


Development of a set of quality indicators in paediatric and perinatal care in Japan with a modified Delphi method

Daisuke Shinjo ^{1,2}, Nobuaki Ozawa,³ Naoya Nakadate,⁴ Yutaka Kanamori,⁵ Kimikazu Matsumoto,⁶ Takashi Noguchi,² Shosuke Ohtera,⁷ Hitoshi Kato⁸

To cite: Shinjo D, Ozawa N, Nakadate N, *et al*. Development of a set of quality indicators in paediatric and perinatal care in Japan with a modified Delphi method. *BMJ Paediatrics Open* 2023;**7**:e002209. doi:10.1136/bmjpo-2023-002209

► Additional supplemental material is published online only. To view, please visit the journal online (<http://dx.doi.org/10.1136/bmjpo-2023-002209>).

Received 28 July 2023
Accepted 30 October 2023

ABSTRACT

Backgrounds Few paediatric and perinatal quality indicators (QIs) have been developed in the Japanese setting, and the quality of care is not assured or validated. The aim of this study was to develop QIs in paediatric and perinatal care in Japan using an administrative database and confirm the feasibility and applicability of the indicators using a single-site practice test.

Methods We used a RAND-modified Delphi method that integrates evidence review with expert consensus development. QI candidates were generated from clinical practice guidelines (CPGs) available in English or Japanese and existing QIs in nine selected paediatric or perinatal conditions. Consensus building was based on independent panel ratings. The performance of QIs was retrospectively assessed using data from an administrative database at the National Children's Hospital. Data between April 2018 and March 2019 were used, while data between April 2019 and March 2021 were also used for selected condition, considering the small number of patients. Each QI was calculated as follows: number of times the indicator was met/number of participants×100.

Results From the literature review conducted between 2010 and 2020, 124 CPGs and 193 existing indicators were identified to generate QI candidates. Through the consensus-building process, 133 QI candidates were assessed and 79 QIs were accepted. The practice test revealed wide variations in the process-level performance of QIs in four categories: patient safety: median 43.9% (IQR 16.7%–85.6%), general paediatrics: median 98.8% (IQR 84.2%–100%), advanced paediatrics: median 94.4% (IQR 46.0%–100%) and advanced obstetrics: median 80.3% (IQR 59.6%–100%).

Conclusions We established 79 QIs for paediatric and perinatal care in Japan using an administrative database that can be applied to hospitals nationwide. The practice test confirmed the measurability of the developed QIs. Benchmarking these QIs will be an attractive approach to improving the quality of care.

INTRODUCTION

Many efforts have been made to improve the quality of care; these efforts include the identification of specific, measurable indicators of quality. Quality indicators

WHAT IS ALREADY KNOWN ON THIS TOPIC

- ⇒ Quality indicators (QIs) and their benchmarks between hospitals are widely used in many healthcare areas to improve the quality of care.
- ⇒ Few paediatric and perinatal QIs have been developed in Japanese settings to date; the quality of care in paediatrics and perinatal care in Japan is not validated.

WHAT THIS STUDY ADDS

- ⇒ A total of 79 QIs for paediatric and perinatal care in Japan were established based on an administrative database using the RAND/UCLA appropriateness method.
- ⇒ The practice test based on one institution confirmed the measurability of the developed QIs.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTISE OR POLICY

- ⇒ Developed QIs can be applied to nationwide hospitals, providing useful information that could contribute to improving quality of care and updating the health provision system and health policy.

(QIs) are measurable elements of practice performance for which there is evidence or consensus and which are systematically developed with guidelines that can be used to assess the appropriateness of specific healthcare services and outcomes. They can highlight and reveal quality issues as a benchmark, which in turn helps provide solutions.^{1–3}

These QIs and their benchmarks between hospitals are widely used worldwide, including in paediatric and perinatal care.^{4–8} For example, Braithwaite *et al* developed paediatric QIs in Australia, focusing on 17 clinical conditions, and revealed that the overall prevalence of adherence to quality-of-care indicators was not high.⁴ However, due to the variations in healthcare systems and culture, direct application of foreign-developed QIs may not be suitable for different countries



© Author(s) (or their employer(s)) 2023. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

For numbered affiliations see end of article.

Correspondence to

Daisuke Shinjo; dshinjo.hci@tmd.ac.jp

like Japan, where limited paediatric and perinatal QIs exist and care quality remains unvalidated.^{9–12}

The hospital information system is a rich data source for QI calculation, but differing vendor data structures pose challenges requiring extensive preparation efforts.¹³ Alternatively, administrative databases with standardised data structures offer an efficient, population-based approach to monitoring care quality and have been used to compare QIs in prior studies.^{6 14–17} Prior to publication and benchmarking, practice tests are essential for evaluating feasibility.¹⁸ Accepting QIs must undergo review, as some may not be measurable or applicable in certain healthcare systems.¹⁹

This study's primary aim was to develop paediatric and perinatal QIs in Japan using an administrative database, facilitating nationwide benchmarking to improve care quality as part of a quality improvement framework. We selected perinatal care instead of obstetrics care, considering its association with children. A secondary objective was to conduct a pilot practice test at one hospital to confirm QI feasibility and applicability.

