Status and Future Perspectives of Utilizing Big Data in Neurosurgical and Stroke Research

Ataru NISHIMURA,[1] Kunihiro NISHIMURA,[2] Akiko KADA,[3] Koji IIHARA,[1] and J-ASPECT Study GROUP

[1]Department of Neurosurgery, Graduate School of Medical Sciences, Kyushu University, Fukuoka, Japan;
[2]Center for Cerebral and Cardiovascular Disease Information, National Cerebral and Cardiovascular Center, Osaka, Japan;
[3]Department of Clinical Trials and Research, Clinical Research Center, National Hospital Organization Nagoya Medical Center, Nagoya, Japan

Abstract

The management, analysis, and integration of Big Data have received increasing attention in healthcare research as well as in medical bioinformatics. The J-ASPECT study is the first nationwide survey in Japan on the real-world setting of stroke care using data obtained from the diagnosis procedure combination-based payment system. The J-ASPECT study demonstrated a significant association between comprehensive stroke care (CSC) capacity and the hospital volume of stroke interventions in Japan; further, it showed that CSC capabilities were associated with reduced in-hospital mortality rates. Our study aims to create new evidence and insight from 'real world' neurosurgical practice and stroke care in Japan using Big Data. The final aim of this study is to develop effective methods to bridge the evidence-practice gap in acute stroke healthcare. In this study, the authors describe the status and future perspectives of the development of a new method of stroke registry as a powerful tool for acute stroke care research.

Key words: Big Data, neurosurgical research, stroke research, Diagnosis Procedure Combination

Introduction

The last decade has seen significant advances in the amount of data routinely generated and collected, as well as in our ability to use technology to analyze and understand this data. The intersection of these trends is called 'Big Data' and it is helping businesses in every industry achieve higher efficiency and productivity. Repeated observations over time and space generate most Big Data; examples include worldwide users' Internet search engine queries (e.g., Google), e-commerce browsing and transactions (e.g., Amazon), and genomic sequencing in biomedical research.[1,2]

In the United States, the advanced bioinformatics technologies have led to the use of Big Data in many fields of healthcare. In stroke care, the use of Big Data has received considerable attention as an important source for creating new evidence.[3–9] Although the use of Big Data for healthcare is considered important in Japan, it is yet to be utilized effectively, including for stroke care research.

The nationwide survey of acute stroke care capacity for proper designation of comprehensive stroke centers in Japan (J-ASPECT study) is the first nationwide survey in Japan on the real-world setting of stroke using Big Data obtained from the diagnosis procedure combination (DPC) based payment system.[10,11] This review discusses the creation and analysis of Big Data using analytics and the current uses that are relevant to stroke and the J-ASPECT study that are challenges to using Big Data in stroke care research.

Big Data Analytics in Healthcare

A simple definition of Big Data is based on the concept of datasets whose sizes are beyond the management capabilities of typical relational database software. A more articulated definition of Big Data is based on the three V paradigm: volume, variety, and velocity.[12,13] The volume of the data requires novel storage scalability techniques and distributed approaches for information query and retrieval. The variety of the data source prevents the straightforward use of neat relational structures. Finally, the velocity, which is the increasing rate at

which data is generated, follows a similar pattern as the volume.[14] In the domain of healthcare, Big Data sources and techniques include structured electronic health record (EHR) data, unstructured clinical notes, medical imaging data,[15] genetic data,[16] and the other data (epidemiology and behavioral data).[17]

Despite the inherent complexities of healthcare data, there is potential and benefit in developing and implementing Big Data solutions within this realm. A report by McKinsey Global Institute suggests that if the United States healthcare were to use Big Data creatively and effectively, the sector could create more than $300 billion in value every year.[18] Two-thirds of the value would be in the form of reducing the United States healthcare expenditure.[18] Advances in computer and networking technology, patient monitoring systems, and EHR systems have allowed hospitals to collect and store a rapidly increasing volume and variety of patient data.[19,20] Increasing the recognition of the potential utility of Big Data in health outcomes research has created an impetus to collect and pool EHR data in national datasets. These large datasets provide access to information regarding rare conditions and outcomes that are otherwise difficult to study without robust sample sizes. The goal of Big Data analytics in healthcare is to build evidences and insights based on the 'real world,' and furthermore, to use these evidences to lower costs and improve outcomes through smarter decisions.

