CASEMIX REFINEMENT BY EXTRACTING AND INTEGRATING MULTI-DIMENSIONAL SEVERITY INFORMATION FROM ROUTINE DATA

Yuichi Imanaka
Department of Health Systems Management and Research, Kameda Medical Center, 929 Higashi-cho, Kamogawa City, Chiba 296-8602, JAPAN
Ph: +81.470.922211, Fax: +81.470.930420, E-mail: imanaka@kameda.or.jp

ABSTRACT
The aim of this study is to develop new methodology to refine casemix classification based on a data set routinely available through extracting and integrating multidimensional clinical severity information. The other purpose of this study is to examine if this resource-based, refined classification can be used for clinical performance indicators such as mortality. The data of the consecutive discharged cases were collected in teaching hospitals in Japan. The data set included only variables routinely available, such as patient demographic variables, principal and secondary diagnoses, dates and routes of admission and discharge, and fee information. Coronary-aorta bypass graft cases equivalent to one HCFA DRG were targeted for the study. Factors influential to the health care charges in the targeted group were identified by multiple regression analysis. These were patient age, presence of important secondary diagnosis, and emergency status. To confine the number of subgroups, a point was assigned to each factor so that a point of one factor was equivalent to that of another factor in terms of resource intensity. By use of this point system, one DRG was divided into four subgroups and the variance in charge was effectively reduced. This classification system was named as Resource Intensity Equivalent Group (RIEG). The same classification was applied to another larger database and, among the resultant subgroups, intra-hospital mortality was found to vary from 4% to 20% and the risk-stratified mortality tended to distinguish the hospital difference of performance more clearly.
KEYWORDS: Severity adjustment, Medical expenditure, Clinical performance indicator, Quality of care, Casemix classification

Acknowledgement
I greatly appreciate the participants of the study for their enthusiastic commitment: Teine Keijinkai Hospital (Sapporo), Nikkou Memorial Hospital (Muroran), Takeda General Hospital (Aizu-wakamatsu), Tachikawa General Hospital (Nagaoka), Kameda General Hospital (Kamogawa), Rakuwakai Otowa Hospital (Kyoto), Hosoki Hospital (Kouchi), and Asou Iizuka Hospital (Iizuka).

BACKGROUND
Social demands are being strengthened for measurement of quality and efficiency of health care. On the other hand, the Japanese government has started a feasibility study of some kind of prospective payment system based upon casemix classification since November 1998, and has faced the basic issues of information infrastructure for health care in the country as had been described before[1]. The health care payment system in Japan is basically based upon fee for service, and hospitals and long-term care facilities are still under differentiation. As a result, longer and more variable length of stay and larger variation in charges can be found in hospital care in Japan within case groups defined by classification algorithm such as HCFA-DRG. More homogeneous casemix classification is demanded for useful indicators of efficiency and clinical quality. This means that some kind of risk adjustment and/or stratification methods are needed. There are already various elaborated methods of risk adjustment[2]. However, they usually requires complicated sets of data elements and classification algorithms are not publicly disclosed. When an efficient and concise set of data easily and routinely collectable can produce patient classification of effective risk/severity adjustment/stratification system, the classification system

will have a great value to the society in terms of standardization and diffusion for quality and efficiency measurement in health care.

OBJECTIVES

The aim of this study is to develop new methodology to refine casemix classification based on a data set routinely available. By extracting and integrating multidimensional clinical severity-related information from a routine data set, this method produces more homogenous in health care charge, proxy for resource consumption. The other purpose of this study is to examine if this resource-based, refined classification can be used for stratifying severity or risk and for producing more valid or useful clinical performance indicators.

