BMJ Open

Updating quality indicators for low-risk labour care in Japan using current clinical practice guidelines: a modified Delphi method

Kayo Ueda,[1,2] Misato Kaso,[1] Shosuke Ohtera,[1] Takeo Nakayama[1]

To cite: Ueda K, Kaso M, Ohtera S, *et al*. Updating quality indicators for low-risk labour care in Japan using current clinical practice guidelines: a modified Delphi method. *BMJ Open* 2019;**9**:e023595. doi:10.1136/bmjopen-2018-023595

► Prepublication history and additional material for this paper are available online. To view these files, please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2018-023595).

Received 13 April 2018
Revised 17 October 2018
Accepted 12 December 2018

[1]Department of Health Informatics in the School of Public Health, Kyoto University, Kyoto, Japan
[2]Graduate School of Nursing Women's Health & Midwifery, School of Medicine Faculty of Nursing Department of Maternal Nursing, Nara Medical University, Kashihara, Japan

Correspondence to
Dr Kayo Ueda;
ueda.kayo.85a@st.kyoto-u.ac.jp

ABSTRACT

Objectives Quality indicators are measurable elements widely used to assess the quality of care. They are often developed from the results of systematic reviews or clinical practice guidelines. These sources are regularly updated in line with new clinical evidence, but there are few articles on updating quality indicators based on clinical practice guidelines. This study aimed to update the quality indicators developed for low-risk labour care in Japan in 2012, mainly drawing on new or updated clinical practice guidelines, and making the process clearly visible and assessable.

Design and setting We used a modified Delphi method for the update. The procedure included four steps: (1) updating the definition of low-risk labour; (2) reviewing the literature published between June 2012 and December 2015 using five guidelines and two quality indicator databases to extract potential candidate indicators; (3) formation of a multidisciplinary panel including mothers and (4) panel ratings (two rounds between February and April 2016) on the validity of the candidate indicators, and judging the validity of the previous quality indicators drawing on the new evidence.

Participants A multidisciplinary panel of 13 clinicians, including obstetricians, paediatricians and midwives, plus 3 non-clinician mothers.

Results The literature review identified 276 new recommendations from 27 clinical practice guidelines including 2 published in Japan and 21 quality indicators. We developed 13 new candidate indicators from these sources and panel recommendations, 12 of which were approved by the multidisciplinary panel. The panel also accepted all 23 existing quality indicators as still valid, resulting in a total of 35 quality indicators for low-risk labour.

Conclusions We successfully updated the quality indicators for low-risk labour care in Japan. The procedure developed may be useful for updating other quality indicators based on new clinical practice guidelines.

Strengths and limitations of this study

- ► Drawing on current clinical practice guidelines, we efficiently updated the quality indicators for low-risk pregnancy care using a modified Delphi method.
- ► After clarifying the method for recruiting a multidisciplinary panel, our panel included 13 clinicians and 3 mothers who were non-clinicians.
- ► The panel composition may have affected the consensus process of developing and updating quality indicators.
- ► This update procedure may be useful for updating quality indicators based on revised or new clinical practice guidelines in other fields.

BACKGROUND

The US Institute of Medicine has defined the quality of care as 'the degree to which health services for individuals and populations increase the likelihood of desired health outcomes and are consistent with current professional knowledge'.[1] Quality indicators are measurable elements that are widely used to assess the quality of care. They are usually drawn up from up-to-date published evidence including systematic reviews and clinical practice guidelines.[2 3]

It has long been recognised as necessary to update clinical practice guidelines and systematic reviews as new relevant studies are published.[4–7] This suggests that it is also important to update quality indicators based on clinical practice guidelines and systematic reviews,[8–13] but few articles have ever reported on this updating process. If clinical practice guidelines make new recommendations, it is essential to discuss whether to incorporate these into existing quality indicators.[4]

