. Discussion

(1) Evaluation of Estimation Results

From Table 3, showing the average shot rate and average copayment, we can see that the price elasticity of shot rate is −0.2606. It appears to be higher than the results of the previous study. That is, the study based on the conjoint analysis which is the most reliable technique with a hypothetical questionnaire indicates −0.02 to −0.04 for the elasticity, and the actual behavior in the '01/'02season indicates −0.1 for the elasticity[4]. Hence, the result in this paper shows that the shot rate is very elastic against price.

There are two main reasons for these differences. Firstly, this study focus on the metropolitan and big cities and so it may be biased toward extremely urban areas, whereas the previous studies cover the whole of Japan. If the residents in the urban areas have higher price elasticity to vaccination than rural areas, our results here may be reasonable. In this sense, the previous studies seem to be more general than this research.

Conversely, the data in this paper covers all residents in an area, while the previous study relied on survey by mail and it did not cover all the residents, of course, and they may not be representative. If the respondents of the questionnaire tend to have inclination toward vaccination for influenza compared with non−respondents, the shot rate may be insensitive to price. In this sense, the result in this study seems to be more reliable than the previous one. Though, it is not certain which estimate and reasoning is more reasonable, We have to remind ourselves that our final goal, namely the analysis with the BCR, is independent of price elasticity of shots as explained before.

On the other hand, the shot rate elasticity of mortality rate is −2.48, and thus mortality is elastic against shot rate. Combining these two estimation results, if copayment would be cut by a thousand yen (eight dollars), it raises the shot rate by 7 percentage points, and reduces the mortality rate due to pneumonia and influenza by .029 percentage point. It seems like a very small number, but since the average mortality rate due to pneumonia and influenza is very small, the effect certainly is quite high. In fact, this means this policy can reduce about 423 death in an average big city.

Since *F* statistics in the first equation is higher than ten, the fitted variables seems to be good instrument[5]. In other words, the reason of insignificance of the shot rate in the crude weighted least square can be inferred as positive simultaneous bias which offsets the shot rate effect on the mortality rate. Therefore, the instrument variable can solve this bias and it is a more appropriate method for this problem.

(2) Evaluation for BCR

The obtained BCR, 22.4, is quite high compared with the other countries or other vaccinations. In some other countries, since it is 1.93[6] for children before school and 1.81[7] or 2.92[8] for healthy adults, the obtained IBCR is much higher. Compared with other diseases, it is 2.5[9] for measles in Japan and it is just 1.4[10] in the case of hepatics B for all infants in China where there are epidemic areas. Overall, the policy of subsidy for shots for the elderly is quite cost–effective and there is concrete evidence for this.

(3) Problem and limitation in this analysis

At first, there are some differences in the definition of population among areas for the policy targeting or/and for the shot rate calculation. Especially, this policy also subsidizes the non–elderly, i.e. between 60 and 64 years old, who have heart, kidney, and respiratory problem or HIV career. Moreover, each city sometimes extend the target population more than the national policy requirement. Typically, some cities subsidize the institutionalized elderly even if they are younger than 65. These additional target populations are included in the denominator in some cities, but are not in other cities. The subsidized number in the numerator of the shot rate include such additional targeting populations, and thus the shot rate may be different depending on whether the denominator include such additional targeting populations or not. However, these additional target populations are quite small compared with the elderly, and it is less than just one percent. Therefore, such an inconsistency in the denominator of the shot rate may not substantially affect the result.

Moreover, the starting date of subsidy is not the same among areas. In particular, it is remarkable in the first season of this policy, i.e. the '01/'02 season. Our data of shot rate only includes those who received the subsidy, and does not include those who did not receive subsidy but received shots. So the shot rate may be lower than the actual rate in the area where the starting date of subsidy was delayed. In this sense, the data of shot rate is always lower than the actual shot rate among the elderly. This measurement error may lead to upper bias of the estimated coefficient of the shot rate in the second estimation. Hence, it also lead upper bias in IBCR. However, it is not sure how many elderly people received a shot but were not subsidized, and so we cannot evaluate this effect in detail.

