them, 95 (15.1%) were obese and 117 (18.5%) were would-be obese ($23 \leq BMI < 25\ kg \cdot m^{-2}$).

To derive dietary patterns, we conducted a principal component analysis and used the eigenvalue greater than 1.0 criterion, which is widely accepted [19]. We put 18 principal components through this procedure that accounted for 66.5% of the total variance and assigned them to each purchase as dietary patterns. Because 18 components were too many for further analysis, we used 5 principal components as main dietary patterns (eigenvalue > 1.35; cumulative contribution, 29.2%) according to the scree plot. A total of 98 440 purchase data (63.0%) of 156 345 collected corresponded to the 5 main dietary patterns.

The main 5 dietary patterns are shown in Table 3. The most frequent pattern (principal component 1) was labeled as the "healthy" pattern due to the high correlation with side dishes, miso soup, vegetables, grain, and preparation by chilling. The second pattern was characterized by high intakes of Japanese main dishes, eggs, grains, dressing, preparation by simmer, and preparation by cooking; and therefore, called *traditional Japanese*. The third pattern, "Chinese," was correlated with Chinese main dishes,

Table 3
Factor loadings contributing to the 5 major dietary patterns among the AutoMealRecord system registrants identified by the principal component analysis

		Principal components[a]				
		1	2	3	4	5
		Healthy	Traditional Japanese	Chinese	Japanese noodles	Pasta
Food group	Japanese main dish	−0.222	0.438	–	–	–
	Western main dish	–	–	–	−0.317	–
	Chinese main dish	–	–	0.441	–	–
	Japanese noodles	–	–	–	0.200	−0.270
	Pasta	–	−0.205	–	–	0.276
	Chinese noodles	–	−0.214	–	–	−0.209
	Japanese rice	–	–	–	–	–
	Western rice	–	–	–	–	–
	Chinese rice	–	–	0.210	–	–
	Japanese combo meal	–	–	–	–	–
	Western combo meal	–	–	–	–	–
	Chinese combo meal	–	–	–	–	–
	Sandwich	–	–	–	–	–
	Garnishing	–	–	–	–	–
	Side dish	0.414	–	–	0.209	–
	Miso soup	0.366	–	–	–	–
	Dessert	–	–	–	–	–
	Bento	–	–	–	–	–
	Sauce	–	–	–	–	–
Main ingredient	Beef	–	–	0.257	–	–
	Pork	–	–	0.316	–	–
	Chicken	–	–	–	–	–
	Minced/processed meat	–	–	–	−0.226	−0.337
	Seafood	–	–	–	–	0.321
	Vegetables	0.324	–	–	–	–
	Eggs	−0.276	0.307	–	–	–
	Soy products	0.220	–	–	0.297	–
	Grains	0.224	0.263	–	–	–
	Fruits	−0.235	–	–	0.381	–
Cooking method	Simmer (niru)	–	0.270	–	–	–
	Stir-fry (itameru)	–	–	0.589	–	–
	Grill/roast (yaku)	–	–	−0.254	–	0.308
	Deep-fry (ageru)	–	–	–	–	−0.478
	Steam (musu)	–	–	–	–	–
	Dress (aeru)	–	0.213	–	–	–
	Chill (hiyasu)	0.269	–	–	0.358	–
	Boil (yuderu)	–	–	–	0.410	–
	Cook (taku especially rice)	–	0.437	–	–	–
	Marinade (hitasu)	–	–	–	–	–
Contribution[b]		9.7%	7.1%	4.6%	4.2%	3.6%

Values are factor loadings (>.200) contributing to the 5 dietary patterns. We analyzed the dietary patterns based on 39 classification codes (Table 1) using the principal component analysis, and the 5 most frequent princial components are shown (eigenvalue > 1.35; cumulative contribution, 29.2%).

[a] Principal components are labeled according to the factor loadings.

[b] Contribution shows to what extent the principal component accounts for the variance.

Table 4
BMI prediction model constructed by multiple linear regression analysis with stepwise selection using demographic data, dietary patterns, annual mean of dietary intake, and body composition

			Explanatory variables[a]									
			Demographics[b]		Preference of dietary pattern[c]							
			Sex	Age	Healthy	Japanese	Chinese	Japanese Noodles	Pasta			
Whole	–	–	0.089	−1.698				0.040				
	Sex	Age (y)										
Subgroups	Male	20s	–	–				0.214				
		30s	–	–								
		40s	–	–								
		50s	–	–								
	Female	20s	–	–				0.097				
		30s	–	–								
		40s	–	–				0.066				
		50s	–	–					−0.050			

			Explanatory variables[a]									
			Dietary intake[d]					Cafeteria use				
			Energy (Kj)	Protein (% energy)[e]	Fat (% energy)[e]	Fiber (g)	Salt (g)	Frequency of cafeteria use[f]	Lunchtime use (%)[g]	Preference of service company (%)[h]	Frequency of body measurement[i]	R^2
Whole	–	–	0.001	21.887		−0.244		−0.041	−1.085		0.088	0.219
	Sex	Age (y)										
Subgroups	Male	20s	0.006			−0.963						0.359
		30s	0.0009							1.796	0.138	0.045
		40s	0.0008	25.650				−0.057		−2.123		0.097
		50s									−0.537	0.060
	Female	20s			11.463					4.166	0.113	0.530
		30s		34.575							0.108	0.126
		40s									0.072	0.211
		50s	−0.001	−32.612		−1.262		−0.418	14.299	−11.991		1.000

The BMI prediction models are constructed for whole subjects (whole) and each age-sex subgroups (subgroups) by multiple linear regression analysis with stepwise selection. Values are regression coefficients that are selected as explanatory variables of each model.

[a] All variables used for calculation of the models are listed.
[b] Sex and age were not used as explanatory variables in the subgroup analysis.
[c] (Proportion of each dietary pattern) = (annual frequency assigned to the dietary pattern)/(annual total frequency of cafeteria use).
[d] Individual mean of annual dietary intake during October 1, 2007, to September 30, 2008.
[e] Energy-adjusted protein/fat content in percentage of energy intake.
[f] Individual frequency of monthly cafeteria use during the data collection period.
[g] Proportion of meals eaten at lunchtime.
[h] Two different food service companies serve meals at the cafeterias, and the preference of one company was calculated as a variable.
[i] Three-month frequency of body measurement at the body measurement booth of the company cafeterias.

Table 5
Comparison of BMI prediction models by AIC, BIC, and cross validation

		Age	n[a]	Measured BMI ≥ 23 (%)[b]	R^2	AIC[c]	BIC[d]	Cross validation[e]
Whole	–	–	634	33.4	0.219	1249.6	1294.1	0.688
Subgroups	Male	20s	55	25.5	0.359	113.8	121.8	0.709
		30s	153	35.3	0.045	314.9	327.0	0.569
		40s	192	47.4	0.097	378.3	394.6	0.589
		50s	63	50.8	0.060	148.8	153.1	0.476
	Female	20s	24	0.0	0.530	15.9	21.8	0.958
		30s	86	9.3	0.126	130.0	137.4	0.895
		40s	48	16.7	0.211	73.6	79.2	0.854
		50s	9	33.3	1.000	−70.6	−69.0	0.778

The BMI prediction models are constructed for whole subjects (whole) and each age-sex subgroups (subgroups). The fitness of the models are compared using multiple correlation coefficient (R^2), AIC, BIC, and cross validation.

[a] Number of participants in the model.

[b] (Measured BMI ≥ 23 kg·m^{-2}) = (number of participants whose BMI are > 23 kg·m^{-2})/(number of all participants in the model).

[c] Akaike information criterion. Smaller values mean better model fitness.

[d] Bayesian information criterion. Smaller values mean better model fitness.

[e] Values are the prediction accuracy rate calculated by comparing measured BMI with predicted one, which are dichotomized (cutoff point: BMI = 23 kg·m^{-2}). Predicted BMI is computed using leave-one-out cross validation method.

Chinese rice, beef, pork, and preparation by stir-fry. The fourth was called *Japanese noodles* due to the high amount of Japanese noodles and "boiling" method that included Japanese and Chinese noodles. This fourth pattern ("Japanese noodles") was also correlated with side dishes, soy products, fruits, and preparation by chilling, which meant consumption of chilled side dishes such as tofu, natto, and fruit cups. The fifth was labeled as the "pasta" pattern and was correlated with pasta, seafood, and preparation by grilling/roasting, which indicates that pasta or grilled seafood was purchased.

Individual portions of the 5 dietary patterns for each person were calculated by dividing the frequency of each pattern with total frequency throughout the year, called *preference of dietary patterns*. Any one of the 5 dietary patterns was preferred by 569 (89.8%) participants. Comparing the distribution of the dietary patterns by age-group and sex, the "healthy" ($\chi^2_{21} = 41.0$; $P < .01$) and "pasta" ($\chi^2_{21} = 77.8$; $P < .001$) patterns were significantly biased among groups, but "traditional Japanese" ($\chi^2_{21} = 29.3$; $P = .11$), "Chinese" ($\chi^2_{21} = 21.0$; $P = .46$), and "Japanese noodles" ($\chi^2_{21} = 18.6$; $P = .61$) were not significantly different among sex and age-groups.

According to the Pearson's correlation coefficient between BMI and the dietary patterns, a high proportion of "pasta" pattern correlated with low BMI ($r = -0.15$; $P < .001$). Men ($r = -0.36$; $P < .001$) and younger people ($r = 0.26$; $P < .001$) tended to have higher BMI.

Current BMI was predicted by multiple linear regression analysis with stepwise selection using the preference of dietary patterns calculated as mentioned earlier and 11 variables collected by the AutoMealRecord system (Table 4). Preference of the 5 major dietary patterns did not correlate with any macronutrients ($|r| < 0.2$; $P < .001$), and we believe it is justifiable to put the variables together into a regression equation. For all 634 participants (the whole model), the analysis indicated that current BMI positively correlated with male gender, preference of "Japanese noodles," mean energy, protein content, and frequency of body measurement. Current BMI correlated negatively with age, dietary fiber, and lunchtime cafeteria use ($R^2 = 0.22$). This regression model predicted "would-be obese" participants (BMI > 23) with 68.8% accuracy.