There are four reasons to update quality indicators. The first is because it is necessary to reassess the evidence, for example, because new clinical practice guidelines have been published.[14–17] The second is during a review of concepts or definitions of expert terms, such as revision of definitions and metrics.[18 19] The third is when adjusting the data collection system or method used by healthcare providers, such as updates to indicator content or clarifying and adjusting a

particular procedure.[14 19] The fourth is when there is a ceiling effect. We defined the 'ceiling effect' as a phenomenon in which the scores for the quality indicator are near their maximum value. If a ceiling effect occurs, it is impossible to increase the score significantly. It is, therefore, also very difficult to evaluate the quality improvement or detect differences between the measured score and other hospitals or over time.[20 21] Very few previous reports have sufficiently described the methods and processes used to update quality indicators.[6]

In Japan, quality indicators have rarely linked to clinical practice guidelines. There is no national data collection or reporting on patient safety in Japan, and this is not monitored nor shared across hospitals. The authors developed quality indicators for low-risk labour care in 2012. The annual number of live births in Japan is about 980 000, of which 92% are normal term infants. In 2016, the maternal mortality rate was 3.4 and the perinatal mortality rate was 3.6.[22]

The majority involve low-risk labour,[23] which seldom results in complications. We used clinical practice guidelines and quality indicators which were related to low-risk birth and published between July 2007 and June 2012. The results were published in English in 2017.[24]

Quality indicators for low-risk labour care are necessary for several reasons. First, healthcare providers need to share a common understanding of checks that will keep birth-related risks low, reducing potential complications and problems caused by insufficient communication. Second, healthcare providers can compare the quality of care with other facilities, to improve their practice. Third, quality indicators can be used to ensure midwives' accountability for low-risk labour care.

More recently in Japan, there has been an increasingly uneven distribution of obstetricians, a decrease in facilities specialised for childbirth. It has become difficult for obstetricians to be responsible for all low-risk births in the hospital, so systems allowing midwife-led care are being examined. It is, therefore, even more important to guarantee the quality of care for low-risk labour.

Since the development of the original quality indicators in 2012, new clinical practice guidelines on low-risk labour have been published in Japan, which suggests the need to reconsider the quality indicators.[25–27] This study, therefore, aimed to update the quality indicators for low-risk labour care drawing on new clinical practice guidelines and to propose an efficient way to update quality indicators.

METHODS

Indicator update

Review of the literature to assess whether any updating was necessary

We assessed whether the information in recently published clinical practice guidelines touching on low-risk labour care could affect the original quality indicators. We also made an abbreviated search of previously cited clinical practice guidelines to identify any new recommendations or information. Enough new information emerged that we decided that it was necessary to update the quality indicators for low-risk labour.

The update process

We used a modified Delphi method to integrate scientific evidence with expert panel opinions. This method has been widely used to develop quality indicators.[2 6 28–33] To develop a consensus, panel members individually rated candidate indicators in two rounds, through a postal questionnaire, with the second round used to resolve any issues emerging from round 1. In the second round, the rating results and all comments, including agreement and disagreement from the first round, were shared anonymously with the panel members. Disagreements were discussed by email or telephone between a facilitator (a researcher: KU) and the panel member to assure anonymity of comments. We used the consensus-based reporting standards for guideline-based performance measure development and re-evaluation proposed by the Guidelines International Network Performance Measures Working Group.[32]

Step 1: update of the definition of low-risk labour

The new clinical practice guidelines included a revised definition of low-risk labour. We, therefore, revised the definition of women having pregnancies with no special high-risk factors or complications (table 1).[25 26 34 35] In particular, we did not include 'mother treated for infertility' as a risk factor, because infertility treatment does not necessarily result in high-risk labour. However, we did include 'IVF' and 'after extensive infertility treatment' as high-risk factors because of the potential risk of complications associated. Panel members confirmed the definition used in clinical settings and were asked to approve the new definition at the panel meeting.