On the other hand, it is questionable whether our sample in the metropolitan and big cities represents the whole of Japan or not. The coverage of the elderly population in our data is 21% of Japan, but it may not be the average population. Especially, there may be big differences from those in the rural areas as mentioned before. So as to check the robustness of the obtained result, we should extend our analysis to the other areas.

Additionally, the effect of the influenza epidemic on the mortality rate is measured by excess

mortality which is defined by the difference between the actual number of deaths and the hypothetical number in the case of no influenza epidemic[11–14]. Therefore, we have to replace the mortality definition from the crude number of death to the excess mortality. In particular, excess death should be defined regardless of the cause of death[14] because it is very well known that the influenza epidemic raises the mortality rate from other causes than pneumonia or influenza and these death can be prevented by the vaccination and control of the influenza epidemic. Moreover, if we can limit the number of deaths to those of more than 65 years old, it would be a more precise measure. In this sense, the excess mortality of those older than 65 years old in all causes of death is the best measure to evaluate the vaccination effect.

At the same time, we also need more explanatory variables which affect the shot rate or mortality. For example, the hortative measure for vaccination may be much different among local governments and it may affect the shot rate. Even in this case, if such a measure did not change in an area in the two season, this effect can be controlled out completely by the area dummies and it does not affect the estimated coefficient.

On the other hand, there are many implicit assumptions in BCR. First of all, since we limit the effect of vaccination to the prevention of death, and thus it is certainly a finer measurement than the prevention of the severe conditions like hospitalization as emphasized. Since it is difficult to obtain the data of the number of patients and the hospitalized, these numbers would be based on the similar estimation. Hence, these are far less precise than the number of death. In other words, we choose preciseness rather than broadness in the definition of effectiveness. Obviously, this limitation lower BCR. If we take the effects of vaccination on the number of patients and the hospitalized into consideration, BCR definitely becomes higher than the BCR discussed in this paper. It strengthens our conclusion in favor of the subsidy and has never change it.

Conversely, the ignorance of opportunity cost for vaccination or side effects certainly raises BCR. However, almost all of them are retired, and suffered from chronic disease and thus they usually visit a doctor, their additional opportunity cost for vaccination seems to be small. Concerning side effects, on 28 August, 2003, the Ministry of health and welfare reported only two fatal cases and 18 severe side effects from 1998 to 2003. Therefore, we can safely ignore these cost and the obtained conclusion is probably not affected by the introduction of these costs.

Finally, we can extend the effectiveness of vaccination to the number of patients or the medical cost. The data limitations of these variables are already mentioned. Moreover, since the primary purpose of vaccination is the prevention of severe cases, if we extend to these aspects, the results may not be clear and BCR may decline. In extreme case, the fatal case may use less medical resources compared with severe cases where the patients survives. In this sense, the limitation of effectiveness on the number of deaths seems to be more appropriate for considering the vaccination policy. Nevertheless, the research on the number of patients and medical cost are unambiguously important and we need to overcome the data limitations.

V.2. Example 2: Smoking Cessation Program

We show *ex ante* policy evaluation for smoking cessation program as another example of cost–effectiveness analysis in policy for medicine or public health which is originally in Ohkusa and Sugawara(2005a).

V.2.1. Objective

There are many programs to cease smoking, such as group therapy, individual therapy by professional staff such as medical doctors or nicotine replacement therapy. The Tobacco Use and Dependence Clinical Guideline Panel, Staff, and Consortium Representatives (2000) recommend the use of nicotine replacement therapy for nicotine dependence. In Japan, nicotine replacement therapy uses nicotine patch or nicotine gum, but the former requires a prescription written by a medical doctor. The later does not require it and we can buy it as an OTC (Over–the–counter) drug at any pharmacy without consultation by medical doctors. The Nicotine patch is used in more than 60 countries, and is an OTC in more than 30 countries. However, it has not switched to the OTC, yet in Japan. On the other hand, individual consultation by a medical doctor is not covered by health insurance in Japan.