Different models were constructed for each age-sex subgroup (Table 4). Akaike information criterion and BIC were the highest in the whole model and relatively lower in the female models (Table 5). The would-be obese prediction accuracy rate was more than 70% in the models for males in their 20s and females of all ages (Table 5).

4. Discussion

This exploratory data analysis has shown that the data accumulated in the AutoMealRecord system could explain, to some extent, current BMI. Although further evaluation will be required, we can assume that the AutoMealRecord system is reliable and valuable as a tool for health promotion and lifestyle disease prevention.

We observed 5 major dietary patterns, but there was not a "Western" pattern. This makes it difficult to compare our results with other Western studies [20-22]. The "healthy" and "traditional Japanese" patterns in this study were nearly identical to the healthy and Japanese traditional patterns, by Okubo et al [23]. The "healthy" pattern, which had a high loading of vegetables, fruits, fish, and soy products, was associated with lower BMI, whereas the "Japanese traditional" pattern, which had a high loading of rice, miso soup, and soy products, was associated with higher BMI among Japanese female students [23]. Although we could not see

the direct correlation between these patterns and BMI, it is suggested that the "healthy" and "traditional Japanese" patterns are the major dietary patterns in a wide age range of Japanese people.

In our study, BMI was correlated only with the "pasta" preference, calculated as annual proportions of the "pasta" pattern for each person. Higher "pasta" preference was associated with lower BMI ($r = -0.15$; $P < .001$). Those who preferred the "pasta" pattern tended to be female ($r = 0.29$; $P < .001$) and younger ($r = -0.13$; $P < .001$). These data indicate that the association between "pasta" and BMI was not a causal association but a self-selection bias. In Japan, pasta is regarded as fashionable and healthy food. This might explain why the "pasta" pattern includes grilled seafood, which is relatively lean rather than high-fat dish such as meat or fried dish.

The multiple linear regression analysis illustrated that higher BMI was associated with higher frequency of the "Japanese noodles" pattern, which is likely due to the combination of Japanese noodles (soba or udon) with side dishes such as tofu, natto, or fruit cups. This association may be partly understood by the role of dietary glycemic index (GI) and glycemic load (GL). A positive association of dietary GI and GL with BMI among American and Japanese people has been reported [23-25]. Udon (wheat noodles) are a major high GI/GL food in Japan and soba (buckwheat noodles) have a relatively low GI/GL. Furthermore, those who like to eat Japanese noodles may have little time to enjoy their meal because they can be quickly prepared and eaten. Eating quickly has been shown to be associated with being overweight in Japanese men and women [26]. Most Japanese regard Japanese noodles as "daddy" food as there are many stand-up-eating soba shops all over Japan filled with working-age men. Although we cannot know the type of noodles and speed of eating from the data, the association of BMI and preference of Japanese noodles seems to support previous studies.

The obesity prediction model for all participants indicated that males and younger participants tended to have a higher BMI. This is consistent with the National Health and Nutrition Survey [2]. However, preference of dietary patterns was not equally distributed among sex and age-groups. An example of this is the "pasta" pattern mentioned above. Thus, we constructed BMI prediction models by sex and age-groups. More than half of the models contained the annual mean of total energy and frequency of body measurement as explained variables. Rolls et al [27] has shown that large food portion sizes leads to excess energy intake. The company cafeteria we observed served 2 portion sizes (normal and small) for almost every main dish, and the high energy intake was probably the result of choice as there was an indication that a tendency to choose high energy content foods was associated with higher BMI. Frequency of body measurement seems to reflect body consciousness. Among females in their 20s to 40s and males in their 30s, more frequent measurement was associated with higher BMI. It is possible that these participants had already been conscious of their body shape and kept assessing themselves. On the other hand, less frequent measurement was associated with higher BMI among males in their 50s. This group is the central target of the annual medical checkup act mentioned above. It would appear that body conscious men had tried to keep or improve their shape, although this point regarding body consciousness and satisfaction needs to be examined in more detail in future studies.

The models for males and females in their 20s and females in their 40s and 50s showed relatively higher multiple regression coefficients. The more variables that were selected, the higher multiple regression coefficients tended to be, and so we did not think it was appropriate to use multiple regression coefficients to compare the models. We examined the model fit by AIC and BIC as these criteria are independent of the number of variables [12-15]. Compared with the whole model, AIC and BIC values were relatively smaller in the sex and age-group models, which means that these models were better than the whole model. The data fitness of models for males in their 30s and 40s was relatively lower, despite their frequent cafeteria use. It was suggested that lifestyle outside the cafeteria had a greater impact on these age-groups.

On the model validation, we used BMI of 23 $kg \cdot m^{-2}$ or greater as the cutoff value instead of the Japanese standard obesity criteria, BMI of 25 $kg \cdot m^{-2}$ or greater [17]. We did this for 3 reasons. First, this was in accordance with a previous study on the AutoMealRecord system by Ishida [18]. That study showed that a dietary education program had more effect on people with a BMI between 23 and 25 $kg \cdot m^{-2}$. The result suggested that obese people more than 25 $kg \cdot m^{-2}$ could not easily change their dietary lifestyle and that preventive intervention for nonobese people would have a greater impact on public health. Secondly, in one study [3] founded on a longitudinal analysis of data from a population-based cohort study, it was reported that Japanese men and women aged 40 to 49 years tended to gain their weight for a period of 10 years. In the same study, nearly 10% of men aged 40 to 49 years who had a BMI of less than 25 $kg \cdot m^{-2}$ at baseline became obese (BMI $\geq$ 25 $kg \cdot m^{-2}$) during the 10-year period. Therefore, we decided to focus on prediction of "would-be obese" people with a BMI more than 23 $kg \cdot m^{-2}$ for model validation. The third reason we used BMI of 23 $kg \cdot m^{-2}$ or greater as the cutoff value is that our study participants were a relatively lean population compared to the general Japanese. Few obese people whose BMI were more than 25 $kg \cdot m^{-2}$ were in the female subgroups (n = 0-4), which might destabilize the prediction capability. The models tended to predict lower BMI, and "would-be obese" (BMI $\geq$ 23 $kg \cdot m^{-2}$) prediction accuracy was better among females and young males whose measured BMI was relatively low (prediction accuracy, 70.9%-100.0%). The models for males in their 30s to 50s were less well fitted, as indicated by AIC and BIC. We must consider lifestyles outside the company cafeteria to check their health status more precisely.

It should be noted that our study has several limitations because of the nature of the exploratory data analysis. First of all, the data we used were not collected for this study. We only "mined" the data accumulated in the AutoMealRecord system database, which were used for the system users to check the feedback on their dietary record on a weekly basis. We could not make the dish classification or nutrient calculation. Second, the dataset included only the purchased data from a company cafeteria. We cannot know their whole dietary life and what the participants did or did not eat for certain. In addition, the participants were not randomly selected but were voluntary registrants from employees of a major electric company. They may have higher incomes, higher education, or better health consciousness. Although high education had no association with dietary consciousness in the study by Carrera [28], it is possible that our study participants were more health conscious or at least healthier because of their lower BMI than the national average in Japan. We must be cautious about generalization of the results. Also, we had no way of knowing the disease history of the participants and could not eliminate effects on dietary life from any previous or concurrent disease. The company we studied had employees who were at risk for lifestyle diseases according to dietary education programs using the AutoMealRecord system and personal comments from dietitians. The at-risk employees were identified by system ID. We believe that almost everyone who registered in September 2008 was healthy. To overcome those limitations, we are conducting a Web-based questionnaire survey asking socioeconomic status, lifestyle, disease history, and whole nutrition using a validated food frequency questionnaire.

In addition, we used self-reported height and weight that are said to be subject to substantial measurement error [29]. However, Stommel et al [30] recently suggested that nonobese people are not likely to overreport their weight according to data from the continuous National Health and Nutrition Examination Survey. It is reported that 90.4% of people with "normal weight," whose BMI was between 18.5 and 25 using physically measured weight, were classified correctly as "normal weight" using self-reported weight [30]. It was also suggested that high-income respondents tend to report their weight more accurately [31]. In our study, participants were relatively lean (85.0% [539/634] of them had BMI < 25) and most of them might be college educated according to our recent study on the AutoMealRecord system users (data not shown). It cannot be denied that self-reporting bias exists within our data, but we can assume it is minimally suppressed.

Furthermore, our dietary pattern analysis has some limitations. We conducted a principal component analysis to reduce the number of purchase-related variables and make easy-to-understand dietary patterns. Principal component analysis itself is mathematically clear-cut as it does not depend on subjective decisions but is based on a simple mathematical rotation of axes. However, we made a subjective decision to use 5 major principal components for analytical simplicity and interpretability. An exploratory factor analysis with 5 factors had shown a similar pattern to the result of principal component analysis, and we confirmed the reliability and robustness of the dietary patterns beyond analytical methods.

Beyond these limitations, it is notable that we could show the ability of the AutoMealRecord system to predict would-be obese people accurately by only using the data in the system. Although further evaluation will be required, we can assume that the data were reliable and valuable as tools for lifestyle disease prevention. We are now conducting a Web-based questionnaire survey and a messaging intervention using e-mail for the AutoMealRecord system users.

Segmentation and message tailoring are the most important methods for effectively changing health behavior by health communications [32]. Previous dietary education for adults has paid little attention to age and sex but has focused more upon BMI or health consciousness. This study indicates that tailoring the health message, with consideration given to age and sex, is also important for improving the effectiveness of dietary education to prevent lifestyle diseases.

In conclusion, we are justified in believing that the AutoMealRecord system is valuable for further consideration as a health care intervention tool by analyzing the data with data mining approach. With the spread of smart cards, the AutoMealRecord system could be a powerful infrastructure to maintain a healthy dietary lifestyle not only for company employees but also for the general public in the future.