In Japan, most births are managed in hospitals (530 172, 54.3%) or private obstetric clinics (439 371, 45.0%). Very few take place in a midwifery-led unit (5968, 0.6%) or at home (1168, 0.1%).[23] In this study, low-risk labour care was, therefore, limited to care provided by a midwife with obstetrician approval in the late stages of pregnancy, and low-risk labour was defined as labour suitable for in-hospital midwifery with obstetrician approval in the late stages of pregnancy. Low-risk labour is expected to result in normal childbirth and excludes the high-risk factors listed above. Items related to abnormalities during labour or after delivery are excluded

Step 2: identification of new candidate indicators

To identify candidate indicators, we extracted existing recommendations from clinical practice guidelines relating to care for low-risk labour, and not limited to those published in Japan. Between November 2015 and January 2016, we searched for data sources using the terms 'pregnant women', 'mothers', 'infant', 'perinatal care', 'prenatal care', 'postnatal care', 'delivery', 'obstetric' and 'surgical procedure'. If the databases allowed searches by category,

Table 1 Categories for high-risk factors*

Category	High-risk factors
Physical findings	Age ≥40 years, body weight >80 kg before pregnancy, primiparas with body mass index >25 in the antepartum period.
Medical treatment for complication	Thyroid disease, connective tissue disorder, kidney disease, mental disorder, epilepsy, bronchial asthma, neurological disorder, blood-type incompatible pregnancy, haematological disease, heart disease, uterine cancer, Rhesus (D) alloimmunization in pregnancy, high blood pressure, hypertensive disorders of pregnancy, HIV positive, diabetes, gestational diabetes mellitus, antiphospholipid syndrome, pelvic fracture, placenta previa, pregnancy following conisation, premature birth, non-cephalic presentation after 36 weeks' gestation, multiple pregnancy, intrauterine growth retardation, pregnancy following myomectomy, high-grade cervical dysplasia, abdominal surgery other than caesarean section performed or planned during the pregnancy.
Pregnancy course	IVF, pregnancy after extensive fertility treatment, undergoing treatment for sexually transmitted disease, risk of mother-to-child transmission, two or fewer pregnancy check-ups, oligohydramnios, polyhydramnios, placenta previa because of previous caesarean section, received definitive diagnosis of fetal malformation or chromosomal abnormalities.
History of gynaecological diseases	Large uterine fibroids, postuterine surgery, caesarean section in previous delivery, placental abruption, underwent or plans to undergo abdominal surgery other than caesarean section, cervical incompetency, two or more spontaneous abortions, congenital disease, history of blood-type incompatible pregnancy, eclampsia/HELLP syndrome, gestational diabetes mellitus, stillbirth, neonatal death, delivery of infant <2500 g, severe gestational hypertension ≥160/110 mm Hg, history of delivering infant with major malformations.

*We defined low-risk labour as labour without any of these high-risk factors.
HELLP, Hemolytic anemia, Elevated liver enzymes, Low Platelet count.

these were used. We searched five databases for guidelines, the US Agency for Healthcare Research and Quality (AHRQ) National Guideline Clearinghouse, Guidelines International Network, Minds from the Japan Council for Quality Health Care, UK National Institute for Health and Care Excellence (NICE) and WHO Guidelines. Two databases, the AHRQ National Quality Measures Clearinghouse and NICE quality standards, were also searched for quality indicators. Several of these databases have inclusion criteria on development procedures, including AHRQ National Guideline Clearinghouse,[36] AHRQ National Quality Measures Clearinghouse,[37] WHO Handbook[38] and Minds.[39] All clinical practice guidelines and quality indicators obtained using these databases were therefore expected to fulfil certain methodological criteria.

We screened the literature to identify quality indicators and clinical practice guidelines published in English or Japanese between June 2012 and December 2015. We excluded any for which the title or abstract included the keywords '16 years old and younger', '40 years old and older', 'premature delivery', 'multiple births', 'breech presentation', 'prepregnancy obesity', 'pregnancy complication', 'obstetric history', 'abnormal pregnancy progress', 'infant congenital disease', 'diagnosis and treatment for infant disease', 'birth weight under 2000 g or over 4000 g', 'anaesthesia', 'operation and examination procedure', '28 or less weeks gestation', 'more than 1-month post partum', 'medical care provided only for high-risk labour' or 'does not apply in Japanese clinical practice'. We selected for recommendations on clinical practice by searching for: 'recommend', 'do not recommend' and 'suggest'. We did not include 'weak recommendation or suggestion', in line with the Grading Recommendations, Assessment, Development and Evaluations method.[40 41] The other exclusion criteria were 'content similar to previous candidate quality indicator', 'already basic standard in Japanese clinical practice', 'unmeasurable in Japanese clinical systems' or 'medical care provided only for high-risk labour'. When we found a revised clinical practice guideline which had been cited in the original development process, we confirmed which recommendations were new or updated, and considered whether the definition of the quality indicators, including the numerator and denominator, should be changed. We also considered the feasibility of data collection and the measurability of data.