This paper tries to conduct an ex ante cost effective analysis to evaluate new policies for smoking cessation, such as switching the nicotine patch to OTC (PO) and insurance coverage for individual therapy by medical doctors (PI).

V.2.2. Material and Method

The survey collected information through the web site in December 2004. The respondents were limited to smokers aged 20 to 59, and randomly drawn stratified are, age, and gender which replicate the national average from the list of the contracted members with the survey company.

It employs the hypothetical questionnaire which is used in the Conjoint analysis (Halpern, Berns and Israni(2004), Ratcliffe, Buxton, McGarry, Sheldon and Chancellor(2004), Maddala, Phillips and Johnson(2003) Schwappach(2003), Phillips, Maddala and Johnson(2002), Gyrd–Hansen and Slothuus(2002), Aristides, Chen, Schulz, Williamson, Clarke and Grant(2002), Bryan, Roberts, Heginbotham and McCallum(2002), Ratcliffe, Van Haselen, Buxton, Hardy, Colehan and Partridge(2002), Telser and Zweifel(2002), Gabriel, leung, Chan, Chau and Chua(2001), Johnson, Banzhaf and Desvousges(2000), Ratcliffe(2000), Tilley and Chambers(2000)). It asks the respondent to choose visiting a doctor or going to a pharmacy under the hypothetical situations: about cost of medical services and OTC, traveling time to visit a doctor, insurance coverage of individual consultation by a medical doctor, and explanation by a pharmacist on now to use the OTC drug.

In each attribute, the levels are set as follows; traveling time to visit a doctor: 30 minutes, 60 minutes, and 120 minutes, cost for both medical service and OTC: 100 thousand yen (800 dollars) to 500 thousand yen (4000 dollars) by 100 thousand yen (800 dollars), explanation by pharmacist: none, 5 minutes and 10 minutes, insurance coverage and switching to OTC: yes or no.

In the case of insurance coverage for medical service, costs for medical services are reduced to be 30%, which is the coinsurance rate in Japan. The Cost for medical services and OTC are selected from estimations of cost in the current situation. Expected rate of those who quit smoking is supposed to be the same among programs.

Hence there are 900 possible scenarios. Of these, we select 50 scenarios orthogonally. Then we allocate 10 questions to each respondent and set 5 patterns.

We adopt random effects Probit model which is very common to estimate the Conjoint analysis. Especially, we estimate it separately, whether nicotine patch is switched or not, so as to fully evaluate its effect. In each estimation, the dependent variable is binary; if i th the individual choose OTC for j (j=1,2,…,10) th question, then $O_{i,t}$=1, and is zero otherwise. The estimation equation is

$$O^*_{i,j} = \alpha_0 + \alpha_M \log M_{i,j} + \alpha_T \log T_{i,j} + \alpha_I I_{i,j} + \alpha_C \log C_{i,j} + \alpha^5_E E^5_{i,j} + \alpha^{10}_E E^{10}_{i,j} + \mu_i + \varepsilon_{i,j}$$
$$O_{i,j} = \begin{cases} 1 & if \quad O^*_{i,j} > 0 \\ 0 & otherwise \end{cases} \tag{12}$$

where $M_{i,j}$, $T_{i,j}$, $I_{i,j}$, $C_{i,j}$, $E^5_{i,j}$, $E^{10}_{i,j}$ are respectively the cost for medical service, the traveling time to visit a doctor, dummy for insurance coverage, the cost for OTC, dummy for a five minute explanation by a pharmacist, and dummy for a ten minute explanation by pharmacist. μ_i is the random effect that captures individual effects and $\varepsilon_{i,j}$ is a stochastic disturbance term.

Next, we perform a cost effective analysis of these two new policies, PO and PI, based on the estimated demand curve. We calculate the incremental benefit cost ratio (IBCR) with and without such an externality among the current situation and switching the nicotine patch to OTC, insurance coverage for smoking cessation therapy by a doctor, and both of them. Moreover, we refer to net benefit in this policy so as to evaluate its amount of gain or loss in monetary term.