Acknowledgment

This project was funded by Pfizer Health Research Foundation. We thank Susumu Murakami for technical advice regarding the AutoMealRecord system.

References

[1] Kelly T, Yang W, Chen C, Reynolds K, He J. Global burden of obesity in 2005 and projections to 2030. Int J Obes 2008;32:1431-7.
[2] Ministry of Health Labour and Welfare. Annual Report of the National Nutrition Survey in 2008. Tokyo (Japan): Ministry of Health , Labour and Welfare; 2008 [in Japanese].
[3] Matsushita Y, Takahashi Y, Mizoue T, Inoue M, Noda M, Tsugane S. Overweight and obesity trends among Japanese adults: a 10-year follow-up of the JPHC Study. Int J Obes (Lond) 2008;32:1861-7.
[4] Ministry of Health Labour and Welfare. Guideline for operation of annual medical checkup and special health guidance. Tokyo (Japan): Ministry of Health, Labour and Welfare; 2008 [in Japanese].
[5] Kirin Brewery. A Survey on People's attitude toward "health." Kirin Brewery News Release. Tokyo (Japan): Kirin Brewery; 2008 [in Japanese].
[6] The Yomiuri Shimbun. Revival of company cafeterias. Tokyo (Japan): The Yomiuri Shimbun Morning edition; 2008 [in Japanese].
[7] Matsuura T. An overview of Japanese Cafeteria Plans—objectives and results. Tokyo (Japan): NLI Research Institute; 1998. http://www.nli-research.co.jp/english/socioeconomics/1998/li9803.html.

[8] Murakami S. Health Management System Using Meal Record collected by IC Smart Card. Inst Image Inf Telev Eng 2007;61: 1100-3 (in Japanese) See also: http://www.cooca.jp/fdk/system.html.
[9] Witten I, Frank E. Data Mining, second edition—practical machine learning tools and techniques. Massachusets: Morgan Kaufmann; 2005.
[10] Ministry of Education Culture Sports Science and Technology, Ministry of Health Labour and Welfare. The Ethical Guidelines for Epidemiological Research. Tokyo (Japan): Ministry of Education Culture Sports Science and Technology, and Ministry of Health Labour and Welfare; 2007 [in Japanese].
[11] Takeuchi K, Ichikawa S, Ohashi Y, Kishimoto A, Hamada C. Data analysis using SAS. Tokyo (Japan): University of Tokyo Press; 1987.
[12] Akaike H. Information theory and an extension of the maximum likelihood principle. Second International Symposium in Information Theory, Budapest, Hungary: Akademiai Kaiado; 1972.
[13] Sakamoto Y, Ishiguro M, Kitigawa G. Akaike information criterion statistics. Tokyo (Japan): KTK Scientific; 1986.
[14] Agresti A. Categorical data analysis. New York: John Wiley&Sons; 1990.
[15] Schwarz G. Estimating the dimension of a model. Ann Stat 1978;6: 461-4.
[16] Shao J. Linear model selection by cross-validation. J Am Stat Assoc 1993;88:486-94.
[17] Matsuzawa Y, Inoue S, Ikeda Y, Sakata T, Saito Y, Sato Y, et al. The new diagnostic criterion for obesity. J Jpn Soc Stud Obes 2000;6:18-28.
[18] Ishida H. Nutritional Management Guideline and Compliance on Registered Cafeterias. Tokyo (Japan): The Report of Comprehensive Research on Health Science, funded by Ministry of Health, Labour and Welfare; 2003-2005 [in Japanese].
[19] Washio Y, Ohashi Y. Multidimensional data analysis. Tokyo (Japan): Iwanami Books; 1989.
[20] Maskarinec G, Novotny R, Tasaki K. Dietary patterns are associated with body mass index in multiethnic women. J Nutr 2000;130:3068-72.
[21] Newby PK, Muller D, Hallfrisch J, Andres R, Tucker KL. Food patterns measured by factor analysis and anthropometric changes in adults. Am J Clin Nutr 2004;80:504-13.
[22] Panagiotakos DB, Pitsavos C, Skoumas Y, Stefanadis C. The association between food patterns and the metabolic syndrome using principal components analysis: the ATTICA Study. J Am Diet Assoc 2007;107:979-87.
[23] Okubo H, Sasaki S, Murakami K, Kim MK, Takahashi Y, Hosoi Y, et al, the Freshmen in Dietetic Courses Study II group. Three major dietary patterns are all independently related to the risk of obesity among 3760 Japanese women aged 18-20 years. Int J Obes 2008; 32:541-9.
[24] Ma Y, Olendzki B, Chiriboga D, Hebert JR, Li Y, Li W, et al. Association between dietary carbohydrates and body weight. Am J Epidemiol 2005;161:359-67.
[25] Murakami K, Sasaki S, Takahashi Y, Okubo H, Hosoi Y, Horiguchi H, et al. Dietary glycemic index and load in relation to metabolic risk factors in Japanese female farmers with traditional dietary habits. Am J Clin Nutr 2006;83:1161-9.
[26] Maruyama K, Sato S, Ohira T, Maeda H, Kubota Y, Nishimura S, et al. The joint impact on being overweight of self reported behaviours of eating quickly and eating until full: cross sectional survey. BMJ 2008;337:a2002.
[27] Rolls BJ, Roe LS, Meengs JS. Larger portion sizes lead to a sustained increase in energy intake over 2 days. J Am Diet Assoc 2006;106: 543-9.
[28] Carrera PM, Gao X, Tucker KL. A study of dietary patterns in the Mexican-American population and their association with obesity. J Am Diet Assoc 2007;107:1735-42.
[29] Rothman K. BMI-related errors in the measurement of obesity. Int J Obes 2008;32:S56-9.
[30] Stommel M, Schoenborn C. Accuracy and usefulness of BMI measures based on self-reported weight and height: findings from the NHANES & NHIS 2001-2006. BMC Public Health 2009;9:421.
[31] Vaughan C, Sacco W, Beckstead J. Racial/ethnic differences in body mass index: the roles of beliefs about thinness and dietary restriction. Body Image 2008;5:291-8.
[32] Rogers EM. The field of health communication today: an up-to-date report. J Health Commun 1996;1:15-23.

Ishikawa and Kiuchi *BioPsychoSocial Medicine* 2010, 4:18
http://www.bpsmedicine.com/content/4/1/18

BIOPSYCHOSOCIAL MEDICINE

REVIEW Open Access

Health literacy and health communication

Hirono Ishikawa*, Takahiro Kiuchi

Abstract

Health communication consists of interpersonal or mass communication activities focused on improving the health of individuals and populations. Skills in understanding and applying information about health issues are critical to this process and may have a substantial impact on health behaviors and health outcomes. These skills have recently been conceptualized in terms of health literacy (HL). This article introduces current concepts and measurements of HL, and discusses the role of HL in health communication, as well as future research directions in this domain. Studies of HL have increased dramatically during the past few years, but a gap between the conceptual definition of HL and its application remains. None of the existing instruments appears to completely measure the concept of HL. In particular, studies on communication/interaction and HL remain limited. Furthermore, HL should be considered not only in terms of the characteristics of individuals, but also in terms of the interactional processes between individuals and their health and social environments. Improved HL may enhance the ability and motivation of individuals to find solutions to both personal and public health problems, and these skills could be used to address various health problems throughout life. The process underpinning HL involves empowerment, one of the major goals of health communication.

Introduction

Health communication, i.e., interpersonal or mass communication activities focused on improving the health of individuals and populations [1], has emerged as one of the most important public health issues in this century. The Healthy People 2010 project in the US suggests that health communication can contribute to all aspects of disease prevention and health promotion and that it is relevant to a number of domains including (1) health professional-patient relations, (2) individuals' exposure to, search for, and use of health information, (3) individuals' adherence to clinical recommendations and regimens, (4) construction of public health messages and campaigns, (5) dissemination of individual and population health risk information, that is, risk communication, (6) images of health in the mass media and the culture at large, (7) education of consumers about how to gain access to the public health and health care systems, and (8) development of telehealth applications.

People in modern society are expected to actively engage in the management of their health and to make a wide range of health decisions. Sound health decisions require comprehensible health information that is accessible and appropriate to the needs and cultural and social backgrounds of individuals [2]. Although health care professionals have historically been the primary sources of health and medical information, the increase in media reports and the rapid expansion of the Internet have rendered other sources more available to the general public [3-6]. Thus, skills in understanding and applying information about health issues may have a substantial impact on health behaviors and health outcomes. These skills have recently been conceptualized as health literacy (HL).

One of the objectives related to health communication in the US Healthy People 2010 project involves improving the HL of persons with inadequate or marginal literacy skills. Indeed, significant concern that people with limited HL may not be able to adequately understand health information, even in the presence of access to such information and related services, has emerged. In some cases, more information may actually cause feelings of confusion and powerlessness instead of facilitating sound health decisions. Even when health information is not intentionally sought, it may be provided by the media or by anyone with whom individuals communicate. The need for improved HL has become apparent as the number of health information sources that are easily accessed by the general public has increased in the absence of established

* Correspondence: hirono-tky@umin.ac.jp
Department of Health Communication, Graduate School of Medicine, The University of Tokyo, 7-3-1 Hongo, Bunkyo, Tokyo, Japan

assurances of the quality of the information provided by such sources.

The purpose of this paper is to introduce currently used concepts and measures of HL and to present an example of a study of HL in Japan. We then discuss the role of HL in health communication and outline a future research agenda for this domain.

The concept of health literacy

In general, literacy is the ability to read, write, and speak a language in the service of understanding and solving problems with sufficient proficiency to function at work and in society, achieve goals, and develop knowledge and individual potential (US Congress, National Literacy Act of 1991, Public Law 102-73, 1991). The notion of HL is based on this concept of literacy and generally refers to literacy in the context of health and healthcare. Given that basic literacy skills are required for health literacy, it is reasonable to assume that individuals with limited literacy also have limited HL. Indeed, previous studies have reported significant associations between measures of literacy and measures of functional HL, such as the Rapid Estimate of Adult Literacy in Medicine (REALM) [7]and the Test of Functional Health Literacy in Adults (TOFHLA) [8]. On the other hand, it has been noted that even individuals with adequate general literacy might not have adequate HL because the literacy demands in the context of healthcare are frequently more complex than those in the context of everyday life [9].