METHODS

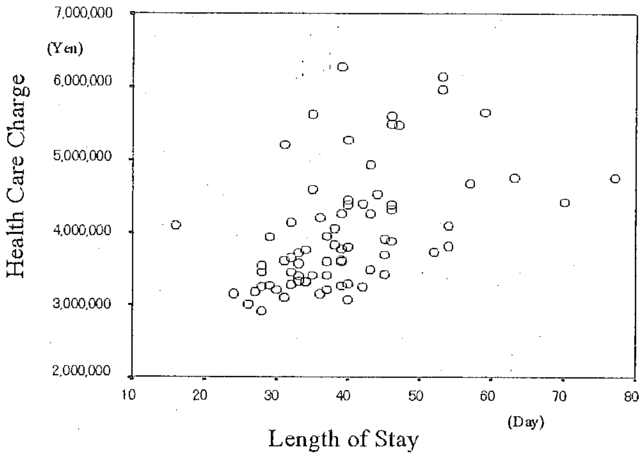
Firstly, the data of the consecutive discharged cases in 1996 were collected in teaching hospitals in Japan. The database of the medical record room and that of the billing system were linked together. The data set included only variables routinely available, such as patient demographic variables, principal and secondary diagnoses in ICD9 and procedures in ICD9CM, dates and routes of admission and discharge, and fee information. CABG (Coronary-aorta bypass graft) cases equivalent to HCFA DRG[3] 106 were targeted for this study. Factors influential to the health care charges in the targeted group were identified by multiple regression analysis and points were assigned to each factor so that the amount of resources were reflected in the number of points. The casemix classification refinement method was developed based upon this point system.

Secondly, the same algorithm for refine classification was applied to the database in the years 1997 to 1998 on the same data set of five hospitals of a voluntary study group. The validity to differentiate health care charges was examined in this database. Next, the validity to use the refinement algorithm as the base for mortality measurement as a clinical performance indicator.

RESULTS

In the CABG cases, equivalence to one HCFA DRG, in the first database, health care charges and length of stay were widely spread with the average length of stay of 39.5 days and the average charge of 4,020,000 yen as shown in figure 1[4]. In this database, four factors presented substantial and statistically significant effects on health care charges in analysis of variance. They were the number of graft, the emergency status (the existence of acute myocardial infarction was used instead, because this variable has stronger explanatory power), the presence of important secondary diagnosis, and the patient age. Conceptually, patient classification could be divided to represent more homogeneous groups by the 4 corresponding dichotomies into 16, 2 by 2 by 2 by 2, subgroups as shown in figure 2[4]. Multiple regression analysis was performed and the three variables of acute myocardial infarction, important secondary diagnosis, and patient age was selected as significant determinants of health care charge and nearly one third of the variance was explained by them. In stead of such multiple divisions, a point was assigned to each factor of which the weight was based upon each corresponding regression coefficient so that a point of one factor was equivalent to that of another factor in terms of resource intensity. By use of this point system, cases in one DRG was divided into four subgroups according to an algorithm shown in figure 3[4], and still the variance in charge was effectively reduced by a third. The median charge of each subgroup was different by about 500,000 yen from its adjacent group as shown in figure 4. This classification system was named as "Resource Intensity Equivalent Groups" (RIEG)[4]. The same classification method was applied to CABG cases in the second database and equivalent results were obtained in terms of health care charges.

Figure 1. Distribution of Length of Stay and Health Care Charges



note) Coronary-Aorta Bypass Graft (CABG) cases (n=73). [c.f., ref.4]

Crude intra-hospital mortality was calculated across five different hospitals in the second database, and the rate differs from 4.3, 5.2, 8.9 and 11.8 to 20.0 percentage (figure 5). There should be substantial variation in the frequency of having very difficult patients for surgical operation among hospitals. As an example, a lot of difficult patients were referred and transferred from other high-level surgery teams to one hospital for surgical operation , and in another hospital, only limited number was transferred from outside experts. Therefore, the crude rate comparison may not be fair, and some kind of risk adjustment method should be used. In this study, the RIEG classification method was used assuming that those who need the more health care resources would have the correspondingly higher risk for mortality. Using the same classification algorithm again, in the RIEG classes of the same database, the mortality differs with 2.0 % in Class 1 (n=150), 5.5% in Class 2 (n=217), 23.5% in Class 3 (n=34), and 66.7% in Class 4 (n=3) (figure 6). To focus on the similar risk group and also to assure the sample size of a group, the cases in Class 3 and 4 were excluded from the analysis in the next step. The resultant, risk-stratified mortality in each corresponding hospital was found to vary from 2.6, 3.1, 5.9 and 13.3 to 15.4 percentage (figure 7), and the ratio of mortality rates of a worse hospital's to a better hospital's tended to be widened.