Bid Data Analyses about Stroke in the United States

An example of health-care-related Big Data efforts in the United States is the nationwide inpatient sample (NIS), which is structured inpatient EHR data. The NIS is one of the major databases compiled and maintained by the healthcare cost and utilization project (HCUP), which is funded by the agency for healthcare research and quality (AHRQ). Federal and state governments along with medical industry fund AHRQ.[21]

The NIS is the largest all-payer database of inpatient discharge data, and hence, it is a useful dataset for outcomes research. Among hospitals participating in the survey, every discharge for the calendar year is included. The database contains discharge-level data instead of patient-level data, and there is no unique patient identifier to identify re-admissions. Several severity measures are also included in the NIS. Researchers can use these measures for risk-adjustment, or to develop their own risk-adjustment models using the diagnosis and procedure codes included in the data. The NIS does not currently identify conditions present on admission. The age, gender, socioeconomic factor and comorbidities index such as Charlson

comorbidities index is a typical adjustment factor for risk-adjustment. Propensity score matched analyses or mixed model analyses adjusted for hospital level variability of disease severity are typical statistical methods conducted in claims-based analysis.

NIS contains data on a stratified sample of over 1000 US hospitals with approximately 8 million hospital stays per year; weights are available to convert NIS data into national estimates.[21] Further, specialty hospitals (e.g., orthopedics or obstetrics-gynecology hospitals) are included, as are long-term acute care hospitals (since 2005). The data have been collected on an annual basis since 1988, and resources are available to facilitate the evaluation of time trends.

For the effective use of Big Data in stroke care, analyses using NIS have been reported in the United States since 1999; further, the number of articles increased rapidly after 2006. Chronological change in stroke care in the United States can be analyzed using NIS database as it provides not only overall nationwide data, but also continuous data per annum.

The NIS provides important information such as nationwide epidemiological and health economical information. In 1999, Williams et al. reported the estimation of the occurrence, incidence, and characteristics of total (first-ever and recurrent) stroke using the NIS database that is representative of all 1995 US inpatient discharges.[3] There were 682,000 occurrences of stroke with hospitalization and an estimated 68,000 occurrences of stroke without hospitalization. The overall incidence rate for the occurrence of total stroke (first-ever and recurrent) was 259 per 100,000 population (age- and sex-adjusted to 1995 US population). This new figure emphasized the importance of preventive measures for a disease that has identifiable and modifiable risk factors and of the development of new and improved treatment strategies and infrastructures that can reduce the consequences of stroke. The impact of new treatments for stroke was evaluated by examining the changes between 1990 to 1991 and 2000 to 2001 in in-hospital mortality rates and hospital charges in adult patients with stroke.[4] There had been an increase in the inflation-adjusted hospital charges for all patients with stroke and a reduction in mortality rates for all stroke subtypes, which was probably related to an increase in the proportion of patients with stroke admitted to urban teaching hospitals.

NIS data can provide clinical data such as clinical background or outcome; therefore, a researcher can analyze these data similar to a normal registry. Outcomes in acute stroke patients treated with thrombolysis were examined using the NIS database available in the United States for the years 1999–2002.[5] The thrombolysis cohort had a higher in-hospital mortality rate compared with nonthrombolysis patients (11.4%

vs. 6.8%). The rate of intracerebral hemorrhage was 4.4% for the thrombolysis cohort and 0.4% for nonthrombolysis patients. Multivariate logistic regression showed advanced age, Asian/Pacific Islander race, congestive heart failure, and atrial fibrillation/flutter to be independent predictors of in-hospital mortality after thrombolysis. Trends in therapy for cerebral aneurysms in the US were identified along with outcomes, using NIS data for the period 1993–2003.[6] Endovascular techniques for aneurysm occlusion have been increasingly used, while the use of surgical clipping procedures has remained stable. Toward the end of the study period, better overall outcomes were observed in the treatment of cerebral aneurysms, both ruptured and unruptured. Large academic centers were associated with better results, particularly for surgical clip placement. It was hypothesized that patients with ICH had a higher mortality risk if they were admitted to the hospital on the weekends than if they were admitted during the week.[7]

One advantage of nationwide claim data such as NIS is that they can provide data about rare diseases or special treatments that are difficult to obtain from a single facility. The incidence, mortality, and risk factors for pregnancy-related stroke in the United States from 2000–2001 were estimated.[8] A total of 2,850 pregnancy-related discharges included a diagnosis of stroke for a rate of 34.2 per 100,000 deliveries. There were 117 deaths or 1.4 per 100,000 deliveries. African-American women are at an increased risk, as are women aged 35 years and older. Risk factors, not previously reported, include lupus, blood transfusion, and migraine headaches. The acute stroke hospitalization rates for children and young adults (aged 15–44 years) and the prevalence of stroke risk factors among children and young adults hospitalized for acute stroke[9] has been examined. The prevalence of hospitalizations of acute ischemic stroke increased among all age and gender groups except females aged 5–14 years. Hypertension, diabetes, obesity, lipid disorders, and tobacco use were among the most common coexisting conditions, and their prevalence increased during the period of study among adolescents and young adults hospitalized with acute ischemic stroke.

About 250 articles are published using the NIS database related to stroke. They become an important evidence for epidemiology or health economy about stroke.