Several definitions of HL are currently used; these share the basic concept of literacy, but vary in scope. The US Healthy People 2010 project refers to HL as "the degree to which individuals have the capacity to obtain, process, and understand basic health information and services needed to make appropriate health decisions" [10]. Compared with earlier definitions of HL that focus on patients in healthcare settings and their understanding of medical information, this definition includes individuals outside of clinical settings and also links health literacy to the promotion of health and preventive behaviors.

Another well-recognized definition, proposed by the World Health Organization (WHO), defines HL more broadly as "the cognitive and social skills which determine the motivation and ability of individuals to gain access to, understand, and use information in ways which promote and maintain good health" [1]. This version also suggests that HL entails a level of knowledge, personal skills, and confidence that enables making changes in personal lifestyles and living conditions to improve personal and community health. Thus, this definition includes issues critical to the empowerment of patients. It also focuses not only on the cognitive elements of comprehending, analyzing, and applying health information to decisions about health, but also on the social skills involved in those interactions with other people and society (e.g., communication, negotiation, and organization) that are necessary for transforming decisions into practice. This conceptualization also refers to *motivation* in addition to *ability*. Based on this conceptualization, Nutbeam [11] proposed a model of HL that includes three levels and assumes the existence of benefits to both the *individual* and the *population* at each level: (1) basic/functional literacy, including skills for reading and writing that enable effective functioning in everyday situations, which is broadly compatible with the narrow definition of HL; (2) communicative/interactive literacy, including more advanced skills that enable active participation in everyday activities, extracting and understanding information from different sources, and applying new information to changing circumstances; and (3) critical literacy, including more advanced skills for critically analyzing information and using this information to exert greater control over life events and situations.

As the field of HL has expanded in scope and depth, the term HL itself has come to have different meanings. Nutbeam [12] distinguished two different approaches to HL: HL as clinical "risk" and HL as personal "asset." According to the former, HL is considered to be a set of individual literacy skills that act as a mediating factor in health and clinical decision making. The definition of HL according to the US Healthy People 2010 project is linked to this conceptualization. In contrast, the conceptualization of HL as a personal asset has evolved from public health and health promotion perspectives. In this context, HL is a means for enabling individuals to exert greater control over their health as well as over the range of personal, social, and environmental determinants of health; this corresponds to the definition issued by WHO. A similar distinction has been proposed by Peerson and Saunders [13]: "medical literacy," which is related to individuals as patients within health care settings, versus "health literacy," which is related to everyday life.

Measurement of health literacy

In general, literacy includes a variety of skills beyond reading and writing, such as numeracy, listening, and speaking, and relies on cultural and conceptual knowledge [9]. Nonetheless, most existing measures of HL have focused primarily on reading comprehension and numeracy. Typically, the REALM, the TOFHLA, and their short versions have been used in the US in clinical situations as screening tools to identify patients with limited HL.

A recent review identified 19 instruments for measuring HL that were published between 1990 and 2008 and examined their content and psychometric properties [14]. These included instruments that directly test an individual's abilities (e.g., the REALM, the TOFHLA, the Newest Vital Sign [15]), self-reports of abilities (e.g., Functional, Communicative, and Critical Health Literacy

Scales [16], the Set of Brief Screening Questions [17]), and population-based proxy measures (e.g., the National Assessment of Adult Literacy [18], and the Health Activities Literacy Scale [19]). They concluded that the composition of the underlying constructs and the nature of the content varied widely across HL instruments, rendering it difficult to interpret and compare studies.

It has also been noted that none of the existing instruments appears to completely measure the concept of HL as defined in the previous section. In particular, measurements of communicative/interactive and critical HL have lagged far behind instruments addressing functional HL [20]. Much work remains in the effort to develop more comprehensive measures that will assess individual HL with respect to an individual's ability to access, understand, and use health information in ways that promote and maintain good health [12]. Although a few recent studies have developed self-report measures of communicative and critical HL and examined their impact on health behaviors and outcomes [16,21], the development of an objective and direct measure for communicative and critical HL may pose a greater challenge than the development of such a measure for functional HL, such as the REALM and the TOFHLA. Skills relevant to clinical encounters may be assessed with a coding system applied to recorded communication between patient and healthcare providers, such as the Roter Interaction Analysis Systems [22], whereas skills relevant to other settings, such as those in which information obtained from the mass media or Internet is sought or used, may be more difficult to assess. Assessment difficulties also derive from the fact that the skills necessary will vary depending on the demands placed on the patient by the environment, including healthcare providers, healthcare systems, and the media. Thus, an HL level that is "adequate" in one situation may be inadequate in another situation, and this context dependence is especially true for communicative and critical HL.

In this sense, it is likely that different measurement tools will be required for measuring HL in different contexts [23]. Although previous instruments have approached HL as a quality characterizing the individual, HL is now seen as also based on interactions between an individual's skills and the demands of the society in which the individual lives, including healthcare providers, the healthcare system, the media, and the community [9,24]. Thus, an individual's HL should be defined and assessed in relation to the ability of the society to communicate health information in a manner appropriate to the audience (i.e., the HL of the population). However, further difficulties may arise in the development of an objective measurement of this type of HL [24].

A health literacy study in Japan

In our previous study, we developed self-rating scales measuring the functional, communicative, and critical HL of patients with chronic diseases [16]. Item content and mean scores from this study are shown in Table 1. Functional HL was assessed with five items that examined the extent to which patients experienced difficulties in reading the instructions or leaflets provided by hospitals and pharmacies (Cronbach's α = 0.84). Communicative HL was evaluated with five items that assessed the extent to which patients extracted and communicated diabetes-related information since they were diagnosed with this disease (α = 0.77). Critical HL was assessed with four items that focused on the extent to which patients had critically analyzed diabetes-related information and used it to make decisions (α = 0.65). Higher scores on this HL scale were generally associated with better knowledge of diabetes, a greater number of information sources, and higher self-efficacy with respect to diabetes self-care. Furthermore, patient HL, especially communicative HL, was related to the process of communicating with physicians during medical visits [25]. Moving beyond previous measures focusing solely on functional HL, this HL scale included three levels of HL, each of which may reflect different effects on health behaviors and outcomes. This measure also proved to be easy to administer in a clinical setting.

Based on this study among diabetes patients, a short version of the communicative and critical HL scale for general populations was validated in our study of office workers [21]. Item content and mean scores are shown in Table 2. In our analyses, higher HL was associated with healthy lifestyles and more effective coping with job stress as well as with fewer somatic symptoms.

One of the limitations of these studies was that we were unable to examine the relationship between our new HL scales and the existing standard measures of functional HL, such as the TOFHLA or REALM, because they were not available in Japanese at the time of our study. We noted that this issue should be explored in a future study with an English-speaking population to further revise and validate our HL scales in the service of enhancing their utility. After publication of these articles, several researchers in the US, Australia, the Netherlands, and Germany contacted us for validation of the HL scales in their countries.

Conclusions

Although studies of HL have increased dramatically during the past few years, a gap between the conceptual definition of HL and its application remains. More specifically, studies on communicative/interactive and critical HL are particularly limited. Furthermore, HL should be considered not only as a characteristic of an individual, but also as a

Table 1 Item content and means of the Functional, Communicative and Critical HL scales

	Mean	SD
Functional health literacy	**3.39**	**0.75**
In reading instructions or leaflets from hospitals/pharmacies, you....		
found that the print was too small to read	3.19	1.12
found characters and words that you did not know	3.41	0.88
found that the content was too difficult	3.43	0.84
needed a long time to read and understand them	3.27	1.04
needed someone to help you read them	3.65	0.86
Communicative health literacy	**2.56**	**0.70**
Since being diagnosed with diabetes, you have....		
collected information from various sources	2.43	1.04
extracted the information you wanted	2.18	1.00
understood the obtained information	2.89	0.88
communicated your thoughts about your illness to someone	2.70	0.91
applied the obtained information to your daily life	2.60	0.99
Critical health literacy	**1.96**	**0.63**
Since being diagnosed with diabetes, you have....		
considered whether the information was applicable to your situation	2.71	0.98
considered the credibility of the information	1.87	0.92
checked whether the information was valid and reliable	1.76	0.96
collected information to make health-related decisions	1.51	0.77

Note: The theoretical range: 1-4.

feature of interactions involving an individual's HL and his/her health and social environments.

In contrast to the high-risk approach adopted in traditional health education, which seeks to protect susceptible individuals [26] through such actions as screening those at high risk during health checkups to provide health counseling, the concept of HL may facilitate the development of a population-based approach. Such an approach would seek to control the causes of health problems, including eliminating the barriers that prevent individuals with limited HL from participating in the healthcare process and improving the HL of the population as a whole. Previous HL intervention programs have frequently attempted to decrease specific barriers affecting those with limited HL, including teaching healthcare providers to better communicate with patients with limited HL or developing simple and attractive health education materials pitched at those with lower reading levels [27-29]. Future interventions should be expanded to include methods for improving popular HL through school-based health education directed at children and adolescents as well as through more general efforts directed at adults. Improved HL could enhance the ability and motivation of individuals to solve personal and public health problems by enabling them to apply skills in response to various health problems arising throughout life. This process of empowerment constitutes one of the major goals of health communication.

Table 2 Item content and means of the Communicative and Critical HL scales

	Mean	SD
1) Seeking information from various sources	4.13	0.80
2) Extracting relevant information	3.92	0.82
3) Understanding and communicating the information	3.56	0.85
4) Considering the credibility of the information	3.52	0.89
5) Making decisions based on the information	3.42	0.95
Total scale score (Mean, SD)	3.72	0.68

Note: The theoretical range: 1-5.