The study also aimed to support to establish information systems of each participant institution for continuous preparation of the necessary data and routine use of casemix classification-based performance indicators. Considering everyday use of such refined casemix classification in many institutions in the country, necessary is the established information infrastructure including standardized coding systems of diseases and procedures, standardized data sets and algorithms. A framework developed for this is shown in figure 8[1,4].

CONCLUSIONS

By using multi-dimensional severity information of a routinely available data set and converting them into resource intensity equivalent points, casemix classification can be refined and improved to explain more variance in health care charges. This classification method can also differentiate the groups in terms of mortality risk, and has a potential to be used for some efficiency and clinical performance benchmarking. These findings were base upon limited data in Japanese settings, and further studies are necessary to assure the consistency of the developed system. Its application to other patient groups is also desired. Despite of the limits of this study, the key concept of this study to utilize multi-dimensional severity-related information and integrate them into a few number of subgroups may produce valuable outputs with regard to efficient classification for quality and resource-related indicator measurement and to socialization of casemix classification in terms of standardization and diffusion.

BIBLIOGRAPHY

1. Imanaka Y, et al. Development of framework of uniform data set for clinical cases. Proceeding of 17th JCMI, 404-405, Nov., 1997.
2. Gross, PA., Severity of illness and other confounders of quality measurement. pp.101-123, in Wenzel RP (Ed.), Assessing Quality Health Care, Williams and Wilkins, 1992.
3. 3M Health Information Systems. Diagnosis Related Groups Definitions Manual. 3M Health Information Systems, 1996.
4. Imanaka Y, et al. Resource Intensity Equivalent Groups - A new methodology to refine DRG/Casemix Classification Proceeding of 18th JCMI, 484-485, Nov. 1998.

Figure 1

Distribution of Length of Stay and Health Care Charges in CABG Cases

note) Coronary-Aorta Bypass Graft (CABG) cases (n=73).　[c.f., ref.4]