MATERIALS AND METHODS

Overview of the development process

The project followed a process based on a modified Delphi technique (the Rand Corporation (RAND)/University of California, Los Angeles appropriateness

method),²⁰ which has been widely used to develop QIs.²¹ We applied and modified the development process based on previous literature regarding developing QIs.^{4 22 23} The method integrates an evidence review, a face-to-face multidisciplinary panel meeting and repeated anonymous ratings to build consensus.²⁴ The development and ratification of QIs are depicted in figure 1.

Multidisciplinary panel

Panel members, comprising seven healthcare professionals—five paediatricians, one obstetrician and one public health specialist—selected the conditions for QI and consensus development. The sampling strategy was a non-random selection aimed at seeking participants who would be informative, with recommendations by DS and approval by TN and HK. We selected members who had held any one of the following responsibilities: being a board or committee member of the medical academic society, engaged in guideline development in selected conditions and having outstanding research achievements in related research fields.

Selection of conditions for QI development

Four categories (patient safety, general paediatrics, advanced paediatrics and advanced obstetrics), followed by nine conditions regarding paediatric and perinatal care, were identified based on published research, the burden of disease, frequency of presentation and

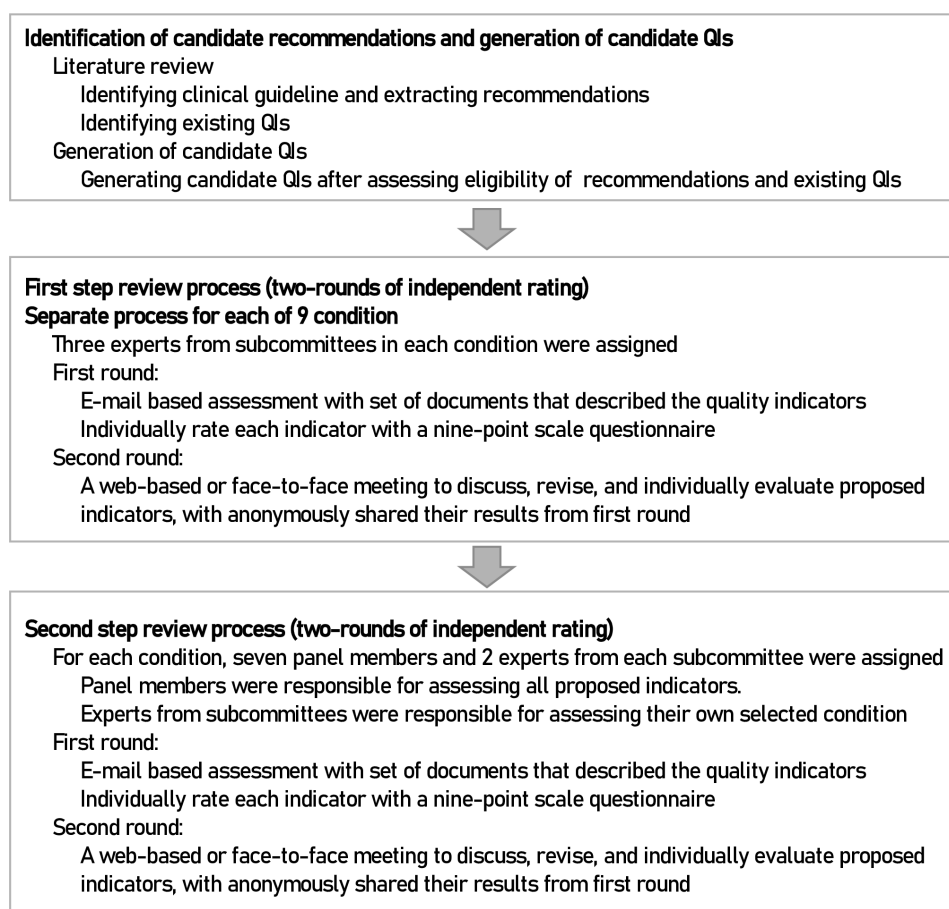


Figure 1 Flow diagram of QI development and ratification. QIs, quality indicators.

national priority areas. The term 'advanced' was used for conditions that are usually cared for specialised centres such as tertiary hospitals. These included conditions with high prevalence, such as paediatric bronchial asthma, neonatal respiratory care, caesarean sections and Kawasaki diseases. Kawasaki disease is one of the major diseases in Japan, with more than 28 500 patients identified between January 2019 and December 2020.²⁵ These categories and conditions were identified through discussions at the panel meeting. Two conditions among these retain two subconditions: (1) 'rare diseases', consisting of acute lymphoblastic leukaemia and congenital diaphragmatic hernia, and (2) 'acute abdomen', consisting of intussusception and appendicitis.