Acknowledgements
This study was supported by a Grant-in-Aid for Young Scientists (B) from Japanese Ministry of Education, Culture, Sports, Science and Technology.

Authors' contributions
HI wrote the first draft of the paper, and TK revised it critically for important intellectual content. Both authors hold final responsibility for the decision to submit the paper for publication.

Competing interests
The authors declare that they have no competing interests.

Received: 15 October 2010 Accepted: 5 November 2010
Published: 5 November 2010

References

1. Nutbeam D: **Health promotion glossary.** *Health Promot Int* 1998, 13:349-364.
2. Kickbusch I, Maag D: **Health literacy.** In *Encyclopedia of public health. Volume 3.* Edited by: Heggenhougen K, Quah S. San Diego: Academic Press; 2008:204-211.
3. Hesse BW, Nelson DE, Kreps GL, Croyle RT, Arora NK, Rimer BK, Viswanath K: **Trust and sources of health information: the impact of the Internet and**

its implications for health care providers: findings from the first Health Information National Trends Survey. *Arch Intern Med* 2005, 165:2618-2624.
4. Napoli PM: **Consumer use of medical information from electronic and paper media.** In *The internet and health communication: Experiences and expectations.* Edited by: Rice RE, Katz JE. Thousand Oaks, CA: Sage Publications, Inc; 2001:79-98.
5. Passalacqua R, Caminiti C, Salvagni S, Barni S, Beretta GD, Carlini P, Contu A, Di Costanzo F, Toscano L, Campione F: **Effects of media information on cancer patients' opinions, feelings, decision-making process and physician-patient communication.** *Cancer* 2004, **100**:1077-1084.
6. Rutten LJ, Arora NK, Bakos AD, Aziz N, Rowland J: **Information needs and sources of information among cancer patients: a systematic review of research (1980-2003).** *Patient Educ Couns* 2005, **57**:250-261.
7. Davis TC, Long SW, Jackson RH, Mayeaux EJ, George RB, Murphy PW, Crouch MA: **Rapid estimate of adult literacy in medicine: a shortened screening instrument.** *Fam Med* 1993, **25**:391-395.
8. Parker RM, Baker DW, Williams MV, Nurss JR: **The test of functional health literacy in adults: a new instrument for measuring patients' literacy skills.** *J Gen Intern Med* 1995, **10**:537-541.
9. Nielsen-Bohlman L, Panzer AM, Kindig DA: **Health literacy: A prescription to end confusion.** Washington, DC: The National Academies Press; 2004.
10. U.S Department of Health and Human Services: **Healthy People 2010: Understanding and Improving Health.** Washington, DC: Government Printing Office;, 2 2000.
11. Nutbeam D: **Health literacy as a public health goal: a challenge for contemporary health education and communication strategies into the 21st century.** *Health Promot Int* 2000, **15**:259-267.
12. Nutbeam D: **The evolving concept of health literacy.** *Soc Sci Med* 2008, **67**:2072-2078.
13. Peerson A, Saunders M: **Health literacy revisited: what do we mean and why does it matter?** *Health Promot Int* 2009, **24**:285-296.
14. Jordan JE, Osborne RH, Buchbinder R: *J Clin Epidemiol* 2010.
15. Weiss BD, Mays MZ, Martz W, Castro KM, DeWalt DA, Pignone MP, Mockbee J, Hale FA: **Quick assessment of literacy in primary care: the newest vital sign.** *Ann Fam Med* 2005, **3**:514-522.
16. Ishikawa H, Takeuchi T, Yano E: **Measuring functional, communicative, and critical health literacy among diabetic patients.** *Diabetes Care* 2008, **31**:874-879.
17. Chew LD, Bradley KA, Boyko EJ: **Brief questions to identify patients with inadequate health literacy.** *Fam Med* 2004, **36**:588-594.
18. Kutner M, Greenberg E, Jin Y, Paulsen C: **The Health Literacy of America's Adults: Results From the 2003 National Assessment of Adult Literacy (NCES 2006-483).** Washington, DC: National Center for Education Statistics; 2006.
19. Rudd RE: **Health literacy skills of U.S. adults.** *Am J Health Behav* 2007, **31(Suppl 1)**:S8-18.
20. Ishikawa H, Yano E: **Patient health literacy and participation in the health-care process.** *Health Expect* 2008, **11**:113-122.
21. Ishikawa H, Nomura K, Sato M, Yano E: **Developing a measure of communicative and critical health literacy: a pilot study of Japanese office workers.** *Health Promot Int* 2008, **23**:269-274.
22. Roter D, Larson S: **The Roter interaction analysis system (RIAS): utility and flexibility for analysis of medical interactions.** *Patient Educ Couns* 2002, **46**:243-251.
23. Nutbeam D: **Defining and measuring health literacy: what can we learn from literacy studies?** *Int J Public Health* 2009, **54**:303-305.
24. Baker DW: **The meaning and the measure of health literacy.** *J Gen Intern Med* 2006, **21**:878-883.
25. Ishikawa H, Yano E, Fujimori S, Kinoshita M, Yamanouchi T, Yoshikawa M, Yamazaki Y, Teramoto T: **Patient health literacy and patient-physician information exchange during a visit.** *Fam Pract* 2009, **26**:517-523.
26. Rose G: **Sick individuals and sick populations.** *Int J Epidemiol* 1985, **14**:32-38.
27. Gerber BS, Brodsky IG, Lawless KA, Smolin LI, Arozullah AM, Smith EV, Berbaum ML, Heckerling PS, Eiser AR: **Implementation and evaluation of a low-literacy diabetes education computer multimedia application.** *Diabetes Care* 2005, **28**:1574-1580.
28. Kalichman SC, Cherry J, Cain D: **Nurse-delivered antiretroviral treatment adherence intervention for people with low literacy skills and living with HIV/AIDS.** *J Assoc Nurses AIDS Care* 2005, **16**:3-15.
29. van Servellen G, Nyamathi A, Carpio F, Pearce D, Garcia-Teague L, Herrera G, Lombardi E: **Effects of a treatment adherence enhancement program on health literacy, patient-provider relationships, and adherence to HAART among low-income HIV-positive Spanish-speaking Latinos.** *AIDS Patient Care STDS* 2005, **19**:745-759.

doi:10.1186/1751-0759-4-18
Cite this article as: Ishikawa and Kiuchi: **Health literacy and health communication.** *BioPsychoSocial Medicine* 2010 **4**:18.

外科治療　Vol. 102　No. 4（2010：4）

消化器癌治療成績のさらなる向上に向けて

大規模臨床データベースの意義と展望

宮田裕章　後藤満一　岩中督
橋本英樹　香坂俊　本村昇
村上新　木内貴弘　兼松隆之
永井良三　里見進　杉原健一
髙本眞一

永井書店

文　献

1) 後藤満一，北川雄光，木村 理，島田光生，冨田尚裕，中越 享，馬場秀夫，杉原健一，川崎 誠治，平田公一，上西紀夫，北野正剛，大津 洋：日本消化器外科学会 消化器外科データベース委員会2007年度調査報告.
(http://www.jsgs.or.jp/modules/oshirase/index.php?content_id=55)

2) 後藤満一，北川雄光，木村 理，島田光生，冨田尚裕，中越 享，馬場秀夫，杉原健一，大津 洋：日本消化器外科学会 消化器外科データベース委員会2008年度調査報告.
(http://www.jsgs.or.jp/modules/oshirase/index.php?content_id=164)

3) Birkmeyer JD, Siewers AE, Finlayson EV, et al: Hospital volume and surgical mortality in the United States. N Engl J Med 346(15): 1128-1137, 2002.

4) van Heek NT, Kuhlmann KF, Scholten RJ, et al: Hospital Volume and Mortality After Pancreatic Resection. A Systematic Review and an Evaluation of Intervention in The Netherlands. Ann Surg 242(6): 781-790, 2005.

5) Edge SB, Schmieg RE Jr, Rosenlof LK, et al: Pancreas cancer resection outcome in American University centers in 1989-1990. Cancer 71: 3502-3508, 1993.

6) Lieberman MD, Kilburn H, Lindsey M, et al: Relation of perioperative deaths to hospital volume among patients undergoing pancreatic resection for malignancy. Ann Surg 222: 638-645, 1995.

7) Wade TP, Halaby IA, Stapleton DR, et al: Population-based analysis of treatment of pancreatic cancer and Whipple resection: Department of Defense hospitals, 1989-1994. Surgery 120: 680-685, 1996.

8) Glasgow RE, Mulvihill SJ: Hospital volume influences outcome in patients undergoing pancreatic resection for cancer. West J Med 165: 294-300, 1996.

9) Imperato PJ, Nenner RP, Starr HA, et al: The effects of regionalization on clinical outcomes for a high risk surgical procedure: a study of the Whipple procedure in New York State. Am J Med Qual 11: 193-197, 1996.

10) Neoptolemos JP, Russell RC, Bramhall S, et al: Low mortality following resection for pancreatic and periampullary tumours in 1026 patients: UK survey of specialist pancreatic units. UK Pancreatic Cancer Group. Br J Surg 84: 1370-1376, 1997.

11) Birkmeyer JD, Finlayson SG, Tosteson AA, et al: Effect of hospital volume on in-hospital mortality with pancreaticoduodenectomy. Surgery 125: 250-256, 1999.

12) Simunovic M, To T, Theriault M, et al: Relation between hospital surgical volume and outcome for pancreatic resection for neoplasm in a publicly funded health care system. CMAJ 160: 643-648, 1999.

13) Gordon TA, Bowman HM, Bass EB, et al: Complex gastrointestinal surgery: impact of provider experience on clinical and economic outcomes. J Am Coll Surg 189: 46-56, 1999.

14) Gouma DJ, Van Geenen RC, Van Gulik TM, et al: Rates of complications and death after pancreaticoduodenectomy: risk factors and the impact of hospital volume. Ann Surg 232: 786-795, 2000.