Two researchers both extracted data independently. One assembled the candidate indicators and updated information about the original indicator set. This work was supervised by two senior researchers: a qualitative research specialist and an epidemiologist and guideline methodologist.

Step 3: convening the multidisciplinary panel

We aimed to construct a multidisciplinary panel of clinicians, public health specialists and lay mothers from across Japan. We selected obstetricians, paediatricians and midwives with5 or more years of clinical experience in supporting labour or neonatal care. We chose at least two panel members from each specialty to reduce bias in individual opinion. We recruited the participants in the previous panel in 2012, and added other candidates recommended by the original members. We obtained written informed consent from all panel members, and finally,

convened a multidisciplinary panel of obstetricians, paediatricians, midwives, public health specialists and mothers who were non-clinicians.

Step 4: panel ratings (rounds 1–2)

During each round, panel members rated the appropriateness of each candidate indicator on a nine-point Likert-type scale ranging from 1 'least suitable' to 9 'most suitable'. The panel also provided comments on the new indicators and the original 23 and suggested additional candidates. Based on criteria from the US National Quality Forum Measure Evaluation Criteria[42] and the American College of Cardiology/American Heart Association,[43] panel members rated the appropriateness of new quality indicators by: (1) usefulness in improving outcomes among mothers and infants; (2) whether the measure is clinically relevant; (3) validity; (4) reliability; (5) feasibility of measure implementation and (6) the overall assessment of the indicator (see online supplementary file 1).

In round 2, panel members were shown the anonymous results from round 1 and asked to evaluate any additional candidate indicators and those new indicators on which they had not reached an agreement in round 1. The candidate indicators and their descriptions were modified after discussion with panel members by email or telephone. We provided one or two reminders to any non-respondents. The process was completed in July 2016. All panel members approved the final list of updated indicators.

Adoption criteria

Candidate indicators were adopted if the median rating for 'overall assessment' during round 1 or round 2 was greater than 7, and if 3 or fewer of the 16 panel members rated it as less than 3.[44]

Patient involvement (involvement of mother)

The lay members of the panel were mothers. They all had to: (1) have given birth, (2) be interested in low-risk labour and midwife-led care and (3) not be clinicians. We first invited the mothers who had participated in the development of the original quality indicators[24] and added additional panel members recommended by the original members. Before asking them to agree to participate, we provided information about the aim, significance and methods of the study, the expected work and schedule of participation, and the advantages and burden of participating. After the mothers agreed to participate, we provided them with more information, for example,

Table 2 The composition of the multidisciplinary panel

	Specialty	Sex	Age	Professional environment/ location	Type of clinical or residential area	Clinical practice in birth centre for low-risk labour care
1	Paediatricians/neonatologist	Male	60s	Perinatal medical centre/ Osaka	Suburban	Experienced
2	Paediatricians/neonatologist	Male	50s	Clinic/Kyoto	Suburban	
3	Obstetrician	Female	40s	University/Kyoto	Urban	
4	Obstetrician	Male	40s	General hospital/Kyoto	Urban	
5	Obstetrician	Male	50s	Clinic/Tokyo	Urban	
6	Obstetrician	Male	40s	University hospital, perinatal medical centre/Nara	Suburban	Experienced
7	Midwife	Female	30s	Local perinatal medical centre/Saitama	Suburban	
8	Midwife	Female	40s	Perinatal medical centre/ Shizuoka	Suburban	Experienced
9	Midwife	Female	60's	University hospital, perinatal medical centre/Nara	Suburban	Experienced
10	Midwife	Female	40s	University hospital, perinatal medical centre/Nara	Suburban	Experienced
11	Midwife	Female	40s	University hospital, perinatal medical centre/Nara	Suburban	Experienced
12	Public health specialist	Female	40s	University/Osaka	Urban	
13	Public health specialist	Female	50s	University/Kyoto	Urban	
14	Non-clinician (mother, economist)	Female	50s	University, Faculty of Economist/Nara	Suburban	
15	Non-clinician (mother)	Female	30s	Association supporting mothers/Hyōgo	Suburban	
16	Non-clinician (mother)	Female	60s	Association supporting mothers/Osaka	Urban	