V.2.3. Results

We collected information from 2,839 individuals and the response rate was 51.9%. The estimation results are summarized in Table 5. It shows that all coefficients are significant and the variance of the random effect is significantly more than zero. Therefore, its consideration is important.

Table 5 :Demand for OTC and Medical Services which Assists Quit Smoking

	Nicotine Patch		Nicotine Gum	
	Marginal Effect	*p* value	Marginal Effect	*p* value
Medical Cost(log)	0.09595611	.000	0.23414278	.000
Traveling Time(log)	0.08233906	.002	0.09190193	.000
Insurance Coverage	−0.06234507	.021	−0.03687187	.001
OTC Cost(log)	−0.33815518	.000	−0.23327342	.000
Explanation(5 min.)[a)]	0.09684908	.000	0.01804711	.089
Explanation(10 min.)[a)]	0.07527379	.050	0.08756433	.000
# of sample	4725		7066	
# of individuals	2375		2377	
p−value for χ^2 test[b)]	<0.0000		<0.0000	
Log Likelihood	−1531.5		−3187.5	
p−value for χ^2 test[c)]	<0.0000		<0.0000	

Note:Dependent variable is binary variable whether they demand for OTC (nicotine patch or nicotine gum) or not.[a)]:"Explanation (5 min.)" and "Explanation (10 min.)" means that how long pharmacist explain about nicotine gum or patch when the consumer buy it at pharmacy.[b)]: Likelihood ratio test for estimation model against constant term only.[c)]: Likelihood ratio test for estimation model against the model without random effects.

The disease burden of smoking has been estimated as 3.7 to 7.3 trillion yen (Institute of Health Economics and Planning(1997), Ohkusa and Sugawara(2005b)). In this amount, externality is the insurance paid for the medical cost of treatment for smoking related disease. While it is not recognized as costs for smokers, it is actually the cost from the societal view point. Such externality is estimated as 0.88 to 1.12 trillion yen and rate of externality is estimated as 7.3/(7.3−1.12)−1=0.18.

The result of cost effective analysis is summarized in Table 6. It shows the results separately in terms of the externality considered. The first row indicates the case of switching nicotine patch to OTC. The second row presents the result of the case where the therapy by doctors does not become popular while it is covered by insurance, and where the traveling time is ninety minutes .The third and fourth rows summarize the result if the traveling time is reduced to be 60 or 30 minutes. Besides, the result if both PO and PI are implemented by traveling time is also shown. Table 7 shows the net benefit in OTC market and market of medical services.

Table 6:IBCR of Switching to OTC of Nicotine Patch and/or Insurance Coverage for Quit Smoking Therapy

Switching to OTC of Nicotine Patch	Insurance Coverage	Traveling Time (min.)	without Externality			with Externality		
				95% CI			95% CI	
			Median	Lower	Upper	Median	Lower	Upper
Yes	No	90	1.46	1.39	1.53	1.72	1.65	1.81
No	Yes	90	0.189	0.039	0.295	0.203	0.024	0.329
No	Yes	60	0.311	0.208	0.386	0.352	0.229	0.442
No	Yes	40	0.461	0.398	0.509	0.534	0.460	0.591
Yes	Yes	90	0.789	0.733	0.863	0.910	0.840	1.00
Yes	Yes	60	0.711	0.698	0.749	0.819	0.803	0.867
Yes	Yes	30	0.665	0.576	0.734	0.767	0.661	0.850

Table 7:Net Benefit of Switching to OTC of Nicotine Patch and/or Insurance Coverage for Quit Smoking Therapy by Doctor