15) Nordback L, Parviainen M, Raty S, et al: Resection of the head of the pancreas in Finland: effects of hospital and surgeon on short-term and long-term results. Scand J Gastroenterol 37: 1454-1460, 2002.

16) Finlayson EV, Goodney PP, Birkmeyer JD: Hospital volume and operative mortality in cancer surgery: a national study. Arch Surg 138: 721-725, 2003.

17) Fink AS, Campbell DA Jr, Mentzer RM Jr, et al: The National Surgical Quality Improvement Program in non-veterans administration hospitals: initial demonstration of feasibility. Ann Surg 236(3): 344-353, 2002.

18) Khuri SF, Daley J, Henderson W, et al: Risk Adjustment of the Postoperative Mortality Rate for the Comparative Assessment Of the Quality of Surgical Care: Results of the National Veterans Affairs Surgical Risk Study. J Am Coll Surg 185: 315-327, 1997.

特集 消化器癌治療成績のさらなる向上に向けて

大規模臨床データベースの意義と展望

Quality improvement initiative based on national clinical database

宮田 裕章[*1] MIYATA Hiroaki
後藤 満一[*2] GOTO Mitsukazu
岩中 督[*3] IWANAKA Tadashi
橋本 英樹[*4] HASHIMOTO Hideki
香坂 俊[*5] KOUSAKA Shun
本村 昇[*6] MOTOMURA Noboru
村上 新[*7] MURAKAMI Arata
木内 貴弘[*8] KIUCHI Takahiro
兼松 隆之[*9] KANEMATSU Takashi
永井 良三[*10] NAGAI Ryozo
里見 進[*11] SATOMI Susumu
杉原 健一[*12] SUGIHARA Kenichi
髙本 眞一[*13] TAKAMOTO Shinichi

臨床データベースは臨床現場が主体となって取り組み，発展している活動である．さまざまな影響を与える事業として，活動の社会的な位置づけを検討することは有用である．集積したデータに基づいた課題の同定・改善を通して，臨床現場が医療の質向上を牽引し，患者により良い医療を提供することは，活動の中心的課題である．データベースを活用した臨床研究や根拠に基づく政策提言もまた，企業や行政，保険者等に影響を与える重要な側面である．

はじめに

臨床データベース(clinical database)はより良い医療を長期的に提供することができる体制を構築するため，臨床現場との連携により体系的なデータ収集と実証的な分析を行う基盤となる，事業である．全国から集積したデータに基づいて課題を同定し，改善に取り組むことにより，専門集団は各分野のプロフェッショナリズムを社会に対して示すことできる．

一方，臨床データベースを活用した研究も，近年多くの学術専門誌に掲載されるようになっており，また根拠に基づいた医療政策を支える基盤にもなっている．このように，臨床データベースの意義が高まっている中で，データベースの評価基準を論じた文献はほとんどみられない．

本稿では社会的文脈における意義について概観するとともに，有用性基準に基づいて，さまざまな立場からみた価値を検証する．

所属は本文末に記載
Key words：臨床データベース／医療の質／医療政策／医療評価／臨床研究

0433-2644/10/¥50/頁/JCOPY

I. 医療の質向上に向けた臨床現場主体の事業

Institute of Medicine が21世紀の医療改革にむけて，「患者のための医療」という概念を主軸の1つとして提示したように[1]，今後の医療においては患者の価値を中心に考えることが重要となる．Society of Thoracic Surgeons は「教育，研究，社会発信を通じて心臓血管外科医の能力を高め，彼らが最高の質の医療を提供できるようにすること」を学会の使命として掲げている．同様に American College of Cardiology では「医療政策を提言し，教育，研究の促進とガイドラインの設定と実施を通して，心疾患医療の質を向上させる」という目的を設定している．また American Cancer Society は「研究，教育，支援活動やサービスの提供を通して，がんを予防し，命を救い，がんによる苦しみを撲滅する」という形で，研究という一側面だけでなく，患者の価値を中心に据えた活動として専門集団としてのプロフェッショナリズムを規定している．

一方で医療をとりまく政策課題において，しばしば医療費の抑制が中心的な課題とされることも多い．しかしながら医療の主たる目的は患者に最善のサービスを提供することであり，医療費を削減することではない[2]．当然ながら同等に質の高い医療を実現できる2つの方法がある場合よりコストが少ない方が望ましい．ただ，患者に提供するサービスの質の把握した上で，一定の質の提供するためにどのようなコストが必要か，という順序で医療を考えることは有用である．したがって，医療においては患者に質の高いサービスを提供することを第1の目的として設定し，その目的のため診療報酬をはじめとした制度や医療提供システム，実践的取り組みをどのように設計・調整するべきかを検討することが重要となると考えられる．

医療の質向上を考える上では，患者の価値を実現する「品質」を定義・把握し，評価することが必須事項である．この医療の品質を示す指標としては，個々の患者のリスクを調整した治療成績を用いることが重要である[3)-5)]．図1に示したように，米国外科学会，National Surgical Quality Improvement Program においても，リスク調整により治療成績に大きな変化のある施設が少なからずある．一方で日本においては，ほとんどの領域においてリスク調整の議論が行われておらず，手術死亡率をはじめとした施設の治療成績が

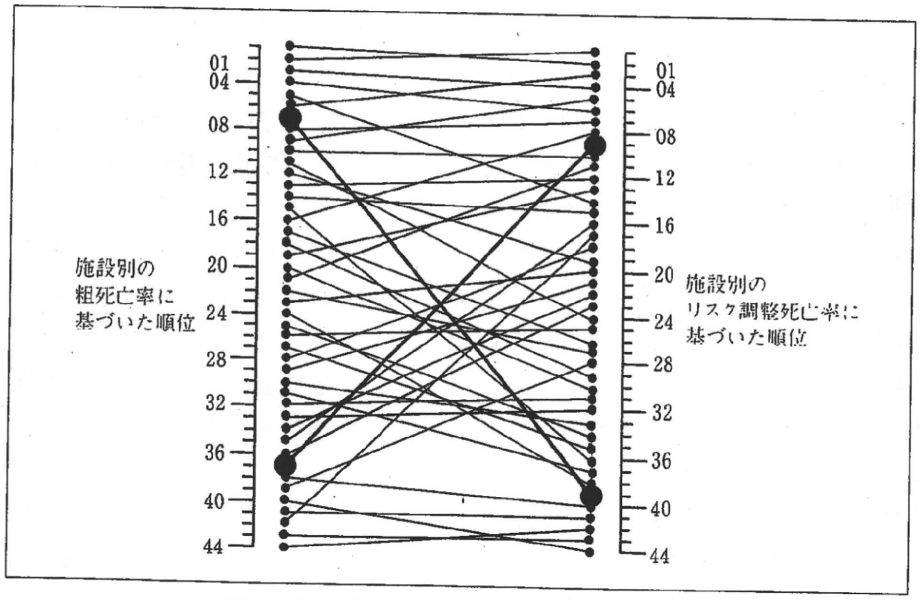


図1 リスク調整による施設別治療成績順位の変化

手術集団の特徴の違いによって左右されることが無視されたものとなっているのが現状である．

このように適切な指標が確立しない状態で情報公開だけが先行した場合には，医療提供者側がリスクの低い患者を回避し，重篤な患者が医療を受ける機会が損なわれてしまうことが，海外の事例からも指摘されている[6)-8)]．情報公開は，医療における透明性を確保し，質の向上を牽引する手段の一つではあるが，それ自体は目的ではない．したがって情報公開の前提として，臨床現場が理解・納得できる正しい情報をフィードバックし，医療の質向上にむけて活用することができるような体制を構築することが必要である[4)]．

このような観点から患者のためのより良い医療を長期的に提供することができる体制を構築するため，臨床現場との連携により体系的なデータ収集と実証的な分析を行う基盤となるのが，各領域の臨床データベース(clinical database)である．

II. 臨床データベースの意義・課題を考えるための視点

本稿では臨床データベースの事業としての社会的意義と課題を有用性の観点から検討する．これは評価において用いられる有用性基準(Utility standard)，実現可能性基準(Feasibility standard)，正当性基準(Propriety standard)，正確性基準(Accuracy standard)のうちの1つである[9)]．有用性基準は，事業が影響を及ぼす関係者の価値を正確に把握し，ニーズを確定し，その必要性に役立つサービスを行っているかどうかを検討するものである．

今回の検討では有用性基準を，①中心的課題の明確化，②関係者の価値の把握，③プロセスと成果の把握，④さまざまな影響に対する配慮，という下位区分で検討した．

1．中心的課題の明確化

臨床データベースでは，先にあげたように患者の価値を主軸に医療の質向上を牽引することが中心的課題となる．しかしながら医師をはじめとした臨床スタッフが疲弊してしまっては，質の高い医療を提供する上で継続的な供給は難しい．したがって医療提供者が充実した環境で高い質のケアを提供できる環境を整備することや，質の高い医療を提供する医療提供者や施設がむくわれるような支援を提言することは，臨床データベースの重要な目的の一つである．

一方でいくら高い質の医療を提供するためとはいえ，医療機関や保険者に非現実的な財政負担が生じることも避けるべきである．良質な医療を継続的に提供するための現実的な制度・体制の整備上でも，臨床データベースは大きな役を果たすと考えられる．臨床データベースに基づいた課題を同定し，改善に取り組む，臨床現場の取り組みに対して診療報酬加算を設定し，全体の医療の質向上を通して保険者の負担を軽減するという“pay for participation”という政策は米国で行われている取り組みの一つである[4)]．また近年は治療成績の良好な施設に対して診療報酬加算を設定し，医療の質向上の動機づけを高めるという pay for performance も海外では保険者が取り入れるようになってきている[10)11)]．

