A strategic approach for comparing different types of health risks

— A risk assessment of toluene exposure using quality-adjusted life years —

Atsuo Kishimoto

[Translation from *Synthesiology*, Vol.1, No.1, p.31-37 (2008)]

The issue of social demands must be taken into consideration when developing a methodology for risk assessment of chemical substances. We classified the objectives of risk assessment into the following three categories: (A) to derive reference values, (B) to establish screening assessment to remove chemical substances of low importance, and (C) to set priorities based on comparisons of health risks and cost-effectiveness of risk reduction measures for different chemical substances. In this categorization, we demonstrated that while the existing risk assessment methods fulfill objectives A and B, they do not satisfy objective C. Therefore, the steps of risk assessment methodology were revised to fulfill objective C, and the effectiveness of the new methodology was demonstrated using toluene as an example. By adopting the loss of quality-adjusted life years (QALYs) as an indicator of human health risks, we could compare the health risks for different chemical substances as well as compare chemical substances and other risks such as accidents and infectious diseases.

Keywords **:** Chemical substances, risk assessment, social demands, quantification, comparisons, cost-effectiveness

1 Introduction

Concept of risk and methodology for risk assessment are essential in rationally addressing environmental, safety, and health issues. Human health risks from chemical substances are determined by two factors, toxicity (hazard) and exposure (intake). Low exposure to a highly hazardous substance may result in low risk, although high exposure to a substance of very low hazard may become high risk. In an established chemical risk assessment methodology, comparisons are conducted between NOAEL (no observed adverse effect level) derived from animal tests or epidemiological studies and actual or potential exposure levels. Then, judgment is made whether the exposure level of target substance is acceptable or not. This thinking is actually applied to establish various environmental and safety standards, as well as in screening assessments for removing chemical substances of low importance.

However, this is not the only expected role of risk assessment. To achieve the ultimate goal of minimizing the total risk in our society, it is necessary to quantify and compare the health risks of each chemical substance, estimate the cost-effectiveness of each countermeasure to reduce risk, and rank them accordingly. Methodology of risk assessment suited to these purposes has not yet been developed.

In this paper, we propose an alternative framework of risk assessment to address emerging social demands, by returning to the scheme of how conventional risk assessment was established. We then present an application of this framework to chemical substance, toluene [1]. It was a process of trial and error involving returning to social demands, reconsidering each element in a consistent manner, and reintegrating each step into one consistent risk assessment process.

Toluene is a clear, colorless, and highly volatile liquid at room temperature. Most of the toluene emitted into the environment disperses into the atmosphere. Its volume released into the atmosphere is largest among the 354 chemical substances reported in the PRTR (Pollutant Release and Transfer Registers) system in Japan. Toluene is also known as one of the major indoor air pollutants, and a non-binding guidance value was set by the Ministry of Health and Welfare in 2001. In the second section, we review the relationship between social demands and risk assessment methods. In the third section, we present improvements in the steps of risk assessment and their application to toluene. In the final section, the implications are discussed.

2 Social demand and risk assessment methodology

The current procedures for chemical risk assessment are conducted according to the gray arrows in Figure 1 (from bottom to top). They are divided into two types depending on their objectives (social demands). First is to derive reference values, such as environmental standards or acceptable daily intakes (ADIs), which are compared with measured or estimated exposure levels. Examples are environmental standards and guidance values set by the Ministry of the Environment, and also the indoor guidance values, the tolerance values for pesticide residue, and the standard control concentrations for occupational settings set by the Ministry of Health, Labour and Welfare. The social demands they meet include protecting vulnerable people (such as

Research Center for Chemical Risk Management, AIST Tsukuba West,Tsukuba,Ibaraki 305-8569, Japan E-mail : kishimoto-atsuo@aist.go.jp

children) and most highly exposed groups in our society. The reference values are derived from NOAEL in animal tests divided by sufficiently large uncertainty factors. Therefore, even if the exposure level slightly exceeds the reference values, no adverse effects will appear in most people.