Systematic search for evidence

Because de novo development of evidence-based QIs is costly and time-consuming, methods using existing clinical practice guidelines (CPGs) have gained interest as viable alternatives.²⁶ Thus, for selected paediatric and perinatal conditions, we retrieved existing CPGs and QIs available in English or Japanese from QI/guideline databases and one medical literature database (PUBMED) in 2020, which was published between April 2010 and March 2020 (details of the search formula are shown in online supplemental table 1). The following QI/guideline databases were used: Agency for Healthcare Research and Quality (USA), National Quality Forum (USA), National Institute for Health and Care Excellence (NICE, UK), Australian National Health and Medical Research Council, Minds with the Japan Council for Quality Healthcare and the National Hospital Organisation (Japan). Websites regarding these selected conditions, including those of related paediatric associations, were also reviewed. Furthermore, we manually searched to identify literature that might be relevant to this study.

Indicator development

Recommendations in the selected guidelines were extracted from the CPGs. Each recommendation was screened for eligibility from the viewpoints of (1) strength of recommendation (relatively strong in each condition); (2) validity and adequacy in actual clinical practice in Japanese settings; and (3) feasibility of defining indicators using administrative databases. Recommendations that did not match the above-mentioned three criteria were excluded. The remaining recommendations were then converted into a standardised indicator format using the modified American College of Cardiology/American Heart Association methodology.²² Existing QIs were also converted into a standardised indicator format. We designed QIs based on the format of the Japanese administrative database.

Subcommittees for each condition

An expert coordinator was appointed to review the proposed indicators under each condition. The subcommittee consisted of five experts recruited to undertake a

review of the proposed indicators. Three experts participated in the first-step ratings, and two experts participated in the second-step ratings. They assessed the proposed indicators based on their selected subcommittees' conditions, while members of the multidisciplinary panel were responsible for assessing all proposed indicators.

Expert consensus process

The proposed indicators were reviewed and ratified by experts from subcommittees in each condition and a multidisciplinary panel with two-step, two-round independent ratings. Three experts from subcommittees reviewed the proposed indicators in the first-step rating, while nine members (two experts from subcommittees (in each condition) and seven panel members) reviewed the indicators in the second-step rating. During each step/round, members rated the appropriateness of each QI candidate on a 9-point scale, where 1 and 9 represented 'least suitable' and 'most suitable', respectively. QIs were adopted according to the following criteria: the median individual rating in each round/step was >7; the number of members who scored <3 was one or fewer in the first-step rating and <3 was two or fewer in the second-step rating. In addition, members were given the opportunity to provide comments or suggest additional candidates, especially with regard to supporting panel members for ratings. In round 1, members individually evaluated indicators with a set of documents that described the QIs adapted from nine domains: evidence-based, interpretable, actionable, denominator, numerator, validity, reliability, feasibility and overall assessment.²² In round 2, they convened for a web-based or face-to-face meeting to discuss, revise and individually evaluate the proposed indicators, anonymously sharing their results from the first round. If additional candidates were presented after the meeting, they discussed them via email using the same postal questionnaire.

We have explained the development process to panel and subcommittee members in detail, and all members agreed to the process. The features of the study are as follows: (1) QIs would be developed in a wide range of paediatric and perinatal care; (2) QI candidates were developed by guidelines and existing QIs so that each condition may have different quality standards and/or levels of evidence and (3) assessment of QI candidates was two-round-two-step ratings by subcommittees' specialists (first step and second step) and panel members (second step).

Pilot practice test for feasibility

We used data from the Japanese Administrative Database, the Diagnosis Procedure Combination (DPC) per-diem payment system (details of the DPC have been described elsewhere).²⁷ In brief, the DPC is a case-mix patient classification system linked to payments at acute-care and mixed-care hospitals in Japan. Anonymous clinical and administrative claims data were included in the database. Clinical data comprised baseline patient information,

diagnosis (based on ICD-10) and detailed medical information, including all major or minor procedures, medication and device use. The DPC database was used for many epidemiological research including QIs.^{9–12 27 28} A programme for data checking was provided by the government to support the quality of the data, and the data were uniformly collected.

We collected DPC data from the National Center for Child Health and Development, which is the only national children's hospital, between April 2018 and March 2019. We further collected data on QIs for rare diseases between April 2019 and March 2021, considering the small number of patients. For each indicator, percentage scores (QIs) were calculated as follows: number of times the indicator was met/number of participants (excluding those who had obvious reasons for not implementing the process as defined by the indicator) ×100. The medians of the indicator scores were also computed as the overall quality score of the programme. To ensure feasibility, we further checked the data in cases where the percentage scores were lower than 5%. Data processing was performed using the Microsoft SQL server (Microsoft Corporation, Seattle, WA, USA).

RESULTS

Development of quality indicators

From the literature review, we extracted 124 CPGs and 193 indicators. A total of 132 QI candidates pertaining to the Japanese healthcare system were aggregated from the CPGs and existing QIs after the exclusion of those screened for eligibility. All panel members and subcommittee experts from 21 institutions agreed to participate in this study. All responded to the email-based surveys, attended meetings and owned their responsibilities. The consensus development process was completed in October 2021.