2．関係者の価値の把握

1）患者・一般住民

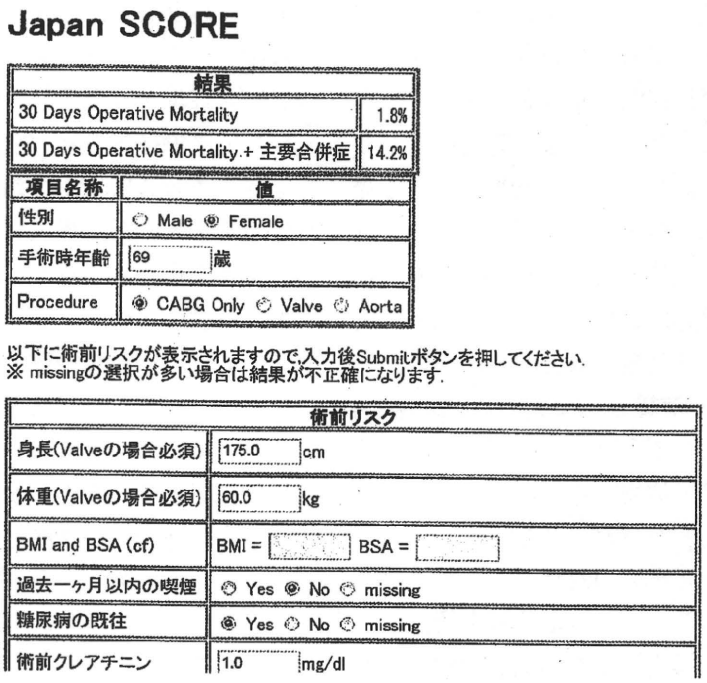
患者および一般住民の利益は，臨床データベースに基づいた改善の取り組みを通じて，全体としての医療の質が底上げされ，より良質な医療の提供を受けることである．一方で，各施設や専門医について公開された情報を基に，自分自身が納得できる施設選択を行うことも，患者側のメリットとしてあげることができる．公開される情報形式としては，

①施設や専門医の認定の有無とその根拠

②医療の質に関わる施設条件(人員配置や症例数)や臨床プロセス(臨床指標の施設別の遵守率)

③重症度補正した施設別の治療成績

などさまざまである．また同一の患者を正確に同定することができるような情報を臨床データベ



Japan SCORE

結果	
30 Days Operative Mortality	1.8%
30 Days Operative Mortality.+ 主要合併症	14.2%

項目名称	値
性別	○ Male ◉ Female
手術時年齢	69 歳
Procedure	◉ CABG Only ○ Valve ○ Aorta

以下に術前リスクが表示されますので,入力後Submitボタンを押してください.
※ missingの選択が多い場合は結果が不正確になります.

術前リスク	
身長(Valveの場合必須)	175.0 cm
体重(Valveの場合必須)	60.0 kg
BMI and BSA (cf)	BMI = BSA =
過去一ヶ月以内の喫煙	○ Yes ◉ No ○ missing
糖尿病の既往	◉ Yes ○ No ○ missing
術前クレアチニン	1.0 mg/dl

図2　心臓外科領域における術前リスク予測機能

ースが保持することができれば，異なる施設に受診した場合や，退院後長期間経過していた場合でも患者情報を引き出し，診療により有意義に活用可能することもできる．

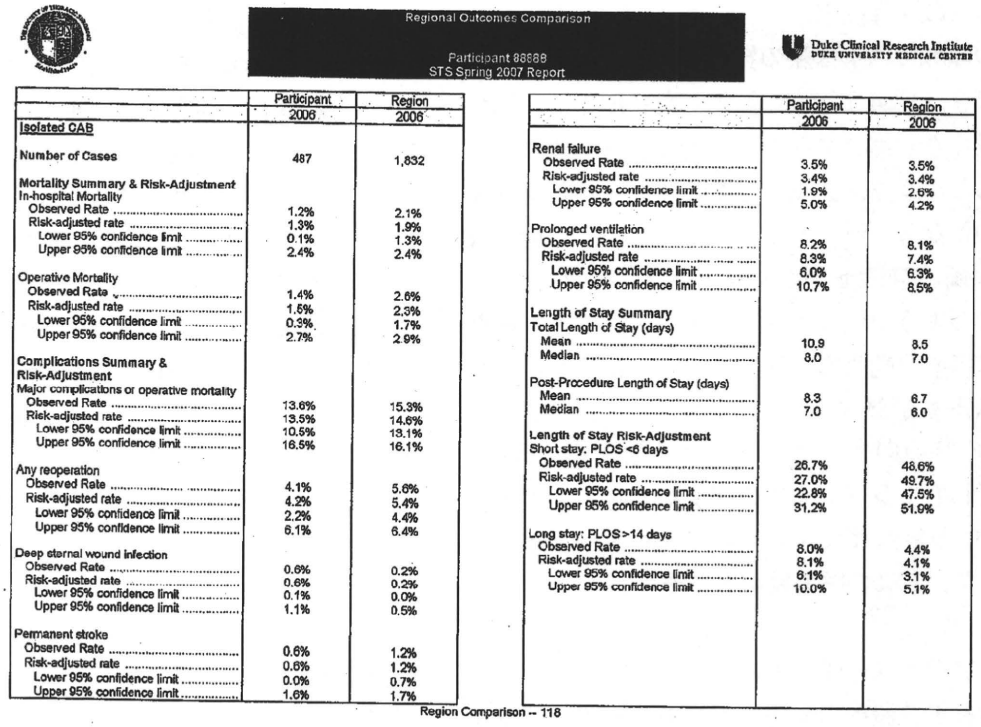
2）医療提供者

臨床現場の医療提供者は全国で統一された基準でデータを入力・管理することにより自施設での取り組みを，全国の状況と対比して把握することができる(図2)．全国のデータに基づいた重症度分析に基づいて，目の前の患者がどのようなリスクを有しているのかを事前に同定フィードバックすることができる．これにより医療提供者は，より客観的な情報に基づいて治療適応の判断やインフォームドコンセントを行うことができる．また標準化された情報を症例レポートとして再出力し，カンファレンスでの情報共有にも活用することも可能である．個々の施設で入力されたデータは，専門医をはじめとした各種臨床学会の資格申請に活用することができ，各スタッフの事務手続きの負荷を軽減することが可能である．

一方で自施設のデータを活用し，追加の項目を加えることにより発展的な臨床研究を実施することもできる．

3）参加施設

参加施設には定期的に全国データと対比した形で，重症度補正を行った治療成績を含んだ施設レポートが配布される(図3)．この施設レポートに基づいて，参加施設は自施設の特徴と課題を把握することができる．また施設は自施設の位置づけを参考に，施設としての戦略やスタッフのマネジメントを行うことも可能となる．また臨床データベースを活用したベンチマーキング事業に参加していること自体が，施設としての一定の質を保証する[12)13)]．したがってデータベース事業への参加を施設の広報に活用する(例：米国の循環器内科のベストホスピタルのうち95%が American college of cardiology の臨床データベース参加施設)，学会からの施設認定を受ける，データベース参加により診療報酬加算を受けるなど施設の活動を後押しするものとして位置づけることも可能



Regional Outcomes Comparison

Participant 88888
STS Spring 2007 Report

Duke Clinical Research Institute
DUKE UNIVERSITY MEDICAL CENTER

	Participant 2006	Region 2006
Isolated CAB		
Number of Cases	487	1,832
Mortality Summary & Risk-Adjustment		
In-hospital Mortality		
Observed Rate	1.2%	2.1%
Risk-adjusted rate	1.3%	1.9%
Lower 95% confidence limit	0.1%	1.3%
Upper 95% confidence limit	2.4%	2.4%
Operative Mortality		
Observed Rate	1.4%	2.6%
Risk-adjusted rate	1.5%	2.3%
Lower 95% confidence limit	0.3%	1.7%
Upper 95% confidence limit	2.7%	2.9%
Complications Summary & Risk-Adjustment		
Major complications or operative mortality		
Observed Rate	13.6%	15.3%
Risk-adjusted rate	13.5%	14.6%
Lower 95% confidence limit	10.5%	13.1%
Upper 95% confidence limit	16.5%	16.1%
Any reoperation		
Observed Rate	4.1%	5.6%
Risk-adjusted rate	4.2%	5.4%
Lower 95% confidence limit	2.2%	4.4%
Upper 95% confidence limit	6.1%	6.4%
Deep sternal wound infection		
Observed Rate	0.6%	0.2%
Risk-adjusted rate	0.6%	0.2%
Lower 95% confidence limit	0.1%	0.0%
Upper 95% confidence limit	1.1%	0.5%
Permanent stroke		
Observed Rate	0.6%	1.2%
Risk-adjusted rate	0.6%	1.2%
Lower 95% confidence limit	0.0%	0.7%
Upper 95% confidence limit	1.6%	1.7%

	Participant 2006	Region 2006
Renal failure		
Observed Rate	3.5%	3.5%
Risk-adjusted rate	3.4%	3.4%
Lower 95% confidence limit	1.9%	2.6%
Upper 95% confidence limit	5.0%	4.2%
Prolonged ventilation		
Observed Rate	8.2%	8.1%
Risk-adjusted rate	8.3%	7.4%
Lower 95% confidence limit	6.0%	6.3%
Upper 95% confidence limit	10.7%	8.5%
Length of Stay Summary		
Total Length of Stay (days)		
Mean	10.9	8.5
Median	8.0	7.0
Post-Procedure Length of Stay (days)		
Mean	8.3	6.7
Median	7.0	6.0
Length of Stay Risk-Adjustment		
Short stay: PLOS <6 days		
Observed Rate	26.7%	48.6%
Risk-adjusted rate	27.0%	49.7%
Lower 95% confidence limit	22.8%	47.5%
Upper 95% confidence limit	31.2%	51.9%
Long stay: PLOS >14 days		
Observed Rate	8.0%	4.4%
Risk-adjusted rate	8.1%	4.1%
Lower 95% confidence limit	6.1%	3.1%
Upper 95% confidence limit	10.0%	5.1%

Region Comparison -- 118

図3 ベンチマークリポートのサンプル(術後アウトカム)

である.