Second is to conduct screening-level risk assessments. Examples are: the Environmental Risk Assessment prepared by the Ministry of the Environment; the Initial Risk Assessment prepared by the Ministry of Economy, Trade and Industry; and the various risk assessments prepared by the Food Safety Commission. Worst-case scenarios are applied not only to toxicity assessment but also to exposure assessment, in which the 95 percentile of measured or estimated values[Term 1] are often adopted as the exposure levels. The social demands that they cover include delisting substances that pose no or very small risk to society even if their toxicity or exposure is overestimated. These are screening-level risk assessments. The substances that are not delisted should be subject to detailed risk assessment, and precautionary measures are sometimes taken at this point.

These methodologies of exposure assessment, toxicity assessment, and risk characterization were developed to cover the social demands for setting reference values or for screening purposes. In contrast to the risk assessment procedures (gray arrows in Figure 1, from bottom to top), each step was developed in the reverse direction (white arrows in Figure 1, from top to bottom). The bottom line is that the current elemental techniques have been optimized to address particular social demand, and are not necessarily optimal responses to other social demands.

Today, there are emerging social demands in the field of chemical risk management. We must compare the health risk of a chemical substance being used now with the risk of a proposed substitute. We must deal with the trade-offs between chronic disease risks and accident risks. We must also assess the cost-effectiveness of risk reduction measures to maximize risk reduction within budget constraints. Current toxicity assessment with conservative assumptions (that are used to overestimate risks in case of uncertainty and variability) and exposure assessment with worst-case scenarios cannot be used to address these new social demands. When a risk assessor does not recognize the close relationship between methodology and social demand, he or she may naively apply the current procedures (Figure 1) to the assessment for some other purposes. As a result, the risk is overestimated, the effectiveness of risk reduction measures is also overestimated, and the cost-effectiveness ratio turns out much better than the expected (average) one. Since the extent of conservativeness (overestimation) is different in each case, we cannot compare the risks for different chemical substances. To address the third social demand in addition to the first and the second, it is necessary to travel downstream of risk assessment procedures, i.e. the social demands, and reconsider proper methods for each step upstream from scratch, just as when the current risk assessment methodology was first developed.

3 Quantification strategy to compare different kinds of risks

3.1 Casting back from social demands

We considered the conditions necessary for each step of risk assessment to address the social demands, in reverse direction of risk assessment operations. The white arrows in Figure 2 show this review process.

It is necessary to present the magnitude of various risks with common metric to compare them with other types of risks and set priorities for various risk reduction measures. The necessary condition of the common metric for health risks is to combine mortality risks (loss of life years) and morbidity risks (loss of quality of life). We decided to adopt quality-adjusted life years (QALYs) proposed and applied in the area

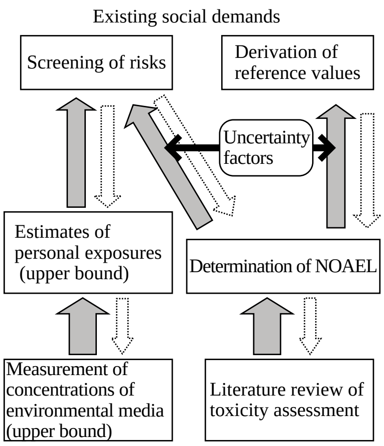

Fig. 1 Conventional process of risk assessment and social demands

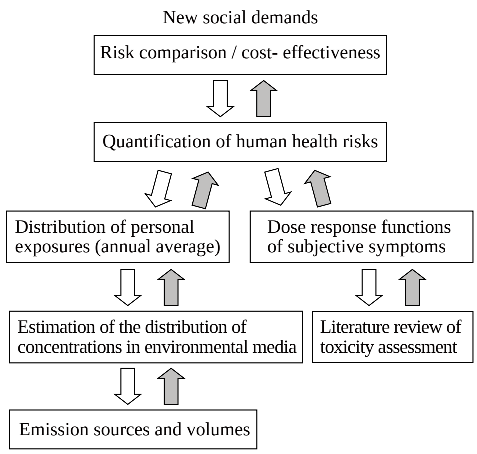

Fig. 2 Reexamination of elemental techniques for new social demands

of medical science. The volume of QALYs is represented by the gray area in Figure 3, where the vertical axis is quality of life (QOL) from 0 (death) to 1 (normal health) and the horizontal axis is age from birth to death. When the QOL is reduced to the dashed line for some reason, the loss of QALYs corresponds to the reduced area. Using QALYs as health risk metrics in risk assessment makes it possible to consistently capture not only loss of life expectancy, but also loss of QOL.