Figure 1 illustrates the QI development process. The first-step ratings resulted in 98 QIs being selected from among the 132 QI candidates; 33 indicator candidates were not adopted. The second-step ratings selected 78 QI candidates and added a new indicator. Some QI candidates were modified and included as additional candidates, following the suggestions of experts, and reassessed. Consequently, 79 QIs were established (figure 2). These 79 QIs comprised 19 QIs for patient safety, 23 for general paediatrics, 23 for advanced paediatrics and 14 for advanced obstetrics (figure 3). Among the 79 QIs, 76 are process measures for expressing the proportion of patients who received appropriate care; the remaining 3 QIs are outcome measures. Examples of the indicators are shown in table 1 (a full list is in online supplemental table 2).

Performance in the pilot practice test

The results of the pilot practice tests are presented in table 2. Seven QIs (rare diseases) used data for fiscal years between 2018 and 2020, and the remaining 72 QIs used

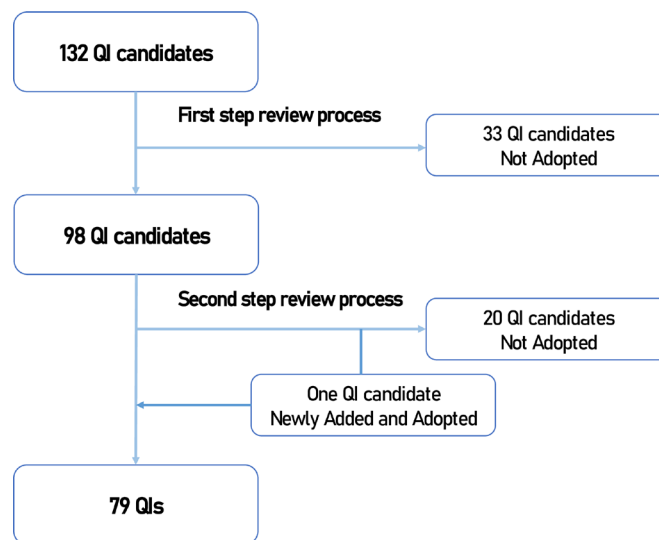


Figure 2 QI candidates and adopted QIs. QIs, quality indicators.

data for fiscal year 2018. Unfortunately, three QIs among the 79 QIs were not calculated (PS12, pulse oximeter monitoring in paediatric sedation during MRI examination; NRC02, budesonide inhalation for neonates at very high risk of bronchopulmonary dysplasia and RDA03, morphological bone marrow examination and diagnosis by a haematologist (for acute lymphoblastic leukaemia)). Most QIs were process-level due to the characteristics of the database and its feasibility. There were wide variations in the process-level performance of QIs in four categories for process measures (patient safety: median 43.9%, IQR 16.7%–85.6%, general paediatrics: median 98.8%, IQR 84.2%–100%, advanced paediatrics: median 94.4%, IQR 46.0%–100% and advanced obstetrics: median 80.3%, IQR 59.6%–100%). The details of each QI and the results of the pilot test are presented in online supplemental table 2.

DISCUSSION

We developed 79 QIs for paediatric and perinatal care using a RAND-modified Delphi method based on CPGs and existing QIs. These QIs are designed to be defined based on the Japanese administrative database. They have the advantages of being a cost-effective source of information, reducing the risk of selection bias and being easily adopted by participating hospitals. The use of QIs has the potential to raise the standards of quality in paediatric and perinatal care in Japan. Our single-site practice test showed that most of the proposed QIs were measurable in real-world clinical practice using administrative databases and that there was a wide variation in their performance.

QIs and their analyses are beneficial for identifying gaps and areas for improvement, informing the creation of future best practices, tracking progress in quality improvement and providing insights into the management of each condition. Poor adherence may affect