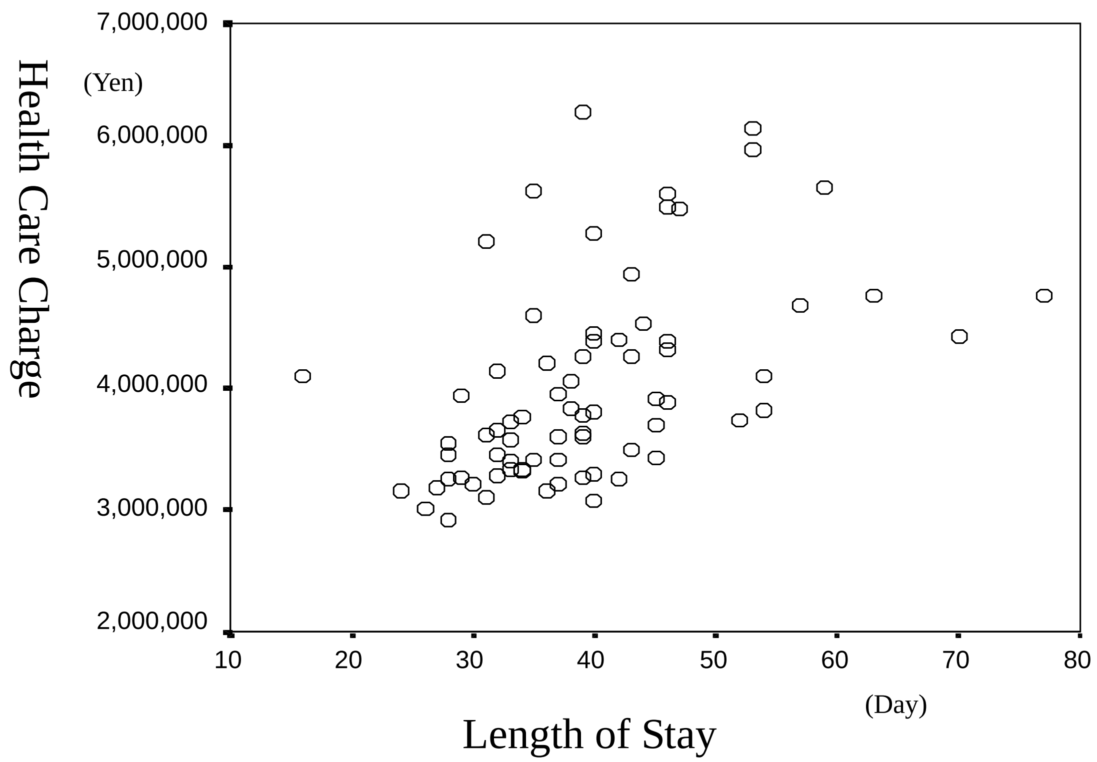


Figure 2

Classification Tree based upon Factors Influential to Health Care Charge

note) The stars (*) indicate the same structure as in the left. AMI stands for acute myocardial infarction. [c.f., ref.4]

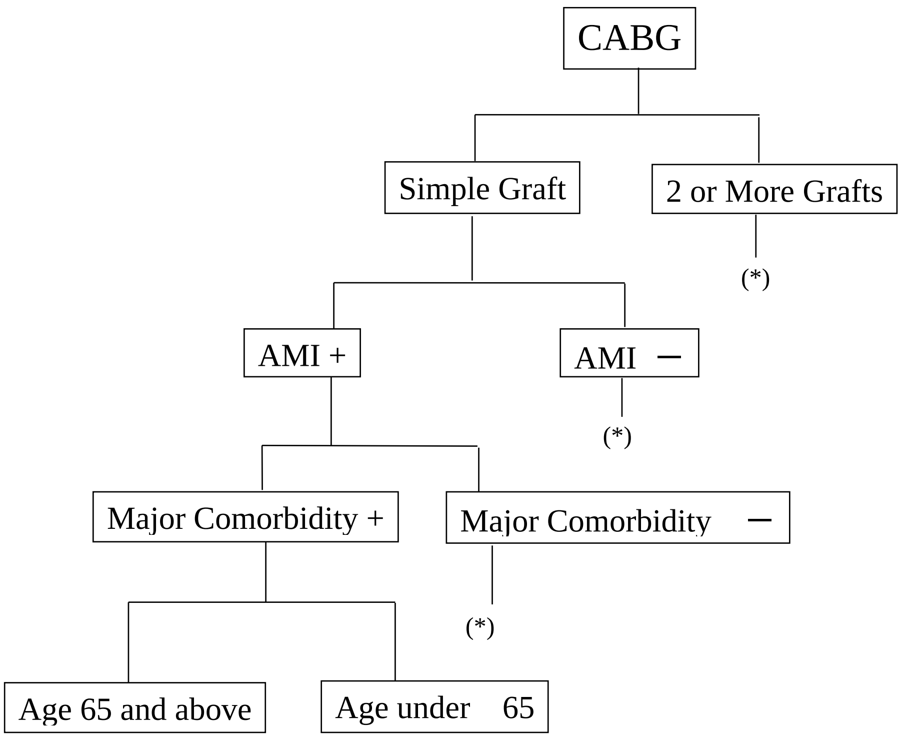
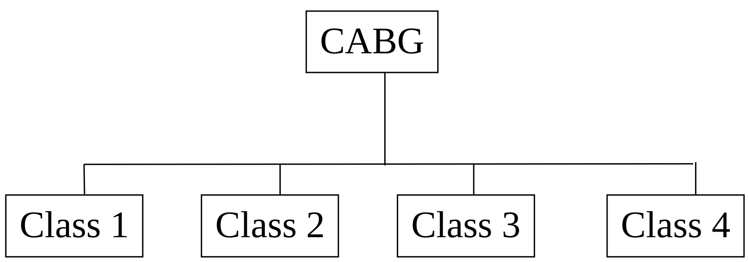