It was necessary to multiply the number of cases of various symptoms, including death, by the QOL weight from 0 (death) to 1 (normal health) to quantify human health risks resulting from exposures to chemical substances as loss of QALYs. The number of cases showing various symptoms was obtained by substituting the annual average exposure levels into the dose response functions[Term 2] that describe the relationship between the intake of substance and the probability of symptom presentation. The dose response functions should be preferably derived from human epidemiology studies with chronic endpoints of established diseases or subjective symptoms. The individual exposure level used in the calculations was expressed as annual average, which is described as the weighted average of indoor and outdoor concentrations. In order to estimate the national distribution of atmospheric concentrations by running the atmospheric dispersion model, the distribution data for emission volumes was necessary. These values should be preferably estimated as a central tendency or average estimates (and distribution profile if possible), instead of mere upper bound estimates.

On that basis, we reexamined the current methodology for each step of risk assessment. As a result, we found that almost all of the existing methodologies were not directly applicable. Therefore, following the gray arrows in Figure 2, we performed risk assessment for toluene step by step by revising slightly, making rough assumptions, or developing new methods.

3.2 Estimation of emission volumes

Initial risk assessments placed less significance on estimation of emission volumes or search for emission sources since they were based on measured concentrations in the environment or human bodies. However, to conduct simulation of the effect of proposed emission control measures, it is necessary to know the emission sources and volumes. For example, suppose that the share of some emission category was assumed to dominate but the actual share was less than half. This means that the effect of the emission control measure was less than half of the assumed effect. Even if the emission sources are discovered, items with large uncertainties, such as unintended emissions and natural sources, are often ignored or underestimated, and this may lead to underestimation of the total emission volume. In case of toluene, there are no estimates of evaporation from automobile fuel tanks and extra emissions (cold-start emissions) that occur before engines are warmed up at the beginning of PRTR.

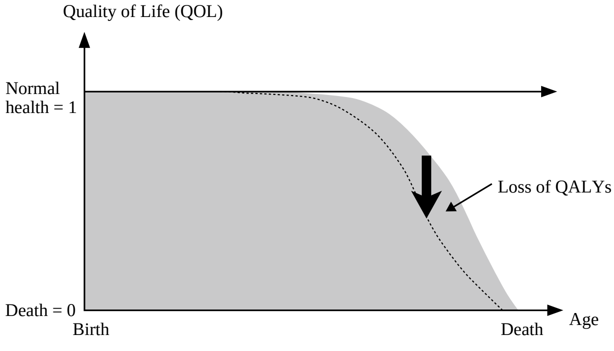

Fig. 3 Concept of quality adjusted life-years (QALYs)

3.3 Estimation of personal exposure distribution

Concentrations of chemical substances tend to be measured at locations and times that the concentrations may become high. The outdoor measurement data are biased toward values measured near the emission sources, and the indoor measurement data are biased toward values measured in newly-built houses. Therefore, we did not have information on the total exposure for Japanese residents. In addition, risks from outdoor and indoor concentrations are usually evaluated separately. Since most of the indoor data are daily averages for houses, information on annual average and inter-house and within-house variabilities are not available.

For outdoor concentrations, AIST-ADMER (Atmospheric Dispersion Model for Exposure and Risk Assessment) has been used to predict the regional distribution of concentrations with a horizontal resolution of 5 km for all Japan [2]. To run this model, estimated emission rates were allocated to 5 x 5 km square grids and meteorological data were collected from National Meteorological Observatories and from the Automated Meteorological Data Acquisition System (AMeDAS). The annual average concentrations within these 5 x 5 km square grids were between 0 and 67 $\mu g/m^3$ and the arithmetic mean was 1.5 $\mu g/m^3$.