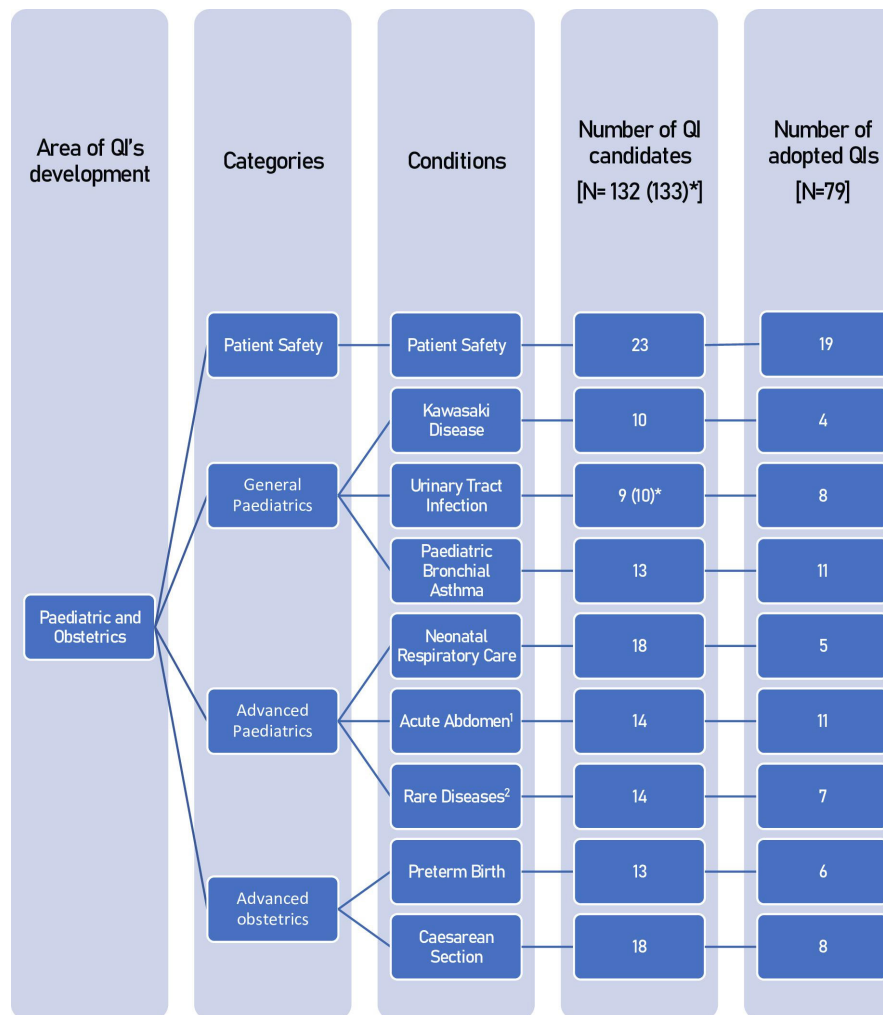


Figure 3 Number of QI candidates and adopted QIs by condition. *Including newly added QI candidates. ¹Acute abdomen consist of intussusception and appendicitis. ²Rare diseases consist of acute lymphoblastic leukaemia and congenital, diaphragmatic hernia. QIs, quality indicators. QIs, quality indicators

patient outcomes or lead to a substantial worsening of disease, death and increased healthcare costs. Adherence gaps and practice variation persist despite decades of development and endorsement of CPGs, and efforts have been made to close this gap.²⁹ Efforts to enhance the quality of guidelines, measure (monitor) and analyse QIs and improve adherence to guidelines are essential for achieving better health service provision. QIs could provide useful information regarding institutional-level and regional-level disparities in quality of care, which could contribute to updating health policy and improving the health provision system. While most QIs are localised considering the difference in healthcare systems and cultures, it would be helpful to consolidate QIs and related information worldwide considering generalisability; comparison of measurable QIs between countries/regions would also help quality improvement.

Our practice test was based on one institution but included all indicators developed in this study. Three proposed QIs were not calculated, partly because the data were not recorded in healthcare insurance regulations. Some QIs had a smaller number of patients

than expected, which is partly because of coding practices (eg, 'Acute focal bacterial nephritis' tended to be coded as 'Acute pyelonephritis'). Results also showed that the Japanese Administrative Database is one of the appropriate sources of information for QIs. Further benchmarking of QIs would be attractive in the Japanese setting, while it seemed hard to compare QIs across countries using administrative databases without a common data structure.³⁰

This study has several implications for future research. First, the feasibility of using these QIs in other hospitals needs to be evaluated. Second, the use of the proposed QIs needs to be evaluated; future studies should assess if and how these QIs contribute to quality improvement, including changes in the behaviour of physicians and the frequency of unwarranted events. Third, implementing these QIs requires continuous updates, evaluations and adaptation.²¹

Strength and limitations

One of the main strengths of this project is the use of routinely collected administrative data, which

**Table 1** Examples and characteristics of the proposed QIs for paediatric and perinatal care in Japan

Category	Condition	No. of QIs	Indicator ID	Description of selected indicators
Patient safety	Patient safety	19	PS08	Therapeutic drug monitoring of vancomycin
			PS14	Sedation without sedative during MRI examination (aged 7–9)
General paediatrics	Kawasaki disease	4	KD02	Recommended initial treatment
			KD03	Examination for hepatitis B and C virus before Infliximab administration
	Urinary tract infection	8	UTI01	Urinary culture and urine Gram staining before administration of antibacterial drugs
			UTI03	Antibacterial drugs for 7–14 days
	Paediatric bronchial asthma	11	PBA03	PaCO ₂ /PaO ₂ examination and chest X-ray for isoproterenol persistent inhalation and/or using mechanical ventilation
			PBA11	unplanned readmission within 7 days after discharge
Advanced paediatrics	Neonates (respiratory care)	5	NRC04	Caffeine administration to facilitate weaning from mechanical ventilation
			NRC05	Avoiding administration of dexamethasone with non-steroidal anti-inflammatory drugs (NSAIDs).
	Acute abdomen	11	AAI03	Avoiding CT scan at admission (for intussusception)
			AAA02	Laparoscopic appendectomy (for acute appendicitis)
	Rare diseases	7	RDA03	Morphological bone marrow examination and diagnosis by haematologist
			RDC03	Avoiding neuromuscular blocking agents at admission
Advanced obstetrics	Preterm birth	6	PB01	Corticosteroids administration for threatened preterm labour
			PB04	Avoiding tocolytic maintenance therapy
	Caesarean section	8	CS04	Planned caesarean section for gestation 38 weeks and more
			CS08	Pneumatic compression and/or compression stockings after caesarean section