Figure 3

Algorithm for Resource Intensity Equivalent Group (RIEG) in Case of CABG

[c.f., ref.4]



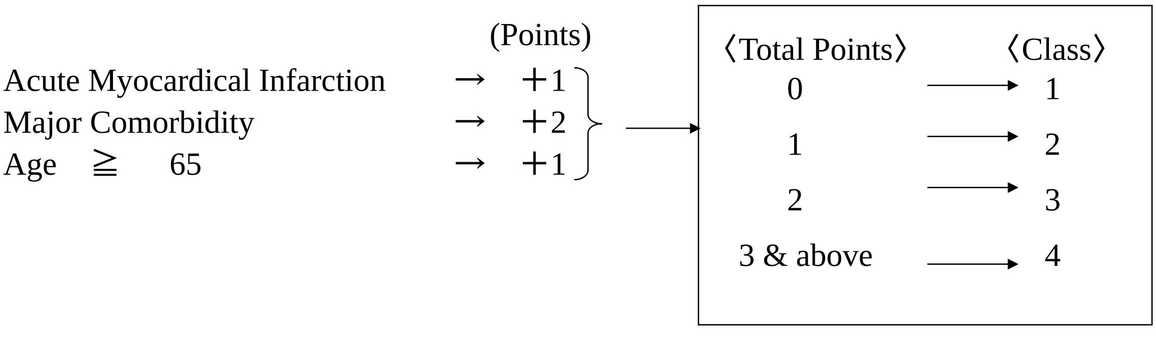


Figure 4

Health Care Charges across Resource Intensity Equivalent Groups (RIEG) in CABG Cases

note) In the box-whisker plot, the bar in the box indicates the median, and the box top and the bottom indicates the 75th and 25th percentiles. Circles are outliers and the whisker indicates the maximum and minimum except the outliers. [c.f., ref.4]