For indoor concentration, it was necessary to obtain the within-house variability in the annual average “indoor concentrations of indoor origin”. The daily average data for 207 houses were obtained through the Freedom of Information Law from the Ministry of Health and Welfare [3]. The concentrations of toluene tended to be higher in indoor environments than in outdoor environments due to multiplicity of indoor emission sources. Since indoor toluene consists of ambient toluene that infiltrates indoors, as well as toluene emitted from indoor sources, “indoor concentrations of indoor origin” were defined as indoor toluene concentrations minus outdoor toluene concentrations. We assumed that these data followed lognormal distribution, which was called the (A) 24-hour average inter-house variability. To estimate (C) annual average inter-house variability, we designed the following procedures.

First, we estimated the variability of the daily average “indoor concentrations of indoor origin” in a house, i.e. (B) 24-hour average within-house variability. Next, we assumed that (A) equaled (B) plus (C), i.e. (C) would be obtained by subtracting (B) from (A). It was assumed that “indoor toluene concentrations of indoor origin” were proportional to the emission rate from indoor sources and inversely proportional to the residential air exchange rate (AER). However, we could not find the annual data for either emission rates or AER. After expert interviews, we judged that the daily variability throughout the year for toluene emission rates from indoor sources was virtually zero under all possible temperature ranges, and that the daily variability throughout the year for 24-hour average AER ranged from 0.5 times per hour to 10 times per hour (range in which 95 % of the values fell). Assuming these variabilities were all independent of each other, the annual average “indoor toluene concentrations of indoor origin” could be described as lognormal distribution whose geometric mean was 15.72 $\mu g/m^3$ and geometric standard deviation was 4.28. The distribution of personal exposure was calculated as a weighted average of indoor and outdoor concentrations, applying the result of a time-budget survey that showed that Japanese residents spent approximately 90 % of their day indoors and 10 % outside [4]. Figure 4 shows that the distribution moves from left to right with addition of each emission source category.

3.4 Derivation of dose response functions from epidemiological study

The goal of toxicity assessment in conventional risk assessment is to derive NOAEL from animal tests or epidemiological studies and set reference values. However, since the type and severity of toxicity endpoint and the extent of conservativeness (extent of overestimation) are different for different chemical substances, it is difficult to compare the risks from the NOAEL or reference values. We attempted to derive the relationship between health effects and exposure levels, i.e. dose response functions, from the results of an epidemiological study [5]. Among various subjective symptoms investigated in the epidemiological study in which subjects were workers at gravure printing factories, eight symptoms that were found to have significant associations between exposure levels and incidence rates were selected, and their dose response functions were derived as shown in Figure 5. The vertical axis indicates the incidence rates, and the horizontal axis shows the exposure levels. As the exposure increased, probabilities of incidence for each symptom as well as the expected number of symptoms increased.

3.5 Quantification of human health risks

The distribution of personal exposure concentrations (Section 3.3) was substituted into the dose response functions (Section 3.4), and the incidence numbers obtained were then multiplied by their severity. The health risks for residents in Japan from exposure to toluene were finally quantified as total loss of QALYs. The severity of each symptom and some combinations of symptoms were expressed in terms of QOL index, where 0 meant death and 1 meant normal health. QOL indices were calculated based on the Health Utilities Index