QIs, quality indicators.

enables the monitoring of quality issues throughout the healthcare system. Second, administrative data are relatively inexpensive to collect compared with primary data collection.³¹ Although QIs may be biased by intrahospital heterogeneity,³² we do not suffer from sampling bias, which avoids non-response and recall

bias.³³ The final strength of the project is the inclusion of a data collection and calculation phase to assess the feasibility of indicator measurement.

Despite the advantages, the study acknowledges certain limitations. The composition of panels and subcommittees may influence consensus development, potentially

Table 2 Performance of proposed QIs for paediatric and perinatal care in Japan

	Process				Outcome			
	Number of QIs	Median (%)	25 Percentile (%)	75 Percentile (%)	Number of QIs	Median (%)	25 Percentile	75 Percentile
Patient safety	17	43.9	16.7	85.6	2	0.2		
General paediatrics	22	98.8	84.2	100.0	1	2.1		
Advanced paediatrics	23	94.4	46.0	100.0				
Advanced obstetrics	14	80.3	59.6	100.0				

QIs, quality indicators.

leading to a biased selection of QIs. Although a two-step, two-round assessment was implemented to reduce bias, the representativeness of panel members could still impact the validity of the consensus method. Notably, patient perspectives were not adequately included due to the absence of patient representation in the panels. Indeed, patient participation in QI development has been limited.^{34 35} This may be a general limitation of QI development based on guidelines; however, it needs to be resolved for further studies, such as updating these QIs.

Furthermore, the study focused on QIs applicable to the administrative database, excluding aspects of quality not relevant to this specific data source. The accuracy of the proposed QIs may have some limitations due to the coding process, even though efforts were made to ensure measurability by medical experts familiar with the database. Although the DPC (administrative) database was used for epidemiological studies including QIs^{9-12 27 28} with efforts on its validation studies,³⁶ differences in coding practices between hospitals could affect QI results, and the limited generalisation of indicator performance due to a single-site practice test also needs to be considered. Further practice tests with more hospitals were also required to strengthen its feasibility and assess its reliability and adaptability. In addition, the proposed QIs did not include timely administration of medications/procedures due to the unavailability of hours-level data.

It would be preferable to a set of QIs where the link between process and outcome is established. However, we have developed QIs without consideration of the linkage between process measures and outcome measures. This is partly because it is not easy to define appropriate outcome measures related to process measures using the administrative database. For example, the validation regarding adverse-related ICD-10 codes (T79-T88) is not established, and the timing of diagnosis is lacking in a Japanese setting. Further efforts focusing on the linkage would be an attractive approach to facilitating quality improvement.

Moreover, evidence-based QIs are only as robust as the underlying evidence they are based on. Existing problems with CPGs, such as redundancy, lack of currency and concerns about the quality of evidence, may systematically under-represent or overrepresent certain aspects in specific clinical areas.

Finally, it is essential to recognise that evidence-based QIs can only be as reliable as the underlying evidence they are based on. CPGs have been extensively studied, revealing issues such as redundancy, outdated information, inconsistent structure and content and overly lengthy documents. Moreover, concerns persist regarding the quality of evidence supporting CPGs, leading to potential under-representation or over-representation of certain aspects in specific clinical areas.

In conclusion, this study offers valuable insights into the development and application of QIs for paediatric and perinatal care in Japan. While the use of administrative

data is advantageous, there are notable limitations related to consensus development, patient perspectives, database relevance, coding accuracy, coding practices and evidence-based QIs. Awareness of these limitations helps to ensure the appropriate interpretation and utilisation of the proposed QIs in enhancing healthcare quality.

CONCLUSION

We have established 79 QIs for paediatric and perinatal care in Japan, drawing from pre-existing international CPGs and QIs and through an informal expert consensus process. These QIs can be applied to nationwide hospitals using an administrative database. While the practice test was conducted at a single site, it proved valuable in confirming the measurability of QIs and demonstrated how incentives, such as insurance coverage, can enhance performance in clinical practice and documentation processes. In the pursuit of improving the quality of paediatric and perinatal care, benchmarking these QIs emerges as an attractive approach.

Author affiliations

¹Department of Health Policy and Informatics, Tokyo Medical and Dental University, Bunkyo-ku, Tokyo, Japan

²Department of Information Technology and Management, National Center for Child Health and Development, Setagaya-ku, Tokyo, Japan

³Center for Maternal-Fetal, Neonatal and Reproductive Medicine, National Center for Child Health and Development, Setagaya-ku, Tokyo, Japan

⁴Division of Medical Security and Patient Safety, National Center for Child Health and Development, Setagaya-ku, Tokyo, Japan

⁵Division of Surgery, National Center for Child Health and Development, Setagaya-ku, Tokyo, Japan

⁶Children's Cancer Center, National Center for Child Health and Development, Setagaya-ku, Tokyo, Japan

⁷Department of Health Economics, National Center for Geriatrics and Gerontology, Obu, Aichi, Japan

⁸San-ikukai Hospital, Sumida-ku, Tokyo, Japan

Acknowledgements The authors wish to extend their gratitude to everyone involved in the study at the National Center for Child Health and Development and Tokyo Medical and Dental University for their cooperation. Special thanks also go to Professor Susumu Yokoya of the Fukushima Medical University for his inestimable comments and suggestions and to Wataru Mimura of the National Center for Global Health and Medicine for his valuable insights and continuous support.

Contributors All the authors contributed to the conception and design of the study. Material preparation, data collection and analysis were performed by DS, NO, NN, YK and KM. The first draft of the manuscript was written by DS; all authors commented on the previous versions of the manuscript and participated in its revision for intellectual content. All authors have read and approved the final manuscript and agreed to be accountable for all aspects of the work. Guarantor: DS.

Funding This study was supported by a Grant-in-Aid for Scientific Research (B) from the Japan Society for the Promotion of Science (JSPS KAKENHI, 20H03921 (to DS)), the Council for Science, Technology and Innovation, Cross-ministerial Strategic Innovation Promotion Program, 'Innovation AI Hospital System' (Funding Agency: National Institute of Biomedical Innovation, Health and Nutrition (NIBIOHN)) (SIP 18087257 (to DS)) and a grant from the National Center for Child Health and Development (2019b-7 (to TN)).

Competing interests None declared

Patient and public involvement The board waived the requirement for the acquisition of informed consent from the patients owing to the anonymisation of patient data.

Patient consent for publication Not required.

Ethics approval This study involves human participants. This study was approved by the Institutional Review Board at the National Center for Child Health and Development (2021-146). The board waived the requirement for the acquisition of informed consent from the patients owing to the anonymisation of patient data.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement No data are available.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID ID

Daisuke Shinjo <http://orcid.org/0000-0002-7495-7409>

REFERENCES

- Albert JM, Das P. Quality indicators in radiation oncology. *Int J Radiat Oncol Biol Phys* 2013;85:904–11.
- Stang AS, Straus SE, Crotts J, et al. Quality indicators for high acuity pediatric conditions. *Pediatrics* 2013;132:752–62.
- Stelfox HT, Straus SE, Nathens A, et al. Evidence for quality indicators to evaluate adult trauma care: a systematic review. *Crit Care Med* 2011;39:846–59.
- Braithwaite J, Hibbert PD, Jaffe A, et al. Quality of health care for children in Australia, 2012-2013. *JAMA* 2018;319:1113–24.
- Ewald DA, Huss G, Auras S, et al. Development of a core set of quality indicators for paediatric primary care practices in Europe, COSI-PPC-EU. *Eur J Pediatr* 2018;177:921–33.
- Sauvegrain P, Chantry AA, Chiesa-Dubruille C, et al. Monitoring quality of obstetric care from hospital discharge databases: a Delphi survey to propose a new set of indicators based on maternal health outcomes. *PLoS One* 2019;14:e0211955.
- Ruiz-Canela-Cáceres J, Aquino-Linares N, Sánchez-Díaz JM, et al. Indicators for childhood asthma in Spain, using the rand method. *Allergol Immunopathol (Madr)* 2015;43:147–56.
- Scanlon MC, Harris JM, Levy F, et al. Evaluation of the agency for healthcare research and quality pediatric quality indicators. *Pediatrics* 2008;121:e1723–31.
- Okubo Y, Horimukai K, Michihata N, et al. Recent practice patterns and variations in children hospitalized for asthma exacerbation in Japan. *Int Arch Allergy Immunol* 2020;181:926–33.
- Okubo Y, Hayakawa I, Nariai H, et al. Recent practice patterns in diagnostic procedures Anticonvulsants, and antibiotics for children hospitalized with febrile seizure. *Seizure* 2019;67:52–6.
- Okubo Y, Miyairi I, Michihata N, et al. Recent prescription patterns for children with acute infectious diarrhea. *J Pediatr Gastroenterol Nutr* 2019;68:13–6.
- Shinjo D, Matsumoto K, Terashima K, et al. Volume effect in paediatric brain tumour resection surgery: analysis of data from the Japanese national inpatient database. *Eur J Cancer* 2019;109:111–9.
- Grunfeld E, Lethbridge L, Dewar R, et al. Towards using administrative databases to measure population-based indicators of quality of end-of-life care: testing the methodology. *Palliat Med* 2006;20:769–77.
- Jenkins KJ, Koch Kupiec J, Owens PL, et al. Development and validation of an agency for healthcare research and quality indicator for mortality after congenital heart surgery harmonized with risk adjustment for congenital heart surgery (RACHS-1) methodology. *J Am Heart Assoc* 2016;5:e003028.
- De Schreye R, Smets T, Annemans L, et al. Applying quality indicators for administrative databases to evaluate end-of-life care for cancer patients in Belgium. *Health Aff (Millwood)* 2017;36:1234–43.
- Geary RS, Knight HE, Carroll FE, et al. A step-wise approach to developing indicators to compare the performance of maternity units using hospital administrative data. *BJOG* 2018;125:857–65.
- Blozik E, Reich O, Rapold R, et al. Evidence-based indicators for the measurement of quality of primary care using health insurance claims data in Switzerland: results of a pragmatic consensus process. *BMC Health Serv Res* 2018;18:743.
- Ewald DA, Huss G, Kraska RA, et al. Feasibility testing of the core set of quality indicators for paediatric primary care in Europe, COSI-PPC-EU. *Eur J Pediatr* 2019;178:707–19.
- Hommel I, van Gurp PJ, Tack CJ, et al. Perioperative diabetes care: development and validation of quality indicators throughout the entire hospital care pathway. *BMJ Qual Saf* 2016;25:525–34.
- Fitch K, Bernstein SJ, Aguilar MD, et al. *The RAND/UCLA Appropriateness Method User's Manual*. Santa Monica, CA: RAND Corporation, 2001.
- Kötter T, Blozik E, Scherer M. Methods for the guideline-based development of quality indicators--a systematic review. *Implement Sci* 2012;7:21.
- Spertus JA, Eagle KA, Krumholz HM, et al. American college of cardiology and American heart association methodology for the selection and creation of performance measures for quantifying the quality of cardiovascular care. *J Am Coll Cardiol* 2005;45:1147–56.
- Ueda K, Ohtera S, Kaso M, et al. Development of quality indicators for low-risk labor care provided by midwives using a RAND-modified Delphi method. *BMC Pregnancy Childbirth* 2017;17:315.
- Campbell SM, Braspenning J, Hutchinson A, et al. Research methods used in developing and applying quality indicators in primary care. *BMJ* 2003;326:816–9.
- Ae R, Makino N, Kuwabara M, et al. Incidence of Kawasaki disease before and after the COVID-19 pandemic in Japan: results of the 26th nationwide survey, 2019 to 2020. *JAMA Pediatr* 2022;176:1217–24.
- Blozik E, Nothacker M, Bunk T, et al. Simultaneous development of guidelines and quality indicators -- how do guideline groups act? A worldwide survey. *Int J Health Care Qual Assur* 2012;25:712–29.
- Hayashida K, Murakami G, Matsuda S, et al. History and profile of diagnosis procedure combination (DPC): development of a real data collection system for acute inpatient care in Japan. *J Epidemiol* 2021;31:1–11.
- Nishimura A, Nishimura K, Onozuka D, et al. Development of quality indicators of stroke centers and feasibility of their measurement using a nationwide insurance claims database in Japan — J-ASPECT study —. *Circ J* 2019;83:2292–302.
- Lieu TA, Wojtowicz A, Stratton K. Closing evidence gaps in clinical prevention: a report from the National academies of sciences, engineering, and medicine. *JAMA* 2022;327:815–6.
- Groene O, Kristensen S, Arah OA, et al. Feasibility of using administrative data to compare hospital performance in the EU. *Int J Qual Health Care* 2014;26 Suppl 1(Suppl 1):108–15.
- Riley GF. Administrative and claims records as sources of health care cost data. *Med Care* 2009;47(7 Suppl 1):S51–5.
- Lelong A, Léger S, Vendittelli F, et al. A quality indicator can be biased by intra-hospital heterogeneity: the case for quality of patient record keeping in France. *Eur J Public Health* 2015;25:787–91.
- Benchimol EI, Manuel DG, To T, et al. Development and use of reporting guidelines for assessing the quality of validation studies of health administrative data. *J Clin Epidemiol* 2011;64:821–9.
- Stelfox HT, Boyd JM, Straus SE, et al. Developing a patient and family-centred approach for measuring the quality of injury care: a study protocol. *BMC Health Serv Res* 2013;13:31.
- den Breejen EME, Nelen WLD, Schol SFE, et al. Development of guideline-based indicators for patient-centredness in fertility care: what patients add. *Hum Reprod* 2013;28:987–96.
- Yamana H, Konishi T, Yasunaga H. Validation studies of Japanese administrative health care data: a scoping review. *Pharmacoepidemiol Drug Saf* 2023;32:705–17.