Those highlighted (numbers 1, 2, 3, 4, 7, 8, 12, 13 and 14) participated in the original indicator development process.

about quality in medicine, how to rate the quality of healthcare, high quality in healthcare, quality indicators and clinical practice guideline. The mothers participated fully as multidisciplinary panel members, on the same terms as the clinicians. They were also asked to propose new candidate indicators. They received no formal training in clinical guidelines and quality indicators, but we provided them with additional explanations and answered any questions face to face or by email or telephone both before and during the panel.

RESULTS

The multidisciplinary panel included 16 people, 13 clinicians (two paediatricians, four obstetricians, five midwives and two epidemiologists) and 3 mothers who were non-clinicians (table 2). All 16 members agreed to participate and responded to the postal questionnaires.

The literature review identified 27 clinical practice guidelines, with 277 recommendations, and 21 existing quality indicators (figure 1), which allowed us to develop

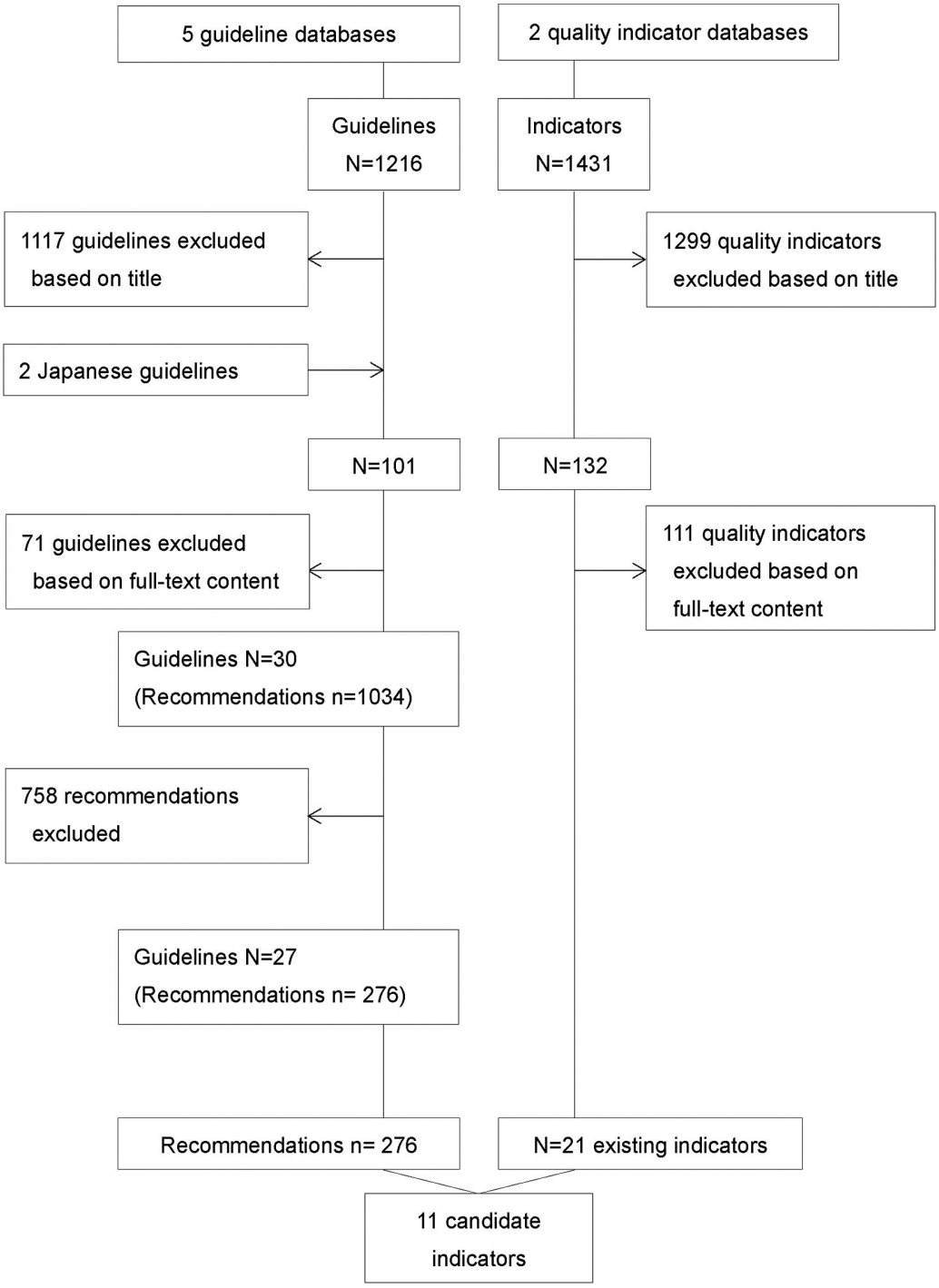

Figure 1 Overview of the literature review process: review of the guidelines and quality indicators extracted to generate candidate indicators.

Table 3 The list of new candidate indicators for low-risk labour care

Revised indicators	Direction for improvement	Rating result	Round 1		Round 2	
			Median	No of panel members rating the indicator less than 3	Median	No of panel members rating the indicator less than 3
1. Women receiving antibiotic prophylaxis during childbirth if maternal group B *Streptococcus* infections are identified at 33 to 37 weeks' gestation.	Higher	Adopted	8	0	–	–
2. Infants offered the necessary resuscitation in the first minutes after birth, evaluating their condition in line with the Japanese Neonatal Resuscitation Algorithm.	Higher	Adopted	8	1	–	–
3. Women receiving uterotonics for the prevention of postpartum haemorrhage during the third stage of labour.	Higher	Adopted	7	0	–	–
4. Women having early skin-to-skin contact with their babies if they wish, soon after birth in secure surroundings.	Higher	Adopted	8	0	–	–
5. Women planning spontaneous vaginal birth in a midwifery ward, and being able to follow that plan.	Higher	Adopted	7	0	–	–
6. Infants given formula supplementation without medical rationale from birth to discharge in term infants, even though mother intended to breast feed.	Lower	Adopted	8	1	–	–