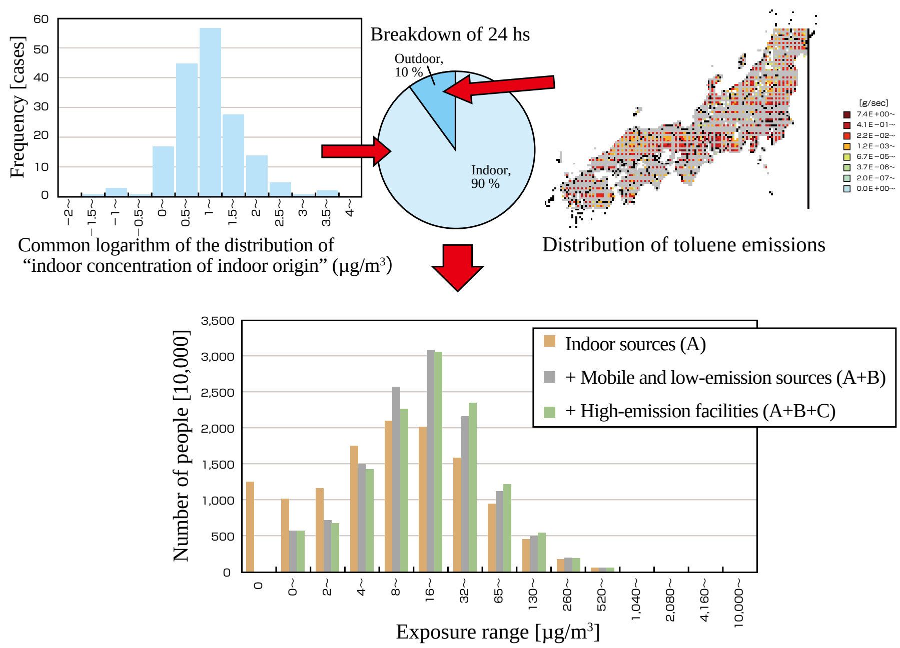

Fig. 4 Distribution of personal exposures (annual average) of Japanese residents

* “High-emission facilities” indicates those emitting more than 300,000 tons per year.

3 (HUI 3), a multiple attribute utility scale which is often used in the medical field [6]. HUI3 has eight attributes: vision, hearing, speech, ambulation, dexterity, emotion, cognition, and pain. Each attribute has five or six levels. This means that, theoretically, total 972,000 health conditions can be assessed using HU13. The health risks for all residents in Japan from exposure to toluene were finally quantified as the loss of 197 years per annum, as shown in Table 1. If similar procedures were applied to other chemical substances or other kinds of risks, we could quantitatively compare their amounts of risk and would be able to allocate resources more rationally.

3.6 Cost effectiveness of risk reduction measures

Since atmospheric modeling and dose response functions clarified the linkage from emission sources to human health risks, we could simulate the health benefits of introducing emission control measures at the sources and express them as "QALYs gained". By estimating the annual costs of emission control measures, "cost per QALY gained" could be calculated by dividing costs by the QALYs gained. In this study, we calculated the cost-effectiveness in the case where 10 % of the total emission volume of toluene reported to the PRTR was reduced by installing Regenerative Thermal Oxidizers (RTOs). The cost per QALY gained was calculated to be approximately 1.6 years. Since it costs about JPY 33,000 to 100,000 to reduce 1 ton of toluene using RTOs (100,000 to 7,500 Nm^3/hr), it costs JPY 430 million to 1.3 billion annually to reduce 13,000 tons (10 % of emission volumes reported to the PRTR). Therefore, the "cost per QALY gained" was estimated to be approximately JPY 270 to 810 million. This indicator of cost-effectiveness can be used to prioritize risk reduction measures for various types of hazards other than chemical substances, such as infectious diseases, accidents, and natural disasters.

4 Discussion

Table 1. Annual loss of QALYs resulting from exposure to toluene in Japan

	Loss of QALYs (years)
Indoor sources	159
Mobile and low-emission sources	28
High-emission facilities	10
Total	197

The conventional method of risk assessment was originally created by combining elemental techniques that were developed by casting back from social demands to set environmental standards or to conduct screening-level risk assessment. These elemental techniques were established as standard techniques of risk assessment practices, and were respected as academic fields with their own experts. Therefore, when new social demands for risk assessment arise, experts of each elemental technique are expected to apply the existing methods without modification. Since the users of risk assessments are usually not experts, they cannot examine the methodologies of each elemental technique. Once the technology is established as a discipline, it evolves independently. However, leading-edge research does not necessarily satisfy new social demands. There is no guarantee that a necessary elemental technique that may fulfill new social demands will emerge endogenously from academic fields.

In this research, we sought the most suitable methods and revised each elemental technique to fulfill new social demands, i.e. different types of risks were compared and priorities were set based on cost effectiveness. We are aware that some of the assumptions lack firm foundations from an expert standpoint. Although these issues must be resolved, it is also important to accurately explain future research requirements to experts of the disciplines. We must also demonstrate that the methodology presented in this paper can be applied to other chemical substances and other types of risks.