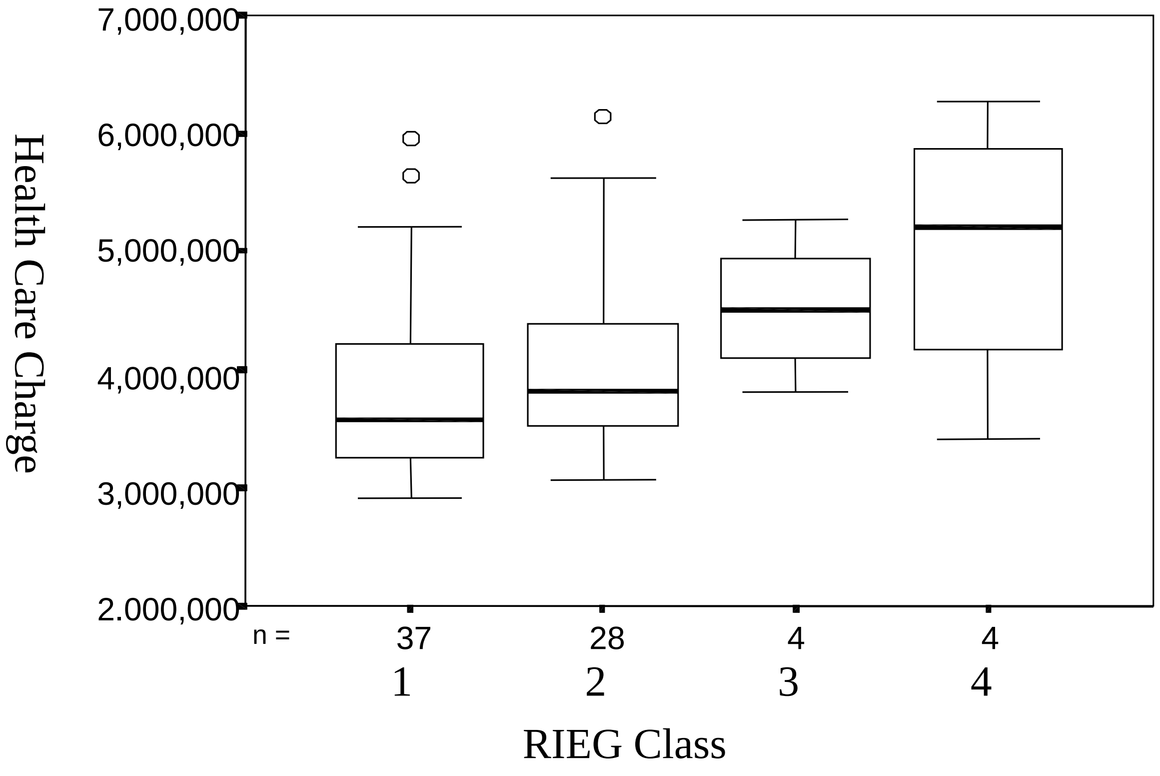


Figure 5

Crude Mortality across Hospitals in CABG Cases

note) The numerator and the denominator in the parenthesis indicate the number of cases dead at discharge and the number of CABG cases respectively.

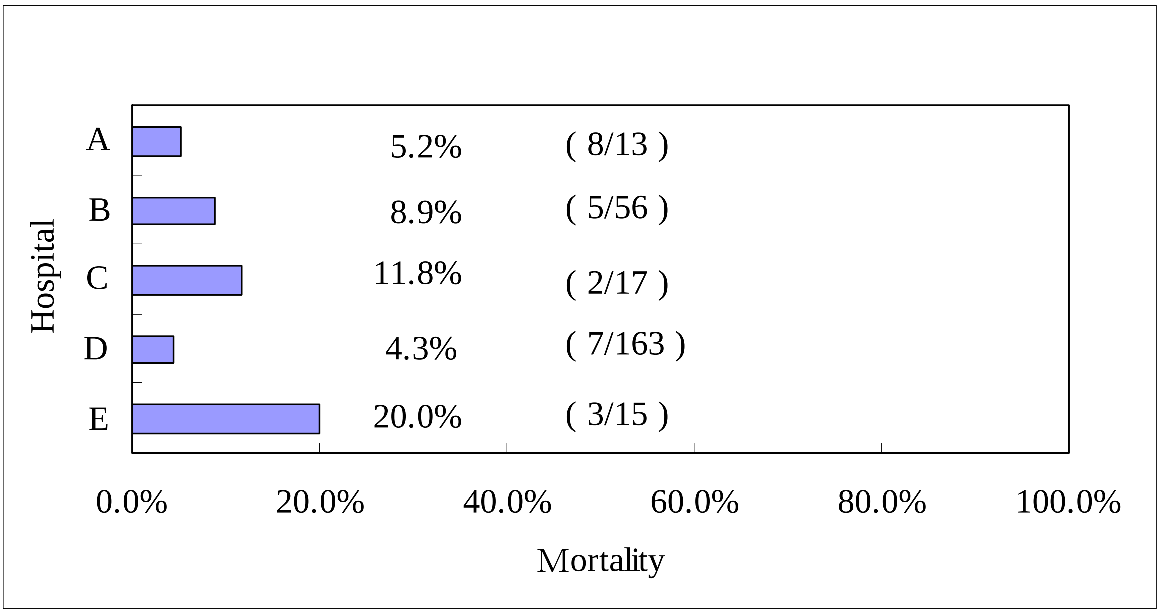


Figure 6

Mortality across RIEG Classes in CABG Cases

note) The numerator and the denominator in the parenthesis indicate the number of cases dead at discharge and the number of CABG cases respectively.