7. Women having a fall during their hospitalisation.*	Lower	Adopted	8	1	–	–
8. Women and infants readmitted within 30 days of discharge.	Lower	Adopted	8	1	–	–
9. Women being screened for antenatal or postnatal depression using a validated questionnaire.	Higher	Adopted	8	1	–	–
10. Women and infants having complete medical records based on all quality indicators.	Higher	Adopted	8	0	–	–
11. Women having a review of their childbirth experience and support with the midwives and other staff who assisted at the birth.	Higher	Added and adopted	–	–	8	1
12. Women having been encouraged and supported to adopt the most comfortable positions throughout second stage labour.	Higher	Added and adopted	–	–	7.5	1
13. Women with non-medically indicated vaginal deliveries or non-medically indicated caesarean sections at greater than or equal to 37 and less than 39 weeks of gestation completed.	–	Not adopted	6	2	7	4

These indicators denote the frequency with which care was provided and recorded for women admitted to a midwifery ward.
'Higher' means that the quality of care in the facility is better when there are a high proportion of patients who received the intervention among the subject group who would benefit from it. 'Lower' means that the quality of care is better when there are a low proportion of patients with negative events among the group who should receive this care.
*An assessment of fall potential at the time of admission is widely used in Japan. The panel considered that the reason why falls seldom happen may be that this preliminary assessment helps to prevent falls. Their focus was on patient safety and they therefore wanted an indicator on this issue. But they also wanted to focus on patient outcomes ,not process. They, therefore, felt that the number of falls was an important indicator although it will not be very sensitive. They, therefore, agreed that this should be adopted as a quality indicator.

11 candidate indicators. We also considered the previous set of 23 indicators for low-risk labour care, all of which were included in the process in their existing form. This provided a set of 34 indicators for low-risk labour care for round 1 of the Delphi process. The panel adopted 10 of the 11 candidate indicators, excluding Indicator 13 (table 3). Two additional candidates (indicators 11 and 12, table 3) were proposed and adopted in round 2 (figure 2). All 23 of the original indicators were readopted (see online supplementary file 2), giving a final set of 35 quality indicators.