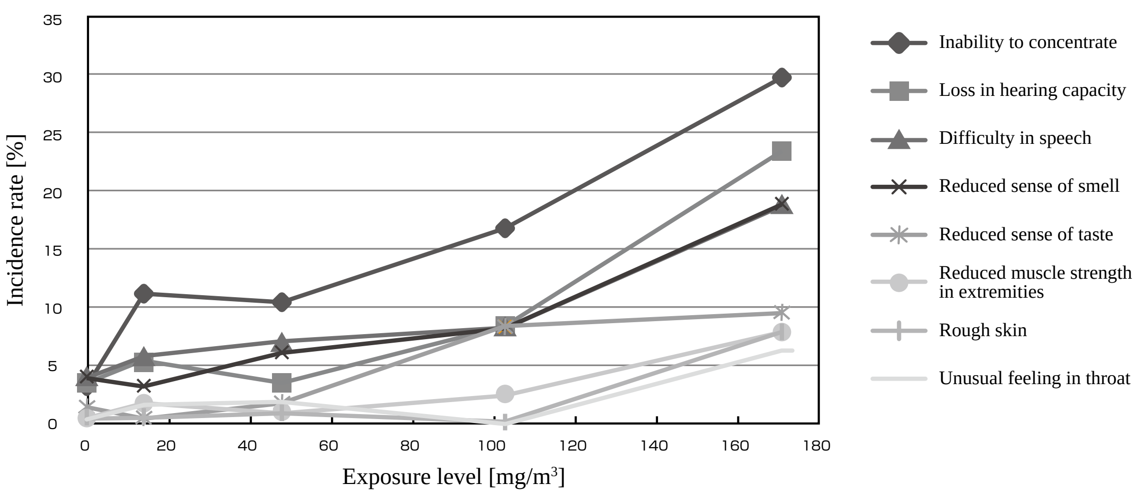

Fig. 5 Dose response relationships of subjective symptoms

Users of risk assessment, or risk managers, must convey the social demands or prescribe the methods of risk assessment to the risk assessment community to ensure that the risk assessment methodology can fulfill newest social demands. Although risk assessment is an interdisciplinary field, this does not mean that it is sufficient to combine works from different fields. We must make sure that each step of risk assessment is consistent, with the goal of fulfilling the given social demands.

Terminology

Term 1. The value is larger than 95 % of the measured values.

Term 2. Percent change in incidence rate of health effects due to intake of chemical substances.

References

[1] J. Nakanishi and A. Kishimoto: *Shosai risuku hyokasho sirizu 3:toruen* (Risk Assessment Document Series 3: Toluene), Maruzen, Tokyo (2005) (in Japanese).

[2] H. Higashino, K. Inoue, K. Mita, H. Shinozaki and H. Yoshikado: Bakuro · risuku hyoka taiki kakusan moderu (ADMER) zenkokuban no kaihatsu to kensho, kankyo kanri (Development and verification of the nationwide version of the atmospheric dispersion model for exposure and risk assessment (ADMER)), *Kankyokanri*, 40, 58-66 (2004) (in Japanese).

[3] Ministry of Health and Welfare: Kyojyu kankyochu no kihatsusei yuki kago butsu no zenkoku jittai chosa ni tsuite (Nationwide survey on the concentrations of volatile organic compounds in the residential environment), (1999) (in Japanese).

[4] M. Shiotsu, S. Yoshizawa, K. Ikeda, and A. Nozaki: Seikatsu jikan ni yoru okunai taizai jikanryo to katsudoryo: shitsunai kuki osen busshitsu ni taisuru bakuroryo hyoka ni kansuru kisoteki kenkyu sono 1 (Survey on human activity patterns according to time and place: basic research on the exposure dose to indoor air pollutants, Part 1), *J. Archit. Plann. Environ.* 511, 45-52 (1998) (in Japanese).

[5] H. Ukai, T. Watanabe, H. Nakatsuka, T. Satoh, S. J. Liu, X. Qiao, H. Yin, C. Jin, G. L. Li, and M. Ikeda: Dose-dependent increase in subjective symptoms among toluene-exposed workers, *Environ. Res.* 60(2), 274-289 (1993).

[6] D. Feeny, F. Furlong, M. Boyle and G.W. Torrance: Multi-attribute health status classification systems: Health utilities index, *PharmacoEconomics*, 7(6), 490-502(1995).