Attempt using Big Data about Stroke in Japan—J-ASPECT study

The DPC is a mixed-case patient classification system that was launched in 2002 by the Ministry of Health, Labor, and Welfare of Japan (MHLW) and was linked with a lump-sum payment system.[22] This system collects important data during hospitalization in addition to the characteristics of the unique reimbursement system. Each patient's background information or discharge summary, which includes principal diagnosis, complications, comorbidities, and outcomes are recorded in the administrative database associated with the DPC system. These patient data are coded using the International Classification of Diseases and Injuries 10th Revision (ICD-10) code. The DPC project collects three types of information: Form 1 is a clinical summary that contains information on diagnosis and severity. The E file has information about the bundled charge of the procedure and the F-file indicates the detail of the bundled procedures. Form 1, E-file, and F-file are matched according to the ID number that is unique for each discharged case. Using these data, we can describe the process of each in-patient treatment. From the point of view of Big Data, the DPC is regarded as a large sample of the structured inpatient EHR data in Japan.

The J-ASPECT study was performed to examine the associations between PSC and CSC capabilities and the impact of CSC capabilities on the hospital volume of stroke interventions. This cross-sectional survey used the DPC discharge database from participating institutions in the J-ASPECT study.

Impact of Comprehensive Stroke Care Capacity on the Hospital Volume of Stroke Interventions[10]

In 2000, the brain attack coalition discussed the concept of stroke centers and proposed two types of centers: comprehensive stroke centers (CSCs) and primary stroke centers (PSCs).[23,24] Most patients with stroke can be treated appropriately at a PSC, and the joint commission has established programs for certifying PSCs and evaluating their performance.[25] The concept and recommended key components of CSCs enable intensive care and specialized techniques that are not available at most PSCs. A set of metrics and associated data elements that cover the major types of care distinguishing CSCs from PSCs have been published previously.[23,24]

In the J-ASPECT study, a 49-question survey was developed on hospital characteristics (i.e., number of beds, academic status, geographic location, and participation in the DPC payment system), PSC and CSC capacity, and hospital volume of stroke interventions. The questionnaire was mailed on February 2011 to 1369 certified training institutions of the Japan Neurosurgical Society, Japanese Society of Neurology, and Japan Stroke Society. This survey included 25 items related to the five

major components of CSCs (personnel, diagnostic programs, specific expertise, infrastructure, and educational components) and five items related to PSC certification (Table 1). CSC scores were divided with/without the availability of a t-PA protocol into quintiles and analyzed the trend with the

Table 1 Number (percentage) of responding hospitals (n = 749) with the recommended items of comprehensive stroke care capacity

Components	Items	n	%
Personnel	Neurologists	358	47.8
	Neurosurgeons	694	92.7
	Endovascular physicians	272	36.3
	Critical care medicine	162	21.6
	Physical medicine and rehabilitation	113	15.1
	Rehabilitation therapy	742	99.1
	Stroke rehabilitation nurses*	102	13.8
Diagnostic (24/7)	CT*	742	99.2
	MRI with diffusion	647	86.4
	Digital cerebral angiography*	602	80.8
	CT angiography*	627	84
	Carotid duplex ultrasound*	257	34.5
	TCD*	121	16.2
Specific expertise	Carotid endarterectomy*	603	80.6
	Clipping of IA	685	91.5
	Hematoma removal/ draining	689	91.9
	Coiling of IA	360	48.1
	Intra-arterial reperfusion therapy	498	66.5
Infrastructure	Stroke unit*	132	17.6
	Intensive care unit	445	59.4
	Operating room staffed 24/7*	451	60.4
	Interventional services coverage 24/7	279	37.3
	Stroke registry*	235	31.7
Education	Community education*	369	49.4
	Professional education*	436	58.6

CT: computed tomography, IA: intracranial aneurysm, MRI: magnetic resonance imaging, TCD: transcranial Doppler.

Data missing: stroke rehabilitation nurse, 9; CT, 1;digital cerebral angiography, 4; CT angiography, 3; carotid endarterectomy, 1; carotid duplex, 3; TCD, 3; stroke unit, 1; operating room staffed, 2; stroke registry, 7; community education, 2; professional education, 5. Reproduced from Iihara et al.[10] with permission from the publisher. Copyright © 2014 National Stroke Association.

Cochran-Armitage trend test and multivariable linear regressions for the hospital volume.

Approximately 749 hospitals responded to the survey. On performing multivariate analysis adjusted for hospital characteristics, the total CSC score, but not the availability of a t-PA protocol, was associated with the volume of all types of interventions with a clear increasing trend (P for trend < 0.001) (Fig. 1).

This study demonstrated a significant impact of comprehensive stroke care capacity represented by the total CSC score on the hospital volume of stroke interventions and unique aspects of comprehensive stroke care capacity in Japan.

Impact of CSC capabilities on in-hospital mortality in patients with stroke[11]

Among the institutions that responded to the questionnaire on CSC capacity, data on patients hospitalized for stroke between April 1 2010 and March 31 2011 were obtained from the Japanese DPC database. In-hospital mortality was analyzed with hierarchical logistic regression analysis adjusted for age, sex, level of consciousness on admission, comorbidities, and the number of fulfilled CSC items in each component and in total. Hierarchical logistic regression models were used to estimate odds ratios (ORs) for in-hospital mortality adjusting for institutional level difference. Each model had two levels of hierarchy (hospital and patient) while considering the random effects of hospital variation, as well as fixed effects of CSC score and patient effects of age, sex, and level of consciousness. The total score and each subcategory score were analyzed separately. CSC scores were divided into quintiles and analyzed the trend with the Cochran-Armitage trend test.

Data from 265 institutions and 53,170 emergency-hospitalized patients were analyzed (Table 2). Mortality adjusted for age, sex, and level of consciousness was significantly correlated with personnel, infrastructural, educational, and total CSC scores in patients with ischemic stroke. Mortality was significantly correlated with diagnostic, educational, and total CSC scores in patients with ICH and with specific expertise, infrastructural, educational, and total CSC scores in patients with SAH.

CSC capabilities were associated with reduced in-hospital mortality rates, and the relevant aspects of care were found to be dependent on stroke type (Fig. 2).

Problems and future perspectives of utilizing Big Data

There are some problems for the DPC-based clinical study. The most frequent problem is the accuracy of information.[26] As DPC data is closely related to
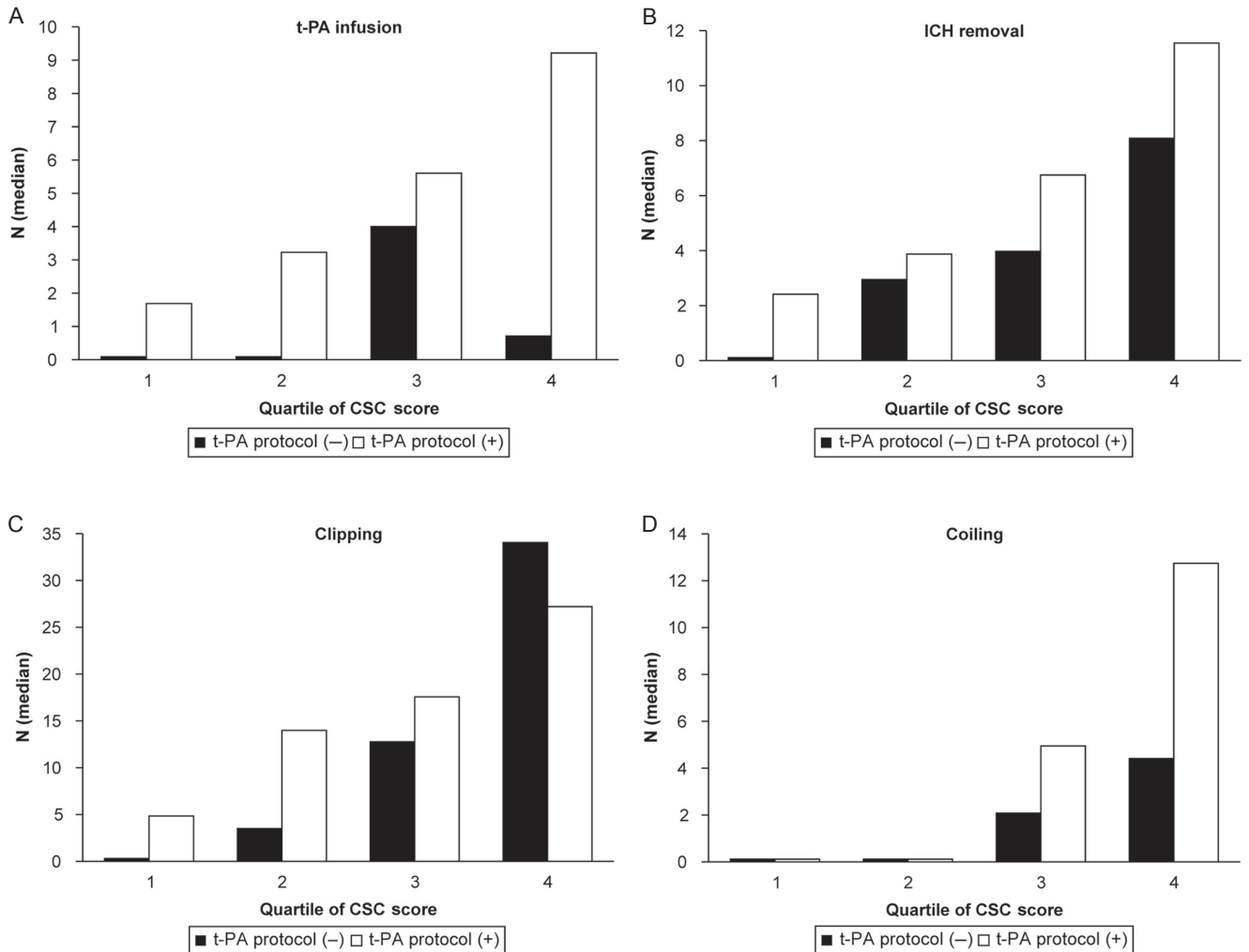
Fig. 1 **Associations between primary and comprehensive stroke care capabilities and case volume of stroke treatment in 2009 in Japan. The inclusion of total comprehensive stroke care (CSC) score, availability of a tissue-type plasminogen activator (t-PA) protocol, and other hospital characteristics in the model revealed that the total CSC score, but not the availability of a t-PA protocol, was significantly associated with the hospital volume of stroke interventions. Q, quintile. Reproduced from Iihara et al.[10] with permission from the publisher. Copyright © 2014 National Stroke Association.**

the payment, there is a possibility that the medical staff may allocate inappropriate diagnosis in order to obtain more reimbursements. The second problem is the possibility of sampling bias. The current DPC database covers only acute in-patient cases.

Furthermore, DPC data do not provide detailed medical information in comparison with general large-scale cohort studies. Because DPC data are limited to hospitalized data, it can provide information about outcomes such as complications or hospitalized death; however, it cannot provide data such as foreign progress or long-term convalescence.

While DPC data have these limitations, the advantage compared with the other databases is that it can cover all types of diseases treated in acute care

facilities. Furthermore, we can obtain data about rare diseases or special treatments, which are difficult to obtain from a single facility.

There have been few academic papers on a large-scale clinical study originating in Japan. One of the reasons is that most Japanese clinical studies are small-scale studies based on close hospital groups, i.e., one university hospital and its associate facilities, and the case registry database is limited to some diseases and domains. Therefore, nationwide statistics about various diseases and treatments are insufficient in Japan. Because nationwide claim data such as DPC data can be obtained from multiple centers with the same format, it is expected as a valid solution to these problems.

Table 2 Demographics of the patient study cohort at the time of diagnosis and hospital characteristics according to stroke type

	Total (n = 53,170)	Ischemic stroke (n = 32,671)	Intracerebral hemorrhage (n = 15,699)	Subarachnoid hemorrhage (n = 4,934)
Male, n (%)	29,353 (55.2)	18,816 (57.6)	9,030 (57.5)	1,584 (32.1)
Age, mean years ± SD	72.5 ± 13.1	74.4 ± 12.2	70.7 ± 13.5	64.7 ± 14.8
Hypertension, n (%)	39,918 (75.1)	22,531 (69.0)	13,281 (84.6)	4,229 (85.7)
Diabetes Mellitus, n (%)	13,725 (25.8)	9,318 (28.5)	3,278 (20.9)	1,174 (23.8)
Hyperlipidemia, n (%)	15,015 (28.2)	11,104 (34.0)	2,529 (16.1)	1,412 (28.6)
Medications during hospitalization				
Anti-renin-angiotensin system agent	34,136 (64.2)	17,694 (54.2)	12,537 (79.9)	4,019 (81.5)
Ca channel antagonist	25,984 (48.9)	10,469 (32.0)	11,719 (74.6)	3,903 (79.1)
Sympathetic antagonist	6,334 (11.9)	3,821 (11.7)	2,172 (13.8)	364 (7.4)
*β-blocker, α,β-blocker	4,357 (8.2)	3,048 (9.3)	1,133 (7.2)	188 (3.8)
α-blocker	2,374 (4.5)	953 (2.9)	1,232 (7.8)	200 (4.1)
Diuretic agent	9,950 (18.7)	5,860 (17.9)	3,074 (19.6)	1,049 (21.3)
Loop diuretic	7,434 (14.0)	4,609 (14.1)	1,912 (12.2)	940 (19.1)
Other diuretic	4,425 (8.3)	2,527 (7.7)	1,653 (10.5)	255 (5.2)
Antidiabetic agent	10,295 (19.4)	6,784 (20.8)	2,473 (15.8)	1,075 (21.8)
Insulin	7,654 (14.4)	4,597 (14.1)	2,044 (13.0)	1,046 (21.2)
Oral antidiabetic agent	5,749 (10.8)	4,459 (13.6)	1,110 (7.1)	197 (4.0)
Antihyperlipidemic agent	12,387 (23.3)	9,264 (28.4)	1,839 (11.7)	1,310 (26.6)
Statin	10,099 (19.0)	7,840 (24.0)	1,366 (8.7)	912 (18.5)
Antiplatelet agent	23,635 (44.5)	21,746 (66.6)	625 (4.0)	1,298 (26.3)
Aspirin	11,929 (22.4)	11,119 (34.0)	378 (2.4)	447 (9.1)
Japan Coma Scale				
0, n (%)	19,635 (36.9)	15,027 (46.0)	3,620 (23.1)	1,024 (20.8)
1-digit code, n (%)	19,371 (36.4)	12,375 (37.9)	5,934 (37.8)	1,117 (22.6)
2-digit code, n (%)	6,937 (13.0)	3,396 (10.4)	2,705 (17.2)	852 (17.3)
3-digit code, n (%)	7,227 (13.6)	1,873 (5.7)	3,440 (21.9)	1,941 (39.3)
Emergency admission by ambulance, n (%)	31,995 (60.2)	17,336 (53.1)	10,909 (69.5)	3,830 (77.6)
Average days in hospital (range)	21 (11–40)	20 (12–38)	22 (10–43)	30 (12–54)
Hospital characteristics (CSC scores)				
Total score (25 items)		16.7 ± 3.8	16.8 ± 3.4	17.1 ± 3.4
Personnel (7 items)		3.7 ± 1.2	3.7 ± 1.2	3.8 ± 1.2
Diagnostic techniques (6 items)		4.4 ± 1.1	4.5 ± 1.0	4.5 ± 1.0
Specific expertise (5 items)		4.4 ± 1.0	4.4 ± 0.9	4.5 ± 0.8
Infrastructure (5 items)		2.8 ± 1.3	2.9 ± 1.3	2.9 ± 1.3
Education (2 items)		1.4 ± 0.8	1.4 ± 0.8	1.4 ± 0.8

CSC: comprehensive stroke center. *A composite variable with a pure beta antagonist and a mixed alpha/beta adrenergic antagonist (e.g., labetalol). Reproduced from Iihara et al.[11] with permission.
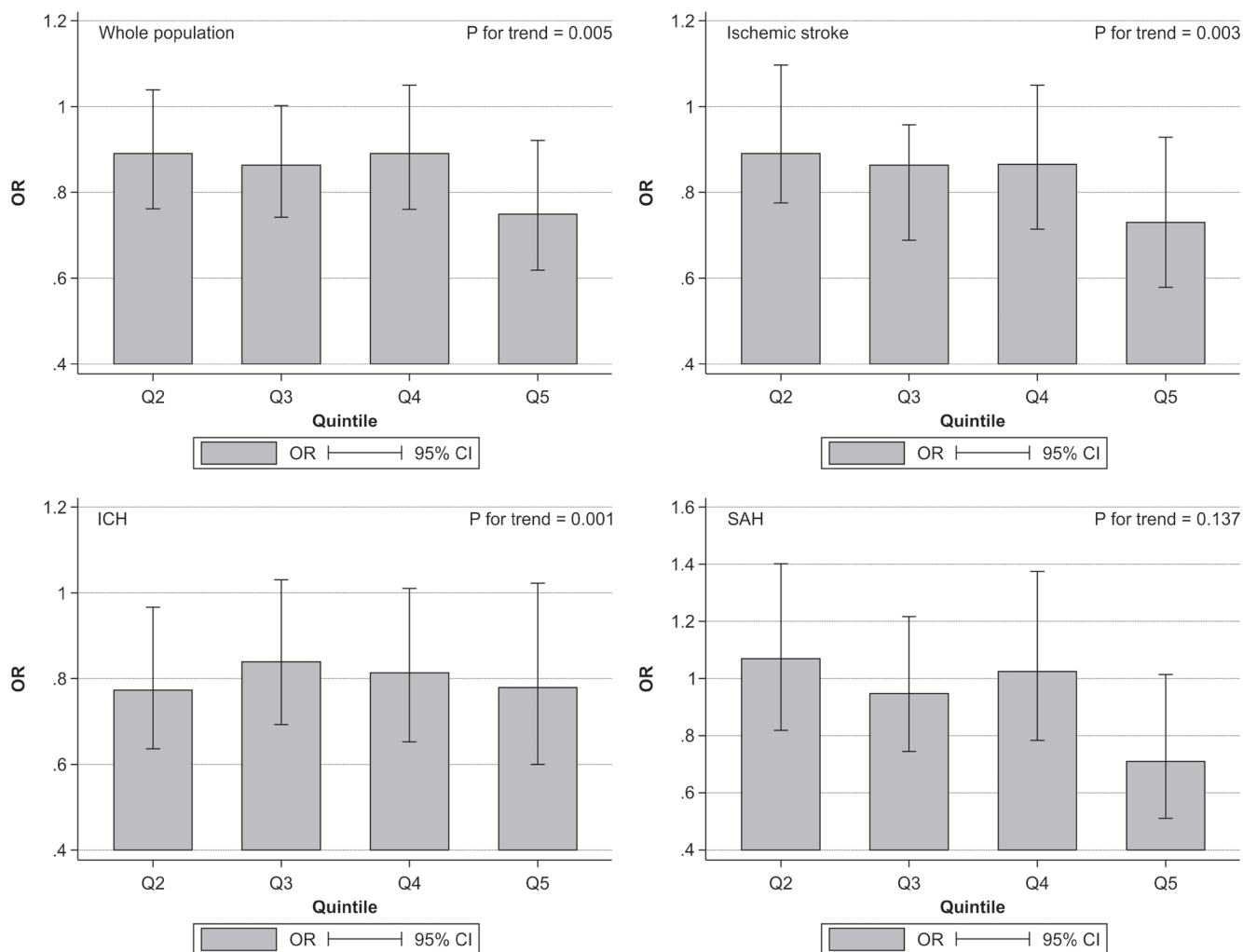
Fig. 2 Associations between total comprehensive stroke care scores separated into quintiles (Q) and in-hospital mortality of patients after all types of stroke. Odds ratios (ORs) and 95% confidence intervals (CIs) of in-hospital mortality for each quintile are depicted compared with that of Q1 as the control (Q1, 4–12 points; Q2, 13–14 points; Q3, 15–17 points; Q4, 18 points; Q5, 19–23 points). Reproduced from Iihara et al.[11] with permission.

As for J-ASPECT study, it is advantage that the overall database about stroke can be established and important information can be provided effectively without bothering stroke physicians. Furthermore, the facilities that participated can feedback their medical treatment process and the outcome of stroke compared to other Japanese hospitals; this benchmark promotes improvements in the quality of the stroke care in each facility.

To make up for the shortcomings of the DPC data and realize a higher quality clinical epidemiologic study, it will be important to link the DPC database with other databases. In the United States, the surveillance, epidemiology, and end results (SEER)-Medicare database is constructed for cancer research. The SEER-Medicare data reflect the linkage of two large population-based sources of data that provide detailed information about Medicare beneficiaries with cancer. The data come from the SEER program of cancer registries that collect clinical, demographic, and the cause of death information for persons with cancer and the Medicare claims for covered health care services from the time of a person's Medicare eligibility until death.[27]

In the field of stroke care, claim data and registry is expected to be linked. In Japan, the construction of the nationwide stroke registry has been planned in order to foresee the development of the basic law for stroke measures. In the future, the linkage of this stroke registry and DPC database can result in a unique population-based source of information that can be used for an array of epidemiological and health services research. Furthermore, it can help

discard the manual work required to enter the data in each facility and reduce the burden on prostrate stroke physicians.

Conclusion

The use of Big Data is expected as an effective modality that establishes new evidence about stroke care. The J-ASPECT study demonstrated the importance of the impact of CSC capacity and CSC capabilities on in-hospital mortality in stroke using the DPC database, one of the Big Data databases in Japan. The advantages of applying Big Data such as DPC to stroke care are that the overall database about stroke can be established without bothering stroke physicians and it can become a large-scale clinical study originating in Japan. The development of a new method of stroke registry using Big Data is expected as it would greatly improve future stroke care.