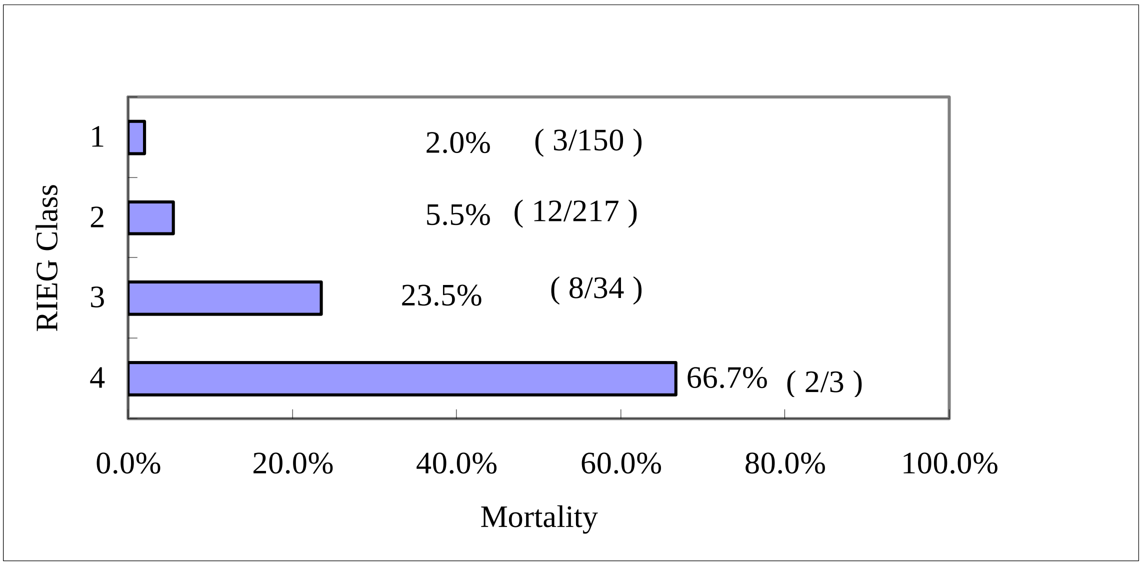


Figure 7
Risk-specific Mortality across Hospitals in CABG Cases

note) The numerator and the denominator in the parenthesis indicate the number of cases dead at discharge and the number of CABG cases respectively.
Only cases in the RIEG Class 1 and 2 are used for this analysis.