4）臨床学会

これまで多くの領域の学会は，各領域での臨床的取り組みに対する体系的なデータを有していなかった.臨床データベースを整備することにより，臨床学会は統一された基準と定義に基づいて，各領域を取り巻く現実の状況を把握することができる．またこのような基準の統一は，収集するデータの再現性を高めるため，情報の科学的精度を高める影響がある．また豊富なサンプルサイズにより，さまざまな研究デザインが可能となる可能性がある．一方で，各種治療の実施状況や効果について正確な情報を得ることで，臨床学会は根拠に基づいた専門医認定や専門医の適正配置，労働環境の改善や診療報酬の設定に向けた政策提言を行うことが可能となる．臨床学会は医療の質向上の牽引により，専門家集団として社会に対する説明責任を果たすとともに，専門医の意義や認定施設の有用性を，社会に対しても広くアピールすることが可能になる.

5）製薬・医療機器関連企業

臨床学会と共同で臨床研究を行うことにより，製薬・医療機器関連企業は，医薬品・医療機器の治験や市販後調査を迅速に行うことが可能となる．また全数調査が原則となる臨床データベースの上で治験を行うことで，サンプリングのコストを削減できるだけなく，登録対象外の患者情報が把握可能となるため，学術的により質の高い検証を行うことが可能となる．一方で医療機器・薬剤の使用状況や効果に対する市販後調査はより効果的な開発や販売促進を行う上でも有益な情報となりうる．ただ日本においては市販後調査や利用状況調査に対して，企業に対する報告義務が米国に比して低い基準であるため，臨床データベースに対する関連企業の認識はそれほど高いとはいえないのが現状である．また治験においても，倫理的にランダム化困難な場合に，臨床データベースに登録された症例を比較対照群として設定して，効

果を検討することが可能となると考えられる．

6）行政・保険者

その領域で何が医療の質を示す指標か把握されないまま，低質な治療が蔓延した状態では，死亡や合併症など高コストを伴う術後有害事象が頻発することにより，医療としてのコストが非常に高いものとなる．一方で臨床データベースは情報収集・評価のコストが新たに発生するものであるが，このような情報のフィードバックにより死亡率や有害事象発生率の減少により，結果として医療コストも削減することが示唆されている[14)15)]．“医療の質の向上”と“医療費の効率的な運用”は必ずしもトレードオフの関係にあるわけではなく，1％の評価コストで医療の質向上を促進させることで，10%の医療費増も可能である．

3．プロセスと成果の報告

1）参加施設への報告，ベンチマーキングレポート

データベース事業の参加施設に対しては，全国のデータと対比した形で，各施設の重症度補正治療成績や，患者の特徴が把握可能なレポートが定期的に配布される．紙ベースのレポートは1年や半年に一度の定期的なものとなるが，近年はWebを通じたフィードバックにより施設の変化や治療成績の推移が即時的に把握できるような形で，フィードバックを行うことができるようになった．

2）学術集会やシンポジウムを通した全参加施設での進捗状況の確認

データを活用した分析結果の報告，データベース事業の運営の状況，専門家集団としての政策対応，入力項目やインターフェイスの改善，各施設の取り組みの支援などさまざまな観点について，運営主体である臨床学会と各参加施設が情報を共有しコミュニケーションを行うことは，発展的な運営に不可欠である．WebやE-mailを利用した情報共有以外にも，シンポジウムや学術集会など定期的な会合で情報や意識を共有することは有用であると考えられる．また各地域や課題別にグループを形成することにより，より活動性の高い活動を構成することも有用である．ACS NSQIPでは政策対策部会以外に，各病院のベストプラクティスを紹介し，共有するグループもある．

3）行政や患者側に対する成果の報告

行政からの金銭的支援を受けた場合には，その結果を報告書として作成することは不可欠である．一方で良質な医療を提供する上で，制度的支援が不可欠である場合には，分析結果を活用し，適宜ロビーイング活動を通して行政・立法府に支援を呼びかけることも有用である．米国胸部外科学会では，毎年医療政策フォーラムを開催するとともに，臨床データベースと連動した形で根拠を検証し，政府に要望書を提出している．

4．さまざまな影響に対する配慮

1）ベンチマーキングを通した成果の確認

全体の治療成績や臨床プロセスの経時的な推移を把握し，事業としてベンチマーキングの重点を適切にデザインすることは重要である．たとえば新たな治療法や治療手段が普及した場合には，その状況を把握するために項目を改善することは必要である．また治療成績についても，周術期死亡の施設間格差が少なくなった場合には，格差が大きな他の合併症に焦点を置き，領域として取り組みを行うことが必要とされる．また負の側面となる影響も考慮しなければならない．とくに治療成績に対して情報公開や診療加算が設定される場合には，重症患者の回避や，早期退院・転科による患者選択などの影響が指摘されており[7)8)]，影響の継続的な検証が必要とされる．

2）データベースに基づいて施行した政策の影響

臨床学会が政策や制度に対して提言を行った場合には，その帰結についても把握し，効果を検証する必要がある．心臓外科領域では施設認定によって生じる患者の移動・それに伴う治療成績への影響を事前・事後に検証を行った事例がある[16)17)]．また一方で，治療成績に対する診療報酬加算や情報の公開を行う場合にも，その効果を把握し，政策の意義を含め，今後に向けたより良い実施方法

(あるいは打ち切り)を検討することが必要とされる[18].

結　論

臨床データベースは，患者のためのより良い医療を長期的に提供することができる体制を構築するため，臨床現場が主体となり体系的なデータ収集と実証的な分析を行う事業である．集積したデータに基づいて課題を同定し，改善に取り組むことにより，臨床学会や現場スタッフは医療の質向上を牽引し，患者や国民により良い医療を提供することが可能となる．加えて臨床データベースを活用した，迅速かつコストパフォーマンスに優れた臨床研究の実施は，医療関連企業にとっても有益になる．

一方で行政や保険者と連携して，臨床データベースを基盤にした政策提言を行っていくことも重要である．しかしながら医療政策の影響は必ずしも望ましいものだけではないため，効果の検証や根拠の確認を行うことも必要である．

*1東京大学大学院医学系研究科医療品質評価学講座 准教授
*2福島県立医科大学医学部臓器再生外科学講座 教授
*3東京大学医学部小児外科 教授
*4東京大学大学院医学系研究科臨床疫学・経済学 教授
*5慶應義塾大学医学部循環器内科
*6東京大学医学部心臓外科 講師
*7東京大学医学部心臓外科 准教授
*8東京大学大学院医学系研究科医療コミュニケーション学 教授
*9長崎大学大学院医歯薬学総合研究科移植・外科学分野 教授
*10東京大学医学部循環器内科 教授
*11東北大学医学部移植・再建・内視鏡外科 教授
*12東京医科歯科大学大学院医歯学総合研究科腫瘍外科学 教授
*13三井記念病院 病院長

文　献

1) Institute of Medicine. Crossing the Quality Chasm : A New Health System for the 21st Century. National Academy Press, 2001.
2) Porter ME, Teisberg EO : How physicians can change the future of health care. JAMA 297 : 1003-1111, 2007.
3) Khuri SF, Daley J, Henderson W, Barbour GJ, Lowry P, Irvin G, Gibbs J, Grover F, Hammermeister K, Stremple JF, Aust JB, Demakis J, Deykin D, McDonald G and Participants in the National Veterans Administration Surgical Risk Study : The National Veterans Administration Surgical Risk Study : risk adjustment for the comparative assessment of the quality of surgical care. Journal of the American College of Surgeons 180 : 519-531, 1995.
4) Birkmeyer NJO, Birkmeyer JD : Strategies for improving surgical quality-Should payers reward excellence or effort? New England Journal of Medicine 354 (8) : 864-870, 2006.
5) Shahian DM, Blackstone EH, Edwards FH, Grover FL, Grunkemeier GL, Naftel DC, Nashef SAM, Nugent WC, Peterson ED : Cardiac surgery risk models : A position Article. Annals of thoracic surgery 78 : 1868-1877, 2004.
6) Shahian DM, Normand SL, Torchiana DF, et al : Cardiac surgery report cards : comprehensive review and statistical critique. Annals of Thoracic Surgery 72 : 2155-2168, 2001.
7) Landon BE, Normand SL, Blumenthal D, Daley J : Physician clinical performance assessment : prospects and barriers. JAMA 290 : 1183-1189, 2003.
8) Schneider EC, Spstein AM : Influence of cardiac-surgery performance reports on referral practices and access to care-a survey of cardiovascular specialists. New England Journal of Medicine 335 : 251-256, 1996.
9) Joint Committee on Educational Evaluation, James R. Sanders (chair) : The program evaluation standards : how to assess evaluations of educational programs. 2nd edition. Sage Publications, Thousand Oaks, CA. 1994.
10) Campbell SM, Reeves D, Kontopantelis E, Sibbald B, Roland M : Effects of pay for performance on the quality of primary care in England. N Engl J Med 361 : 368-378, 2009.
11) Lindenauer PK, Remus D, Roman S, Rothberg MB, Benjamin EM, Ma A, Bratzler DW : Public reporting and pay for performance in hospital quality improvement. N Engl J Med 356 : 486-496, 2007.
12) Jamtvedt G, Young JM, Kristoffersen DT, et al : Audit and feedback : effects on professional practice and health-care outcomes. Cochrane Database Syst Rev 3 : CD00259, 2003.
13) Hall BL, Hamilton BH, Richards K, Bilmoria KY, Cohen ME, Ko CY : Does surgical quality improve in the American college of surgeons national surgical quality improvement program. An evaluation of all participating hospitals. Ann Surg 250 : 363-376, 2009.
14) Berwick DM, James BC, Coye M : The connections between quality measurement and improvement. Med Care 41 (1 suppl) : I30-I38, 2003.
15) Goetzel RZ, Ozminkowski RJ, Villagra VG, Duffy J : Return on investment in disease management : a review. Health Care Financ Rev 26 : 1-19, 2005.
16) Chassin MR : Achieving and sustaining improved quality : lessons from New York State and cardiac