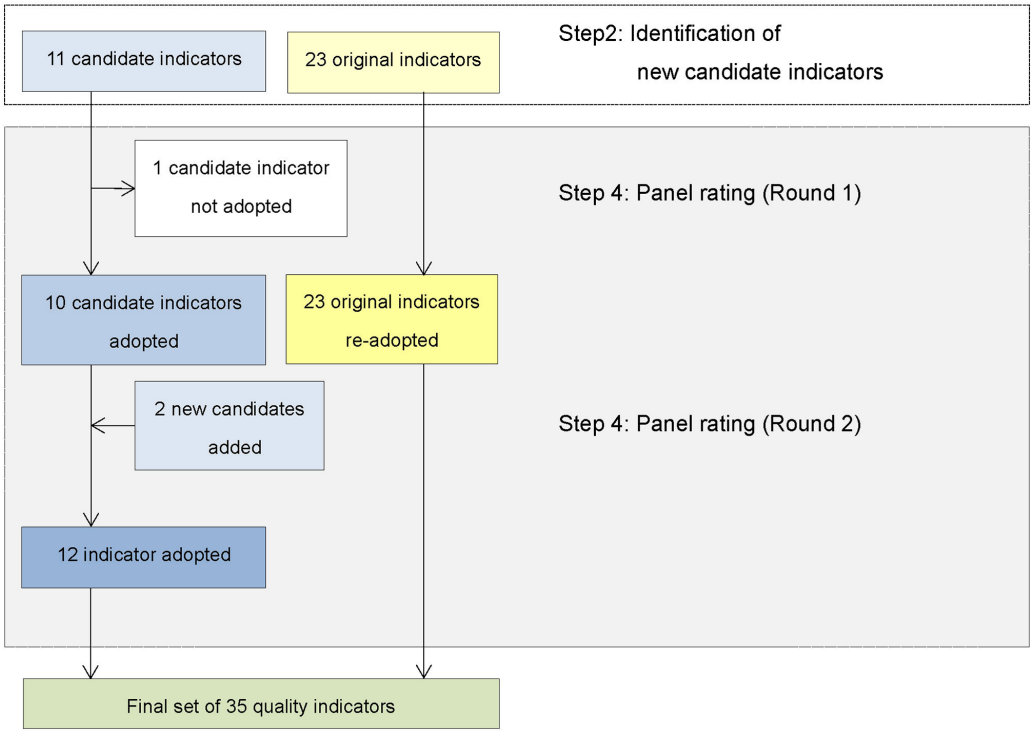

Figure 2 Process used to update quality indicators: a modified Delphi method for low-risk labour care.

DISCUSSION

Using criteria proposed for updating systematic reviews,[4] we decided that the quality indicators for low-risk labour care in Japan needed renewal. We updated them using the modified Delphi method, drawing on new or revised clinical practice guidelines and quality indicators, resulting in a set of 35 quality indicators (the original 23 plus 12 new). To our knowledge, this is the first article to describe a systematic and effective way to update quality indicators.

As new guidelines are developed or revised, it is important to consider whether quality indicators should also be updated. Few previous studies, however, have reported on this process.[14–16 18 19] In this study, we refined search strategies and selection criteria used for candidate indicators. First, we added 'not applicable for Japanese clinical practice' as an exclusion criterion. This allowed us to exclude, for example, some recommendations from WHO guidelines[45] that are primarily designed for use in resource-limited situations in low/middle-income countries. Second, in step 2, we were able to search efficiently by using almost the same strategies and databases as the original research. Third, there was no need to change the original 23 quality indicators once their validity had been checked. Fourth, we kept any unchanged recommendations in revised clinical practice guidelines that had been included in the indicator development work. The work to identify candidates was, therefore, easier than for the original quality indicator development. This strategy could, therefore, be as specific and efficient as updating systematic reviews.[4]