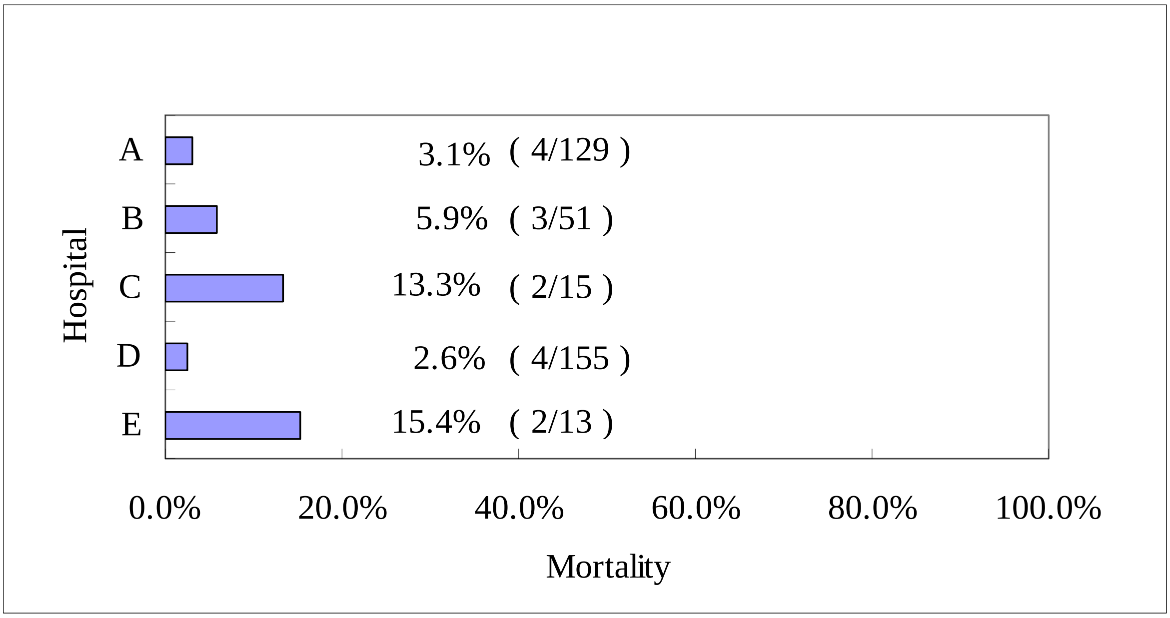
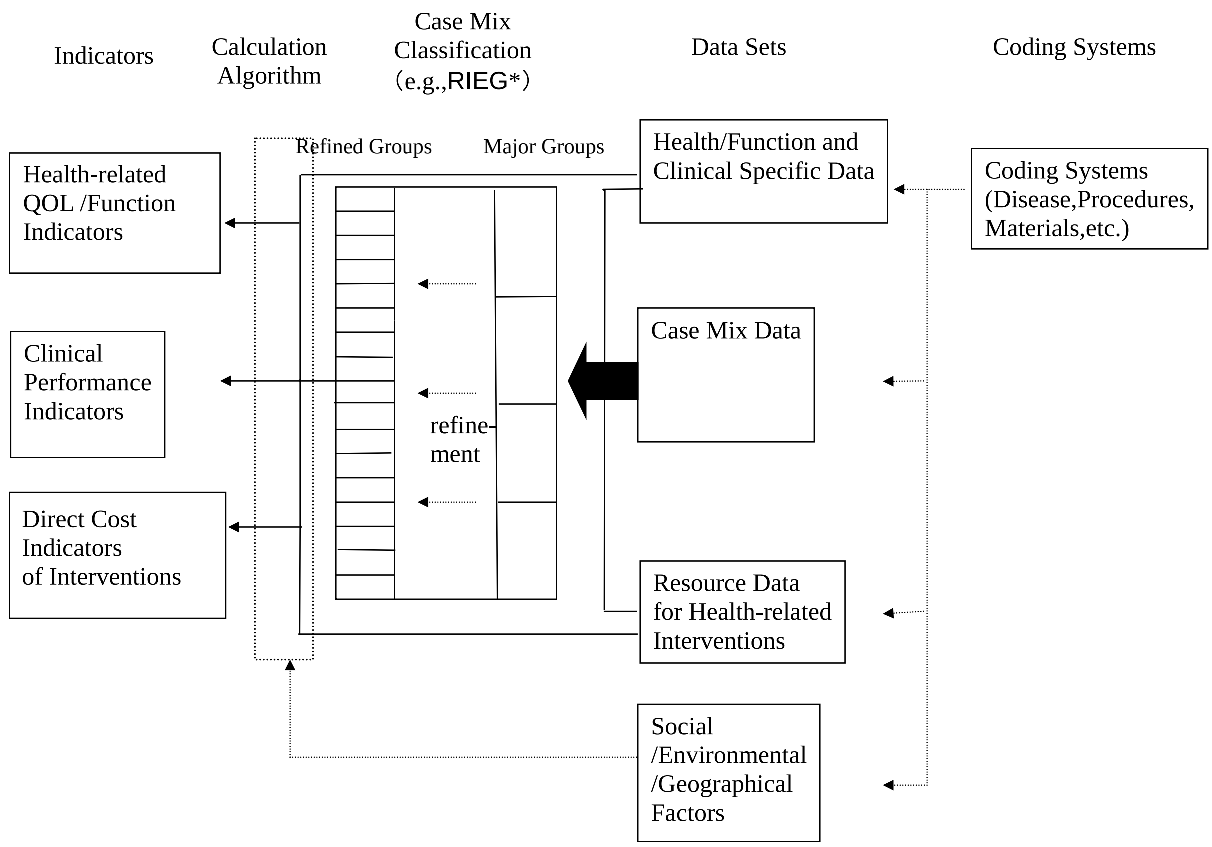


Figure 8

Framework of Information Infrastructure for Casemix Classification and Performance Measures

note) The framework is developed and based upon the Japanese circumstances, but it could be applied to other social settings. [c.f., ref.1, ref.4]



RIEG* :Resource Intensity Equivalent Groups