Insurance Coverage	Traveling Time	After Switching		Before Switching	
		OTC	Medical Service	OTC	Medical Service
without Externality					
No	90	1321.0425 [929.8, 1805]	32.014884 [31.72, 32.27]	193.39717 [145.8, 252.9]	21.691252 [21.46, 21.92]
Yes	90	179.66613 [103.9, 207.5]	−211.5101 [−223.9, −198.7]	1.8866509 [1.193, .9411]	−712.20123 [−718.5, −705.7]
Yes	60	111.0619 [61.72, 191.2]	−214.41094 [−230.9, −197.6]	0.94754236 [.5874, 1.507]	−839.76731 [−848.4, −830.8]
Yes	30	44.870128 [23.23, 82.75]	−260.64714 [−288.6, −232.6]	0.27154143 [.1625, .4476]	−1013.9675 [−1028, −999.5]
with Externality					
No	90	2073.351 [1511, 2744]	47.488069 [47.23, 47.69]	303.77122 [231.9, 392.5]	36.81772 [36.55, 37.07]
Yes	90	340.16637 [206.1, 538.4]	−114.27436 [−128.1, −100.0]	3.4450426 [2.217, 5.282]	−571.11183 [−578.3, −563.7]
Yes	60	218.27756 [127.2, 358.7]	−106.47282 [−125.9, −86.70]	1.7637574 [1.112, 2.7599]	−654.23084 [−664.1, −644.0]
Yes	30	94.090628 [51.14, 165.4]	−9.1588703 [−39.45, 20.92]	0.52249262 [.3182, .8465]	−731.27581 [−747.5, −714.6]

These tables show obviously that the IBCR for PO exceeds one significantly and is 1.46 without externality, and is 1.72 with externality. The net benefit achieves 135 billion yen and it is higher than the current net benefit of 21 billion yen by more than 100 billion yen.

Conversely, the IBCR for PI is less than one, and thus it does not support the implementation of this reform. This policy reduces the net benefit in the OTC market to less than 0.2 billion yen due to the reduction in copayment by insurance coverage and the society loses 71.2 billion yen.

Moreover, if the number of medical institutions which provide smoking cessation therapy by doctors increases and traveling time is shortened to be 60 or 30 minutes, the net benefit in the OTC market is reduced to 30 million yen and the net loss in the market for medical service achieve more than 100 billion yen. Even if we take such an externality into consideration, it leads the negative net benefit amounts to 57 billion yen in the society as a whole. When we test PO and PI simultaneously, its IBCR is not larger than one significantly, even though the upper limit of a case reaches one.

V.2.4. Discussion

If we can assume that the expected benefit of quitting smoking is the same among smokers, the demand curve represents the subjective quit rate. In other words, the smokers who join the smoking cessation program are thought to have a higher subjective quit rate or have more aptitude of these programs than other smokers, those who do not join the program under the same price.

Conversely, almost all the cost effective analyses so far are usually based on the assumption of average individual or some artificial scenarios. Namely, the quit rate is assumed to be a certain level in all smokers in this case. It is true if we consider some nonexclusive public goods because the word "join" or "not join" does not make any sense. However it must not be true if we consider the private goods. In this sense, immunization, medical services or OTC drugs are exclusive private goods.

Even though there is heterogeneity among individuals, these studies ignore this clear fact and assume that they are homogenous. In our context, these typical cost effective analyses ignore the difference in the subjective quit rates which are represented by the demand curve for the smoking cessation program.

Especially, when we evaluate the policy with some subsidies including health insurance coverage, it is very important to recognized that such a policy enforces joining the programs whose subjected benefit is lower than those who join the program even though such a policy is not implemented. In other words, the average benefit among participants must be decreased by such a policy. It is very well known as a deadweight loss. Needless to say, such a policy for private goods must not be recommended because it worsens the welfare. Therefore the cost

effective analysis must be based on the demand curve if we consider the policy for private goods, or the analysis leads to the wrong conclusion.

Moreover, almost all the cost effective analyses so far usually fail to consider externalities. As explained above, the cost effective analyses so far have been based on the average or hypothetical individual and summing them up into the aggregate benefit and cost. Hence the externality which does not count at the individual level may be ignored. This externality does not seem essential in the cost effective analyses and thus it can be incorporated in the analysis, but typically it is ignored.