In selecting candidate indicators, we recognised that some widely used quality indicators for childbirth might not be applicable to low-risk labour. For example, 'women with non-medically indicated vaginal deliveries (induction labour) or non-medically indicated caesarean sections at greater than or equal to 37 and less than 39 weeks of gestation completed'[46] is widely used as an indicator for childbirth care. However, we did not adopt this because pregnant women may have planned normal vaginal delivery even with a clinical judgement of medical indications. They do not undergo caesarean section or induction labour before 39 weeks 0 day, unless an earlier delivery is necessary because of maternal or fetal indications. A pregnant woman planning normal vaginal delivery is diagnosed by an obstetrician as fit for low-risk vaginal delivery. If she shows signs of complications during her pregnancy, she will be treated immediately. Careful consideration is, therefore, needed of all potential quality indicators.

The panel composition might have affected the consensus process of developing and updating quality indicators. Few details, however, are available from previous studies.[6] This study, therefore, clarified the process used to recruit panel members, including defining the required composition and recruiting accordingly. As a result, of the 16 present members, nine had also participated in the original development. The original panel was also multidisciplinary and included various perinatal care providers and mothers. To make the panel construction process more efficient, we first asked the original panel members to participate, because they had a strong interest in improving quality in low-risk labour and were familiar with the whole process. In addition, we included seven new members to bring new experience and opinions. This enabled the panel to discuss care quality thoroughly from different viewpoints. Patients' experience will necessarily be individual. When recruiting mothers to participate in the study, we discussed the potential bias of panel members belonging to associations supporting

mothers. These associations may have preconceptual views on labour and therefore might influence the rating process through their members. However, these women can also draw on the experience of other mothers, which might improve the rating process. To avoid any imbalance, we ensured that at least one member had no relationship with any such association.

Incorporating patient views in discussions about quality and safety is considered important, but patient participation in development of quality indicators has been limited.[6 47–49] In this study, two new indicators ('Women having a review of their childbirth experience and support with the midwives and other staff who assisted at the birth' and 'Women having been encouraged and supported to adopt the most comfortable positions throughout second stage labour') were proposed by the mothers on the panel and were eventually adopted as new indicators. They were advanced indicators, designed to evaluate the mothers' decision-making process. A wide use of these quality indicators, including patients' views, could play a role in improving the quality of care for both mothers and healthcare providers.

Limitations

This study had some limitations. First, not all clinical practice guidelines are quality assured, for example, using the 'Appraisal of Guidelines and Research and Evaluation' (AGREEII) instrument.[8] However, the databases that were mainly used in this study apply their own quality criteria before they register any clinical practice guidelines.[36–38] Second, we did not have the funds or time to organise a face-to-face discussion and needed to add this list of indicators to the next test process. A face-to-face meeting means that all members can offer views, but it has the disadvantage that one member can influence others and affect the ranking. We, therefore, chose the modified Delphi method, rather than the RAND-modified Delphi method, which involves face-to-face meetings. Third, before the Delphi process, we did not have any process or sessions for the three mothers who were panel members to get to know each another and share views. Instead, we provided each of them with information on an individual basis, plus an opportunity to ask questions. However, it is possible that they might have felt more comfortable with the Delphi process, and made more contributions, if we had provided a session for them together in advance. Few quality indicator development processes have involved patients, and even those that have done so have not provided full descriptions of the patient involvement. The strength of this study is that laywomen were included, and a detailed description has been provided of the process. Fourth, despite our intention of developing a method that could be replicable in other clinical fields, there are likely to be topic-specific considerations. The basic conditions for its application would be the existence of solid and sufficient clinical practice guidelines, and guarantee of their quality. With this in place, however, we believe that we have put forward a suitable procedure for how and when to update quality indicators.

CONCLUSIONS

Through a transparent process, we have updated the set of quality indicators for low-risk labour care in Japan, using a modified Delphi method, and drawing on new clinical practice guidelines and quality indicators. As a next step, we are testing the indicators and will revise the list to improve monitoring and help ensure the provision of safe, high-quality maternity and labour care.

Acknowledgements We are grateful to Hiroyuki Kitajima, Isao Kuwahara, Miho Egawa, Fumitomo Nishimura, Masato Takeuchi, Toshiyuki Sado, Yoshimi Miyaki, Kazue Watanabe, Sachie Nishi, Tomoko Terai, