## Original Article

# Comparison of registry and government evaluation data to ascertain severe trauma cases in Japan

Shinji Nakahara,[1] Tetsuya Sakamoto,[1] Takashi Fujita,[1] Tomohide Koyama,[1] Yoichi Katayama,[2] Seizan Tanabe,[3] and Yasuhiro Yamamoto[4]

[1]*Department of Emergency Medicine, Teikyo University School of Medicine, Tokyo,* [2]*Department of Emergency Medicine, Sapporo Medical University, Sapporo,* [3]*Emergency Life-Saving Technique Academy of Tokyo, and* [4]*Foundation for Ambulance Service Development, Tokyo, Japan*

*Aims:* Accurate evaluation of health care quality requires high-quality data, and case ascertainment (confirming eligible cases and deaths) is a foundation for accurate data collection. This study examined the accuracy of case ascertainment from two Japanese data sources.

*Methods:* Using hospital-level data, we investigated the concordance in ascertaining trauma cases between a nationwide trauma registry (the Japan Trauma Data Bank) and annual government evaluations of tertiary hospitals between April 2012 and March 2013. We compared the median values for trauma case volumes, numbers of deaths, and case fatality rates from both data sources, and also evaluated the variability in discrepancies for the intrahospital differences of these outcomes.

*Results:* The analyses included 136 hospitals. In the registry and annual evaluation data, the median case volumes were 120.5 cases and 180.5 cases, respectively; the median numbers of deaths were 11 and 12, respectively; and the median case fatality rates were 8.1% and 6.4%, respectively. There was broad variability in the intrahospital differences in these outcomes.

*Conclusions:* The observed discordance between the two data sources implies that these data sources may have inaccuracies in case ascertainment. Measures are needed to evaluate and improve the accuracy of data from these sources.

*Key words:* Case ascertainment, quality of care, quality of data, trauma

## INTRODUCTION

TRAUMA CARE PROVIDERS should routinely collect clinical data to facilitate quality assurance efforts, as care performance evaluation is a mandate for all health-care professions.[1] The commonly used quality indicators, such as risk-adjusted outcomes (e.g., case fatality rate), are typically derived from administrative or registry data. The Japan Trauma Data Bank (JTDB) is a nationwide trauma registry that provides data for quality assurance.

Appreciable quality improvements rely on valid evaluations that use accurate data, although previous studies have questioned the accuracy of the previously mentioned data sources. In this context, administrative and registry data have significant discordance in information that is used for risk adjustment (coding, severity, or comorbidities) and case ascertainment, which is an essential component of outcome indicators.[2–5] Most comparisons of administrative and registry data consider the registry data as the reference standard,[3] although trauma registry data can include various inaccuracies, including record incompleteness or discrepancies from the original medical records, and variability in coding and scoring across hospitals.[6–8]

The JTDB is known to have data inaccuracy, especially regarding diagnostic misclassification and data incompleteness.[9] Moreover, incomplete case ascertainment may exist, based on the voluntary nature of case registration in the JTDB, although this issue has never been investigated. Another available data source is the annual hospital evaluations that are undertaken by the Ministry of Health, Labour and Welfare (MHLW), although its case ascertainment has never been validated. Therefore, the present study investigated the concordance in ascertainment of trauma cases and deaths in tertiary hospitals from the two data sources.

## METHODS

### Study setting

IN JAPAN, TERTIARY hospitals that are accredited by the MHLW as emergency critical care centers (ECCCs) provide critical care to severely injured and ill patients.[10,11] Requirements for the accreditation are shown in Document S1. Attempts to standardize trauma care practices have been made by introducing Japanese trauma care guidelines,[12] and the MHLW annually evaluates the quality of emergency care in ECCCs. Although ECCCs are expected to provide the highest level of surgical and medical care in each area, the roles of ECCCs in trauma care vary regionally, based on the region's injury patterns and the abilities of the surrounding facilities. Some ECCCs treat a relatively small number of trauma cases.