Received original manuscript September 19, 2007
Revisions received November 16, 2007
Accepted November 16, 2007

Author

Atsuo Kishimoto

Atsuo Kishimoto is a senior researcher of the Research Center for Chemical Risk Management of AIST. He obtained his Ph.D. in economics from Kyoto Univerisity, and joined the National Institute for Resources and Environment as a researcher in April 1998. He is the coauthor and coeditor of "Kankyo risuku manejimento hando bukku (Handbook of Environmental Risk Management)"(Asakura, 2003) and the coauthor of "Shosai risuku hyokasyo shirizu 3 toruen (Risk Assessment Document Series 3: Toluene)"(Maruzen, 2005).

Discussion with reviewers

1 The relationship between QALYs and healthy life expectancy

Comment and question (Akira Ono)

It is significant that different kinds of risks were assessed using common metric of loss of life years to enable comparison of the risks. Until now, discussions about the risk tended to be unrealistic such as, "it is absolutely safe" or "accidents never happen," but the application of QALYs to the risk assessment will enable reasonable judgments about risks and more flexible response to risks.

What is the relationship between the QALY as described in this paper and the so-called "healthy life expectancy?"

Answer (Atsuo Kishimoto)

"Healthy life expectancy" indicates the expected years of life in good or fairly good health, which correspond to the period when the QOL is not far from 1 in Figure 3. In other words, the QOL during period with illness or disability is considered to be virtually 0. In contrast, in the usual concept of "life expectancy," the QOL for the entire period of life is implicitly considered to be 1. The concept of QALYs is positioned between "life expectancy" and "healthy life expectancy," and reflects the health status most accurately.

2 Comparability with other types of risks

Question (Akira Ono)

Is it possible to apply the proposed methodology to accident risks, such as nuclear plant accidents caused by earthquakes, and compare these risks quantitatively with chemical substances? Is it possible to apply it to international risks such as for avian flu, or global risks such as climate change?

Answer (Atsuo Kishimoto)

Since QALYs are indicators that combine mortality effects (loss of life years) and morbidity effects (loss of QOL), it is applicable to almost all human health risks, including ones caused by nuclear plant accidents. In conventional risk assessment procedures, health risks are estimated according to worst-case scenarios. Since the extent of conservativeness in these different risks varies, it is difficult to compare them. The features of the proposed methodology are: 1) risks are expressed using the common metric of QALYs, and 2) risks are estimated as central tendency estimates, i.e. expected values. When these two conditions are satisfied, any risks are comparable to each other. Although it can also be applied to international risks, it seems to me that the issue of who will be the most affected will become the more important point than the aggregate risk values. The application of QALYs to intergenerational issues such as climate change remains to be discussed.

3 Uncertainty in the number of QALYs

Comment (Akira Ono)

I think that the work of making the scenarios for fulfilling new social demands, from top to bottom in Figure 2 was the key to this research. In doing so, it became gradually clear, that the existing elemental techniques must be revised or replaced by new ones. This is a great achievement.

However, I imagine that the work has not yet been completed although elemental techniques are being revised or developed. Assumptions were made without foundation. The quantification of human health risks and calculation of cost-effectiveness were implemented even though there was a huge uncertainty about the results. Is it possible to assess and quantify uncertainty in case sufficient data and information are lacking?

Answer (Atsuo Kishimoto)

In this paper, we set priority on deriving concrete numbers, since it was important to exemplify the whole process of proposed methodology, starting from social demands as indicated by the reviewer. As a result, we admit that the discussion on uncertainty i.e. the range of distributions remains to be addressed. At the same time, it is necessary to develop a simplified methodology that can be applied to cases with less data.

4 Compatibility of this article to the journal

Comment (Masayuki Kamimoto)

The content of this article is consistent with the spirit of *Synthesiology*, since it proposes a new methodology of risk assessment by adopting back-casting process and QALYs, and then applies the methodology to the case of toluene. It also describes the process of revising each step of risk assessment and obtains good results.

5 Comparability with other types of risks

Question (Masayuki Kamimoto)

You state, “We must also demonstrate that the methodology presented in this paper can be applied to other chemical substances and other types of risks”. Is this applicable to other substances with enough data?

Answer (Atsuo Kishimoto)

Yes, it is applicable to other substances with sufficient data. It is also applicable to other types of risks such as traffic accidents, since there is a wealth of statistical data. We are now investigating a methodology that is applicable to cases with scarce information.