On the other hand, we can easily take such an externality into consideration which is represented by the deviation of social benefit from the demand curve. This deviation is the only reasons for policy intervention. In the case of positive externality, since the deviation of social benefit other than individual's utility, the deadweight loss induced by the policy may be compensated. Hence, cost effectiveness analysis for private goods is not for considering whether such a positive externality is sufficiently large enough to compensate the deadweight loss. Therefore, externality is the most important and essential of cost effective analysis for private goods.

In our context, PI means to provide subsidy of 70% of the medical cost and thus it leads to some deadweight loss. Therefore, it needs some evidence that its externality is larger than the deadweight loss. On the other hand, switching the nicotine patch to OTC is a kind of deregulation and thus it does not lead to deadweight loss.

Table 6 and 7 imply that PI does not have evidence but PO has it to support implementation. The point to evaluate PI is whether the externality is so large that this deadweight loss can be compensated. Unfortunately, it cannot. Even if we perform PO and PI simultaneously, though the nicotine patch as OTC partially offset the demand for medical services, this strategy also does not have evidence to support implementation. Therefore we can conclude that PO is strongly recommended but PI is not.

If we calculate IBCR without using a demand curve, but assuming average individual even if a new policy is implemented, it should be as follows: Assume the quit rate is 30% in any program, PO and PI, and cost in both policies is 30 thousand yen (240 dollars), then the additional cost to quit per person is 70 thousand yen (560 dollars). On the other hand, the benefit of quitting smoking is supposed to be 3.7 to 7.3 as mentioned before, and its per capita term is 185–365 thousand yen (1.48–2.92 thousand dollars) if the smoking population is 20 million. We note that this number does not depend on how many person attend due to the new policy. Therefore IBCR is 2.6429(=185/70)– 5.2143(365/70).

Even though this extremely simple calculation ignores discounting, it does not seem to affect the its implication. It is clearly more than one, so this very simple calculation recommends the implementation of both PI and PO. However, we prove that PI make huge deadweight loss and externality cannot compensate it fully, so we cannot recommend it. This small example explain

how such a simple cost effective analysis leads to a wrong conclusion.

However, we have some limitations. The most important of all would be environmental tobacco smoking (ETS). It must raise positive externality of quitting smoking. Even though its medical cost is very small compared with the medical costs of smokers, the bad smell of smoke worsen the QOL of non-smokers. Insurance coverage for smoking cessation therapy by doctors may be changed to be cost effective. Accumulation of knowledge about ETS in epidemiology and health economics is necessary as soon as possible so as to evaluate smoking cessation programs.

VI. Conclusion

As mentioned before,

In other countries, the smoking policy is based on the result of cost effectiveness analysis in the field of policy evaluation for medicine or public health. The more responsibility will be required for policy makers to explain for all citizens about the process of political decision they made. Taking this situation into consideration, additional research and discussion should be needed to respond this requirement.

Acknowledgement

Section 3 was financed by the Japanese Ministry of Education and Science (2000 Scientific Research Grant, entitled "Evaluation for Economic Institutions Based on Empirical Research on Household Behavior" (grant number:12124207), headed by Fumio Hayashi of the University of Tokyo. The author also appreciates the help from Prof. Shunya Ikeda of the Medical School in Keio University and Prof. Shunsuke Ono of the Department of Pharmaceutical in Kanazawa University for their useful suggestions.

Section 5.1 is a part of the outcome of Health and Labour Sciences Research Grants for Research on Emerging and Reemerging Infectious Disease 2003. *Policy Evaluation for Influenza Vaccination based on EBM.* (headed by Prof. Yoshio Hirota, Osaka City University). We appreciate deeply the cooperation for this research by the correspondents of the section of infectious disease in metropolitan and 12 city governments. Moreover,Dr.Kiyosu Taniguchi, Dr. Yuki Tada, Dr. Mika Shigematsu and Dr. Kazuki Masuda of National Institute of Infectious Disease, and Dr.Masayo Sato of National Institute of Social Security and Population.

Section 5.2 is a part of the outcome of Health and Labour Sciences Research Grants for Research on Health Service 2004. *Research for disease burden of smoking and effective quit smoking policy.* (headed by Dr. Tomoike) We appreciate deeply the cooperation for Prof .Ohkubo of Tsukuba University for his assistance.

References

Brooks R. (1996). EuroQol: the current state of play. *Health Polic.* 37(1):53–72.

Ware J.E. and D.C.Sherbourne. (1992). The MOS 36–item short–form health survey. *Medical Care.* 30:473–83.

Le G.C., Buron C., Costet N., Rosman S. and Slama P.R.,(2002).Development of a preference–weighted health status classification system in France: the Health Utilities Index 3. *Health Care Management Science.* 5(1):41–51.

Weeks J., J.O'Leary, D.Fairlough, D.Paliel and M.C. Weinstein. (1994). The Q–tility Index: A new tool for assessing health–related quality of life and utilities in clinical trials and clinical practice. *Proceedings of the American Society of Clinical Oncology*. 13:436.

Van Doorslaer E., A. Wagstaff, et al., 1997, Income–related inequalities in health: Some international comparisons. *Journal of Health Economics.* 16: 93–112.

Marthe R.G., E.S.Joanna, B.R. Louise, and C.W.Milton.(1996) Time preference cost–effectiveness in health and medicine. Oxford University: Oxford.

George,B., A. Harris and A. Mitchell(2001). Cost–effectiveness analysis and the consistency of decision making. *Pharmacoeconomics*.19:1103–1109.

Cutler, D. and E. Richardson (1997). Measuring the Health of the U.S. Population. *Brookings Papers Microeconomics* :217–271.

Donaldson C. S. Brich and A. Gafini(2002). The distribution problem in economic evaluation: income and the valuation of costs and consequences of health care programs. *Health Economics.* 11: 55–70.

Goldman L, Gordon DJ, Rifkind BM, Hulley SB, Detsky AS, Goodman DW, Kinosian B, and Weinstein MC.(1992). Cost and health implications of cholesterol lowering. *Circulation.* 85(5):1960–1968.

Hanemann W.(1991). Willingness to pay and Willingness to pay and Willingness to accept: How much can they differ? *American Economic Review.* 81:635–47.

Horowitz J, McConnell K.(2002). A review of WTA/WTP studies. *Journal of Environmental Economics Management.* in press.

Laupacis A, Feeny D, Detsky AS, and Tugwell PX.(1992). How attractive does a new technology have to be to warrant adoption and utilization? Tentative guidelines for using clinical and economic evaluations. *Canadian Medical Association Journal.*146(4):473–481.

Salkeld G, Randy M, Short L.(2000). The veil of experience: do consumers prefer what they know best?. *Health Economic Letter* .4: 4–9.

Tolley,G., D. Kenkel and R. Fabian(1994). *Valuing Health for Policy*. The University of Chicago Press.

Walton, S.M., P.E.Graves, P.R.Mueser and Johnson, F.R., E.E.Fries and H.S.Banzhaf (2002). The Bias Against New Innovation in Health Care: Value Uncertainty and Willingness to Pay.

Value in Health. 5:67–70.

Welte R, Postma MJ, Bos JM, Van Alphen L(2002). Economic Evaluation of A Large-Scale Meningococcal C Vaccination Program in the Netherlands. present at ISPOR 5th Annual European Congress.

Lothgren M. and N. Zethraus(2004). Definition, Interpretation and Calculation of Cost-Effectiveness Acceptability Curves. *Health Economics* .9:623–630.

Willan A.R. and B.J.O'Brien. (1999). Sample Size and Power Issues in Estimating Incremental Cost-Effectiveness Ratios from Clinical Trials Data. *Health Economics*.:203–211.

Joseph C. Gardiner, Marianne Huebner, James Jetton and Cathy J.Bradley. (2000). Power and sample size assessments for tests of hypotheses on cost-effectiveness ratios. *Health Economics*.:227–234.

Fieller E.C. (1954). Some problems in interval estimation with discussion. *Journal of the Royal Statistical Society, Series B*. 16:175–185.

Willan A.R. and O'Brian, B.J. (1996). Confidence intervals for cost-effectiveness ratios: an application of Fieller's theorem. *Health Economics*. 5:297–305.

Wiiliam A.H. (1985). Economics of Coronary Artere Bypass Grafting. *British Medical Journal* .291, 326–329.

Stratton KR, J.S. Durch, and R.S.Lawrence(2000). *Vaccines for the 21st century*. National Academy Press.

Ohkusa Y. (2005).Policy evaluation for the subsidy for influenza vaccination in elderly. *VACCINE* .23:2256–2260.

Ohkusa Y and Sugawara T (2005a) .Cost effective analysis for switching OTC of nicotine patch and/or insurance coverage of quit smoking therapy at medical institution. *manuscript*

The Tobacco Use and Dependence Clinical Guideline Panel, Staff, and Consortium Representatives (2000). A Clinical Practice Guideline for Treating Tobacco Use and Dependence. *JAMA*. 283(24):3244–3254.

Halpern SD, Berns JS, Israni AK. (2004). Willingness of Patients to Switch from Conventional to Daily Hemodialysis: Looking before We Leap. *The American Journal of Medicine*. 116: 606–612.

Ratcliffe J, Buxton M, McGarry T, Sheldon R and Chancellor J.(2004). Patients' preferences for characteristics associated with treatments for osteoarthritis. *Rheumatology*. 43: 337–345.

Maddala TM, Phillips KA and FR. (2003). An experiment on simplifying conjoint analysis designs for measuring preferences. *Health Economics*. 2: 1035–1047.

Schwappach DLB. (2003). Does it matter who you are or what you again? An experimental study of preferences for resource allocation. *Health Economics*. 12: 255–267.

Phillips KA, Maddala T and Johnson FR. (2002). Measuring preferences for Health Care Interventions Using Conjoint Analysis: An Application to HIV Testing. *Health Services Research*. 37(6):1681–1705.

Gyrd-Hansen D and Slothuus U. (2002).The Citizen's Preferences for Financing Public Health

Care: A Danish Survey. *International Journal of Health Care Finance and Economics*. 2: 25–36.

Aristides M, Chen J, Schulz M, Williamson E, Clarke S, and Grant K.(2002). Conjoint Analysis of a new Chemotherapy. *Pharmacoeconomics*. 20(11): 775–784.

Bryan S, Roberts T, Heginbotham C and McCallum A. (2002).QALY–maximization and public preferences: results from a general population survey. *Health Economics*. 11:679–693.

Ratcliffe J, Van Haselen R, Buxton M, Hardy K, Colehan J, Partridge M.(2002). Assessing patients' preferences for characteristics associated with homeopathic and conventional treatment of asthma: a conjoint analysis study. *Thorox*. 57(6):503–508.

Telser H and Zweifel P. (2002). Measuring willingness–to–pay for risk reduction: an application of conjoint analysis. *Health Economics*. 11: 129–139.

Gabriel M. leung GM, Chan SSC, Chau PYK, Chua SC. (2001).Using Conjoint Analysis to Assess Patients' Preferences When Visiting Emergency Departments in Hong Kong. *Academic Emergency Medicine*. 8(9): 894–898.

Johnson FR, Banzhaf MR and Desvousges WH.(2000). Willingness to Pay for Improved Respiratory and Cardiovascular Health: A Multiple–format, Stated–Preference Approach. *Health Economics*. 9:295–317.

Ratcliffe J. (2000). Public Preferences for the Allocation of Donor Liver Grafts for Transplantation. *Health Economics*. 9:137–148.

Tilley S and Chambers M.(2000). An application of conjoint analysis to the process of psychiatric day hospital care. *Journal of Psychiatric and Mental health Nursing*. 7: 371–372.

Institute of Health Economics and Planning. (1997). *Report on Cost Benefit Analysis for Smoking Policy*(in Japanese).

Ohkusa Y. and Sugawara T.(2005b). Cost effective analysis for quit smoking program (in Japanese), Health and Labour Sciences Research Grants for Research on Health Service 2004. *Research for disease burden of smoking and effcetive quit smoking policy*. (headed by Dr. Tomoike).