### Study design

This retrospective observational study evaluated a merged hospital-level dataset that included registry data from the JTDB and the annual MHLW hospital evaluation data. We examined the concordance in the case volumes, deaths, and case fatality rates between the two data sources. The study protocol was approved by the ethics committee of the Teikyo University School of Medicine (Tokyo, Japan).

### Data sources and collection

The JTDB is a nationwide trauma registry that collects patient-level prehospital and clinical data from tertiary hospitals. Trauma cases with an Abbreviated Injury Scale (AIS) score of ≥3 are registered after they are admitted to participating hospitals, and the data are manually entered using web-based systems at each hospital. The collected data include injury mechanisms, physiological status at the scene and at hospital arrival, time courses, details regarding the examination and treatment, and in-hospital deaths. Most participating hospitals are ECCCs, although a few non-ECCC hospitals that actively practice trauma care also register cases. As participation is voluntary, some ECCCs do not register data. The steering committee of the JTDB annually distribute the pooled data to the participating hospitals, after cleaning the data and sending feedback to hospitals with noticeable amounts of missing data.

The annual evaluation by the MHLW collects hospital-level data regarding indicators of structures, processes, and outcomes from the ECCCs. Scores are assigned to the structure and process indicators (37 items with a total score of 101), and 29 items are directly related to patient care (a total score of 87).[13] The outcome indicators include the number of treated patients and deaths according to disease type. Trauma cases with an AIS score of ≥3 or that are undergoing emergency surgery are included (Doc. S2). The MHLW requests that each ECCC annually report the above information from the previous fiscal year (April to March), and the evaluation does not collect patient-level data.

### Study population

We selected ECCCs that provided data to both the evaluation and the registry regarding patients who were treated between April 2012 and March 2013. The 2013 evaluation collected data from 259 ECCCs for that period. Among these centers, 139 ECCCs provided data to the registry during the same period, although three ECCCs were excluded because they did not register any eligible patients (AIS score of ≥3 excluding cardiopulmonary arrest cases at hospital arrival) (Fig. 1). Thus, the analysis included 136 ECCCs.

### Outcome variables

The primary outcome variable was annual case volumes, and the secondary outcome variables were numbers of in-hospital deaths and case fatality rates (in-hospital deaths divided by case volume). All outcome variables were reported as hospital-level data regarding patients with an
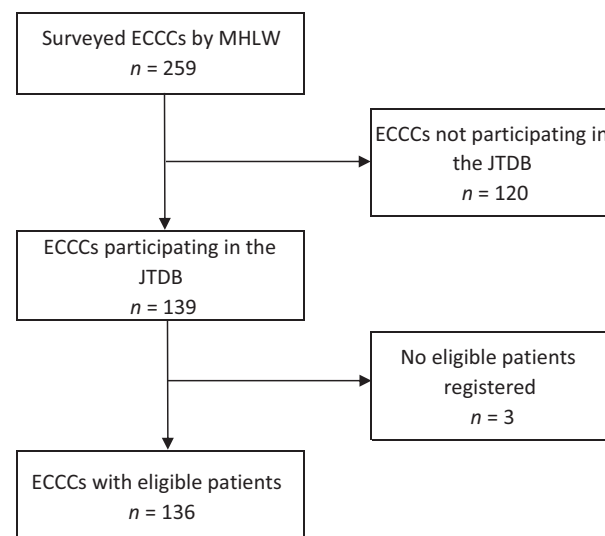


**Fig. 1.** Hospitals included in the analysis. ECCC, emergency critical care centers; JTDB, Japan Trauma Data Bank; MHLW, Ministry of Health, Labour and Welfare.

AIS score of ≥3, which were derived from the registry by aggregating its patient-level data. We excluded cases with cardiopulmonary arrest at hospital arrival, which was defined as patients with a systolic blood pressure and respiratory rate of zero, from the calculations of case volume and the number of deaths. Cases with missing values regarding systolic blood pressure or respiratory rate were included.

## Statistical analysis

The hospitals' characteristics were reported based on the annual evaluation data, which provided information regarding the case volumes, numbers of deaths, and evaluation scores. The characteristics of patients who were registered in the JTDB by the eligible hospitals were reported, including the patients' demographic characteristics, injury mechanisms, and injury severities.

We compared the median hospital-level outcome values between the two data sources using the Wilcoxon test (the data were paired for each hospital), because the outcome variables were not normally distributed. The Wilcoxon test is equivalent to the paired *t*-test, which tests a null hypothesis that the mean difference is zero. According to the methods described by Bland and Altman,[14] we calculated the intrahospital differences in the outcome variables between the two sources (evaluation data subtracted from the registry data at each hospital) and the average of the two values to evaluate the variability in the intrahospital differences. We plotted the differences against the average on a graph, and the average values were categorized into quartiles. We did not perform a scatter plot to avoid the possibility that the analyzed hospitals could be identified based on their values, as hospitals with extreme values would be easily identifiable. These analyses did not require adjustment for the case mix or hospital characteristics, as the hospitals were compared to themselves (intrahospital comparisons).

Using the Wilcoxon test, we calculated the effect size using the method described by Kerby,[15] as well as the *P*-values from null-hypothesis significance testing ($P < 0.05$ was considered significant). The effect sizes indicate the magnitude of differences or associations, and effect sizes of >0.2 are usually considered significant.[16] The sample size was defined as the number of hospitals in the datasets, and the effect size that could be detected based on a sample of 136 hospitals was calculated using G*Power software.[17] With $\alpha = 0.05$ (two-sided) and $\beta = 0.2$, this sample allows a paired Wilcoxon test to detect a standardized effect size of 0.25. All other analyses were carried out using IBM SPSS software (version 23; IBM, Armonk, NY, USA).

## RESULTS

### Hospital and patient characteristics

THE ANALYZED HOSPITALS ($n = 136$) had higher values for case volume, number of deaths, case fatality rates, and evaluation score, compared to the excluded hospitals ($n = 123$) (Table 1). The analyses included data from 21,535 trauma cases that were recorded in the registry by 136 ECCCs. Most patients were male, ≥65 years old, had normal physiological status, and had relatively minor injuries (Injury Severity Score <15) (Table 2). Most of the injuries were unintentional, and falls accounted for the greatest proportion of injuries (followed by traffic accidents). The case fatality rate was 7.8%, and survival information was missing for 959 cases (4.5%).

### Main results

The median case volume and median number of deaths were lower in the registry data, compared to the evaluation data (120.5 versus 180.5 and 11 versus 12, respectively) (Table 3). The median case fatality rate was higher in the registry, compared to the evaluation (8.1% versus 6.4%). There was also broad variability in the intrahospital differences between the two data sources in terms of case volumes, deaths, and case fatality rates (Table 3). Figure 2

**Table 1.** Characteristics of emergency critical care centers based on data from the Japanese Ministry of Health, Labour and Welfare ($n = 259$)

|  | Analyzed centers $n = 136$ | Excluded centers $n = 123$ |
|---|---|---|
| Case volume at each center, median (IQR) | 180.5 (109.5–298.0) | 86 (44–189) |
| Number of deaths at each center, median (IQR) | 12 (7–20) | 5 (2–12) |
| Total evaluation score at each center, median (IQR) | 73 (64.0–83.5) | 66 (57–74) |
| Subtotal patient care score at each center, median (IQR) | 61.5 (52–72) | 54 (46–62) |
| Total case volume; (all centers), *n* | 30,183 | 16,521 |
| Total deaths (all centers), *n* (%) | 2,034 (6.7) | 927 (5.6) |

IQR, interquartile range.

**Table 2.** Characteristics of patients who were extracted from the Japan Trauma Data Bank (*n* = 21,535)

| Characteristics | *n* (%) |
| --- | --- |
| Sex | |
| Male | 12,903 (59.9) |
| Female | 8,623 (40.0) |
| Data missing | 9 (0.04) |
| Age, years | |
| 0–14 | 1,022 (4.7) |
| 15–24 | 1,780 (8.3) |
| 25–34 | 1,429 (6.6) |
| 35–44 | 1,742 (8.1) |
| 45–54 | 1,871 (8.7) |
| 55–64 | 2,782 (12.9) |
| 65–74 | 3,452 (16.0) |
| 75–84 | 4,328 (20.1) |
| ≥85 | 3,099 (14.4) |
| Data missing | 30 (0.1) |
| Glasgow Coma Scale score | |
| 13–15 | 15,746 (73.1) |
| 9–12 | 1,519 (7.1) |
| 3–8 | 2,353 (10.9) |
| Data missing | 1,917 (8.9) |
| Systolic blood pressure, mmHg | |
| ≥90 | 19,921 (92.5) |
| <90 | 1,548 (7.2) |
| Data missing | 66 (0.3) |
| Respiratory rate, per min | |
| <10 | 193 (0.9) |
| 10–29 | 15,637 (72.6) |
| ≥30 | 2,099 (9.7) |
| Data missing | 3,606 (16.7) |
| Injury Severity Score | |
| <15 | 11,276 (52.4) |
| 15–24 | 5,648 (26.2) |
| 25–44 | 4,081 (19.0) |
| 45–75 | 516 (2.4) |
| Data missing | 14 (0.1) |
| Intent | |
| Unintentional | 19,784 (91.9) |
| Self-inflicted | 917 (4.3) |
| Violence | 223 (1.0) |
| Undetermined | 407 (1.9) |
| Data missing | 204 (0.9) |
| Injury mechanism | |
| Fall | 11,277 (52.4) |
| Traffic | 7,658 (35.6) |
| Other blunt | 1,325 (6.2) |
| Penetrating | 405 (1.9) |
| Burn | 352 (1.6) |
| Data missing | 518 (2.4) |

**Table 2.** (Continued)

| Characteristics | *n* (%) |
| --- | --- |
| Survival | |
| Alive | 18,896 (87.7) |
| Dead | 1,680 (7.8) |
| Data missing | 959 (4.5) |

shows the variability in intrahospital differences for the outcome variables in relation to the average values between the two data sources by quartile. Many hospitals had intrahospital differences that were dispersed away from zero. Hospitals with high values for the average outcome variables tended to have greater variability.

## DISCUSSION

THE PRESENT STUDY revealed discordance in case volume, number of deaths, and case fatality data from a trauma registry (JTDB) and the MHLW annual evaluations. We found that the JTDB data had lower case volumes and number of deaths, and higher case fatality rates, compared to the evaluation data. Furthermore, we detected broad variability in the differences between the two data sources for intrahospital comparisons. Similar discordances between data sources have been detected in previous studies. For example, Pasquali *et al.*[18] reported lower case volumes and case fatality rates among pediatric patients undergoing congenital heart surgery in administrative data, compared to registry data (10.7% lower volume and 4.7% fewer fatalities). In addition, Phillips *et al.*[2] reported a lower case fatality rate in administrative data, compared to registry data (3.5% versus 5.2%).

However, we observed broad variability in the intrahospital differences in both directions. Some hospitals had lower numbers (negative differences) whereas others had higher numbers (positive differences) in the JTDB data, compared to the MHLW data. This result suggests that these data sources may have inaccuracy in case ascertainment. Aggregation might have cancelled out the positive and negative differences in each hospital, which would explain the relatively small differences in the median values. In contrast, previous studies by Pasquali *et al.*[18] and Phillips *et al.*[2] have revealed relatively small variability in the scatter diagrams of case volumes and case fatality rates from two data sources.

The different inclusion criteria between the data sources and across hospitals might have introduced some discordances.[2,4] For example, the MHLW criteria for data regarding various disease categories are somewhat arbitrary,[13]

**Table 3.** Comparison of data from the Japan Trauma Data Bank (registry) and annual hospital evaluations by the Japanese Ministry of Health, Labour and Welfare to ascertain severe trauma cases in Japan

| | Registry | Evaluation | *P*-value† | Effect size‡ | Differences§ | |
|---|---|---|---|---|---|---|
| | Median (IQR) | Median (IQR) | | | Median (IQR) | Minimum, maximum |
| Case volume, *n* | 120.5 (63, 237) | 180.5 (109.5, 298) | <0.001 | 0.59 | −44 (−132.5, −0.5) | −456, 402 |
| Number of deaths, *n* | 11 (4, 18) | 12 (7, 20) | 0.01 | 0.26 | −2 (−7, 3.5) | −46, 24 |
| Mortality rate, % | 8.1 (4.4, 12.5) | 6.4 (3.8, 11.7) | 0.03 | 0.22 | 1.5 (−2.3, 4.0) | −37.8, 95.9 |

†Wilcoxon test.

‡Effect sizes were calculated using rank sums in the Wilcoxon test: $\dfrac{|\text{sum of favorable ranks} - \text{sum of unfavorable ranks}|}{(\text{sum of favorable ranks} + \text{sum of unfavorable ranks})}$.

§Values for the evaluation data were subtracted from those of the registry data for each center.

IQR, interquartile range.

which leaves the decision to include or excluded cases up to the discretion of each hospital. Thus, trauma cases with cardiopulmonary arrest at hospital arrival can be arbitrarily categorized as "trauma" or "cardiopulmonary arrest" cases. The MHLW data also have a category for hemorrhagic shock, and severe trauma cases with shock can be categorized as either "trauma" or "hemorrhagic shock" cases. Furthermore, the evaluation data can include patients who are undergoing emergency surgery without an AIS score of ≥3, although there is no clear definition of "emergency."

Nevertheless, the different criteria alone cannot fully explain the observed broad variability, and two factors may increase the differences in both directions, which would result in broader variability. First, incomplete case registration in the JTDB may lead to lower numbers, compared to the MHLW data. A noticeable number of hospitals registered fewer cases in the JTDB, with extremely large discrepancies in some hospitals. This result reflects the fact that participation in the JTDB is voluntary and without external support, as resource-constrained hospitals may not be able to complete the patient registration process, which is a demanding task that requires financial, technical, and human resources.[19]

Second, possible patient registration from outside of the ECCCs may lead to higher numbers in the JTDB, compared to the MHLW data. The MHLW data only cover patients who were treated in ECCCs, whereas the JTDB data may cover patients who were treated outside of ECCCs. Emergency critical care centers are usually part of large hospitals that may also have non-ECCC departments (e.g., surgical, orthopedic, or neurosurgical departments) that treat less severe trauma cases with an AIS score of ≥3. Thus, in some hospitals, cases that were registered in the JTDB may be from both ECCC and non-ECCC departments.

Although certain levels of discordance may be inevitable between different data sources with different purposes and inclusion criteria, the significant discordances with large variabilities that were observed in the present study probably resulted from inaccuracies in case ascertainment. Accurate data is a prerequisite for valid quality evaluation in health care, and inaccurate data may mislead quality improvement and clinical practices.[1,6,7] In particular, case ascertainment provides a foundation for accurate data collection (what cases to target) and the basic components of the outcome indicators (the denominator and numerator).

The findings of the present study imply a serious need to improve the accuracy of case ascertainment in the JTDB and in the MHLW evaluations, in order to improve the data quality. For example, data quality evaluation mechanisms that are based on standardized quality measurements should be integrated into the data collection process for health care quality assurance.[6,8] In addition, it is important to standardize the definitions of cases and data (e.g., diagnoses and severity), and the data collection and registration processes, with specific training for registrars who perform trauma data extraction.[7]

The present study has several limitations. First, the MHLW data did not include patient-level data, and it was not possible to compare the diagnoses, severity scoring, and complications between the two data sources. However, we were able to evaluate the case ascertainment process, which controls the accuracy of the other data. Second, we only analyzed a small proportion of the ECCCs, and most excluded ECCCs did not participate in the JTDB, which was presumably related to resource constraints. However, the exclusion of these facilities likely did not distort our findings, as they would not have completed the case registration even if they had participated in the JTDB.

In conclusion, the present study revealed discordance with broad variabilities in the case ascertainment, number of deaths, and case fatality rates from the JTDB and the
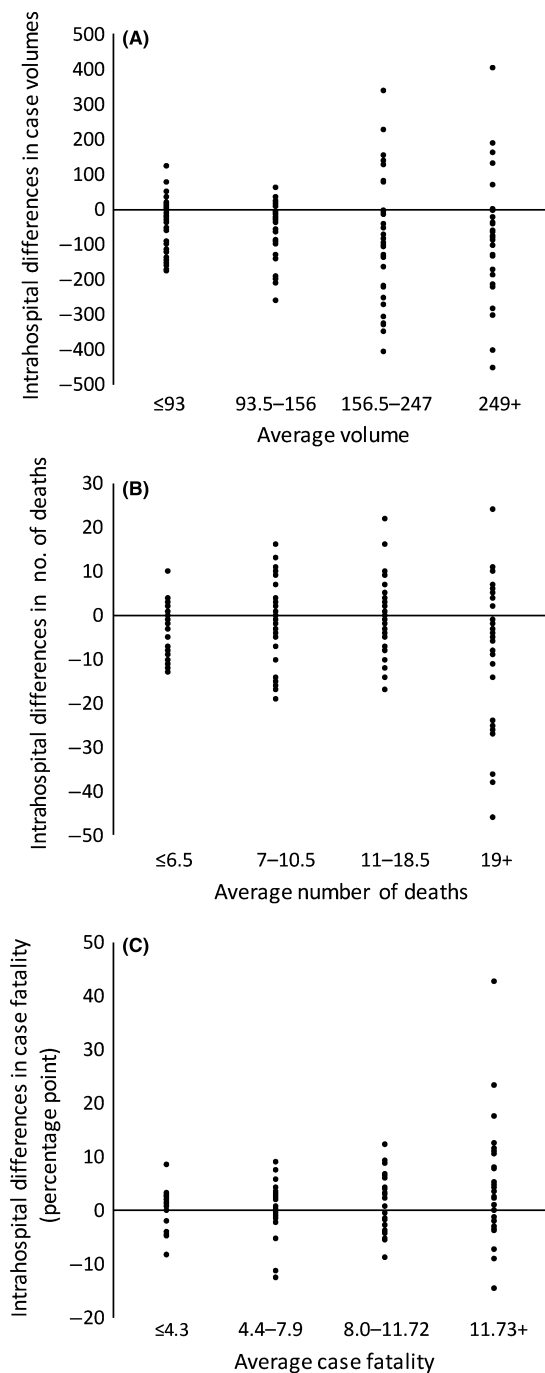
**Fig. 2.** Intrahospital differences in case volumes, number of deaths, and case fatality rates between data from the Japan Trauma Data Bank and annual hospital evaluations by the Ministry of Health, Labour and Welfare. The *Y*-axes indicate intrahospital differences and the *X*-axes indicate the average of the two values by quartile for case volumes (A), number of deaths (B), and case fatality rates (C).

MHLW data. This suggests that either or both of the data sources may have inaccuracies. Therefore, as inaccurate data can compromise the evaluation of care, measures are needed to improve data quality in these data sources.

## DISCLOSURE

Ethical approval: The study's protocol was approved by the ethics committee of the Teikyo University School of Medicine.

Informed consent: The requirement to obtain informed consent from the participants was waived because this study used registry data.

Conflict of interest: None declared.

## REFERENCES

1 Pronovost PJ, Miller M, Wachter RM. The GAAP in quality measurement and reporting. JAMA 2007; 298: 1800–2.
2 Phillips B, Clark DE, Nathens AB, Shiloach M, Freel AC. Comparison of injury patient information from hospitals with records in both the national trauma data bank and the nationwide inpatient sample. J. Trauma 2008; 64: 768–79. discussion 779–780.
3 McCarthy ML, Shore AD, Serpi T, Gertner M, Demeter L. Comparison of Maryland hospital discharge and trauma registry data. J. Trauma 2005; 58: 154–61.
4 Hackworth J, Askegard-Giesmann J, Rouse T, Benneyworth B. The trauma registry compared to All Patient Refined Diagnosis Groups (APR-DRG). Injury 2017; 48: 1063–8.
5 Wynn A, Wise M, Wright MJ, *et al.* Accuracy of administrative and trauma registry databases. J. Trauma 2001; 51: 464–8.
6 Porgo TV, Moore L, Tardif PA. Evidence of data quality in trauma registries: a systematic review. J. Trauma Acute Care Surg. 2016; 80: 648–58.
7 Arabian SS, Marcus M, Captain K, *et al.* Variability in interhospital trauma data coding and scoring: a challenge to the accuracy of aggregated trauma registries. J. Trauma Acute Care Surg. 2015; 79: 359–63.
8 Dente CJ, Ashley DW, Dunne JR, *et al.* Heterogeneity in trauma registry data quality: implications for regional and national performance improvement in trauma. J. Am. Coll. Surg. 2016; 222: 288–95.
9 Tohira H, Matsuoka T, Watanabe H, Ueno M. Characteristics of missing data of the Japan Trauma Data Bank. Nihon Kyukyu Igakukai Zasshi. 2011; 22: 147–55.

10 Hori S. Emergency medicine in Japan. Keio J. Med. 2010; 59: 131–9.

11 Ministry of Health, Labour and Welfare. Kyukyu Iryo Taisaku Jigyo Jisshiyoko. Tokyo: The Ministry, 2015.

12 Japan Association for the Surgery of Trauma. Japan Advanced Trauma Evaluation and Care (JATEC) guidelines, 4th edn. Tokyo: Herusu Shuppan, 2012.

13 Ministry of Health, Labour and Welfare. Kyumei-center no atarashii jujitsudo hyoka nitsuite. Tokyo: The Ministry, 2009.

14 Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986; 1: 307–10.

15 Kerby DS. The simple difference formula: an approach to teaching nonparametric correlation. Compr. Psychol. 2014; 3: 1–9.

16 Ferguson CJ. An effect size primer: a guide for clinicians and researchers. Prof. Psychol. Res. Pract. 2009; 40: 532–8.

17 Faul F, Erdfelder E, Buchner A, Lang AG. Statistical power analyses using G*Power 3.1: tests for correlation and regression analyses. Behav. Res. Methods 2009; 41: 1149–60.

18 Pasquali SK, He X, Jacobs JP, *et al.* Measuring hospital performance in congenital heart surgery: administrative versus clinical registry data. Ann. Thorac. Surg. 2015; 99: 932–8.

19 Ichikawa M, Nakahara S, Wakai S. Factors related to participation in the Japan Trauma Data Bank. Nihon Kyukyu Igakukai Zasshi. 2005; 16: 552–6.

## SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

**Doc. S1.** Requirements to be accredited as an emergency critical care center.

**Doc. S2.** Definitions of the outcome indicator (annual number of critical cases).

**Doc. S3.** Formula to calculate the standardized effect size for a Wilcoxon test.