Disclosures

Acknowledgement

References

1) Laney D: 3D data management: controlling data volume, velocity, and variety. https://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf
2) Mayer-Schonberger V, Cukier K: Chapter 1: Now. In: Mayer-Schonberger V, Cukier K, eds. Big data: A revolution that will transform how we live, work, and think. *Boston: Boston Houghlin Mifflin Harcourt* 1–18, 2013
3) Williams GR, Jiang JG, Matchar DB, Samsa GP: Incidence and occurrence of total (first-ever and recurrent) stroke. *Stroke* 30: 2523–2528, 1999
4) Qureshi AI, Suri MF, Nasar A, Kirmani JF, Ezzeddine MA, Divani AA, Giles WH: Changes in cost and outcome among US patients with stroke hospitalized in 1990 to 1991 and those hospitalized in 2000 to 2001. *Stroke* 38: 2180–2184, 2007
5) Bateman BT, Schumacher HC, Boden-Albala B, Berman MF, Mohr JP, Sacco RL, Pile-Spellman J: Factors associated with in-hospital mortality after administration of thrombolysis in acute ischemic stroke patients: an analysis of the nationwide inpatient sample 1999 to 2002. *Stroke* 37: 440–446, 2006
6) Andaluz N, Zuccarello M: Recent trends in the treatment of cerebral aneurysms: analysis of a nationwide inpatient database. *J Neurosurg* 108: 1163–1169, 2008
7) Crowley RW, Yeoh HK, Stukenborg GJ, Medel R, Kassell NF, Dumont AS: Influence of weekend hospital admission on short-term mortality after intracerebral hemorrhage. *Stroke* 40: 2387–2392, 2009
8) James AH, Bushnell CD, Jamison MG, Myers ER: Incidence and risk factors for stroke in pregnancy and the puerperium. *Obstet Gynecol* 106: 509–516, 2005
9) George MG, Tong X, Kuklina EV, Labarthe DR: Trends in stroke hospitalizations and associated risk factors among children and young adults, 1995–2008. *Ann Neurol* 70: 713–721, 2011
10) Iihara K, Nishimura K, Kada A, Nakagawara J, Toyoda K, Ogasawara K, Ono J, Shiokawa Y, Aruga T, Miyachi S, Nagata I, Matsuda S, Ishikawa KB, Suzuki A, Mori H, Nakamura F: J-ASPECT Study Collaborators: The impact of comprehensive stroke care capacity on the hospital volume of stroke interventions: a nationwide study in Japan: J-ASPECT study. *J Stroke Cerebrovasc Dis* 23: 1001–1018, 2014
11) Iihara K, Nishimura K, Kada A, Nakagawara J, Ogasawara K, Ono J, Shiokawa Y, Aruga T, Miyachi S, Nagata I, Toyoda K, Matsuda S, Miyamoto Y, Suzuki A, Ishikawa KB, Kataoka H, Nakamura F, Kamitani S: Effects of comprehensive stroke care capabilities on in-hospital mortality of patients with ischemic and hemorrhagic stroke: J-ASPECT study. *PLoS One* 9: e96819, 2014
12) McAfee A, Brynjolfsson E: Big data: the management revolution. *Harv Bus Rev* 90: 60–66, 68, 128, 2012
13) Lynch C: Big data: How do your data grow? *Nature* 455: 28–29, 2008
14) Merelli I, Perez-Sanchez H, Gesing S, D'Agostino D: Managing, analysing, and integrating Big Data in medical bioinformatics: open problems and future perspectives. *Biomed Res Int* 2014: 1–13, 2014
15) Tempany CM, Jayender J, Kapur T, Bueno R, Golby A, Agar N, Jolesz FA: Multimodal imaging for improved diagnosis and treatment of cancers. *Cancer* 121: 817–827, 2015
16) Cain AA, Kosara R, Gibas CJ: GenoSets: visual analytic methods for comparative genomics. *PLoS One* 7: e46401, 2012

17) Chui KK, Wenger JB, Cohen SA, Naumova EN: Visual analytics for epidemiologists: understanding the interactions between age, time, and disease with multi-panel graphs. *PLoS One* 6: e14683, 2011

18) Manyika J, Chui M, Brown B, Bughin J, Dobbs R, Roxburgh C, Byers AH: Big Data: The next frontier for innovation, competition, and productivity. *McKinsey Global Institute*, 2011

19) Wolfe PJ: Making sense of Big Data. *Proc Natl Acad Sci U S A* 110: 18031–18032, 2013

20) Costa FF: Big Data in biomedicine. *Drug Discov Today* 19: 433–440, 2014

21) Databases. Healthcare Cost and Utilization Project (HCUP). November 2015. Agency for Healthcare Research and Quality, Rockville, MD. http://www.hcup-us.ahrq.gov/nisoverview.jsp.

22) Yasunaga H, Ide H, Imamura T, Ohe K: Impact of the Japanese diagnosis procedure combination-based payment system on cardiovascular medicine-related costs. *Int Heart J* 46: 855–866, 2005

23) Alberts MJ, Latchaw RE, Selman WR, Shephard T, Hadley MN, Brass LM, Koroshetz W, Marler JR, Booss J, Zorowitz RD, Croft JB, Magnis E, Mulligan D, Jagoda A, O'Connor R, Cawley CM, Connors JJ, Rose-DeRenzy JA, Emr M, Warren M, Walker MD: Brain Attack Coalition: Recommendations for comprehensive stroke centers: a consensus statement from the brain attack coalition. *Stroke* 36: 1597–1616, 2005

24) Alberts MJ, Hademenos G, Latchaw RE, Jagoda A, Marler JR, Mayberg MR, Starke RD, Todd HW, Viste KM, Girgus M, Shephard T, Emr M, Shwayder P, Walker MD: Recommendations for the establishment of primary stroke centers. Brain Attack Coalition. *JAMA* 283: 3102–3109, 2000

25) Reeves MJ, Parker C, Fonarow GC, Smith EE, Schwamm LH: Development of stroke performance measures: definitions, methods, and current measures. *Stroke* 41: 1573–1578, 2010

26) Matsuda S, Fujimori K, Kuwabara K, Ishikawa KB, Fushimi K: Diagnosis procedure combination as an infrastructure for the clinical study. *Asian Pac J Dis Manag* 5: 81–87, 2011

27) National Cancer Institute. SEER-Medicare Linked Database. http://healthcaredelivery.cancer.gov/seer-medicare/

Address reprint requests to: Koji Iihara, MD, PhD, Department of Neurosurgery, Graduate School of Medical Sciences, Kyushu University, 3-1-1 Maidashi, Higashi-ku, Fukuoka 812-8582, Japan.
e-mail: kiihara@ns.med.kyushu-u.ac.jp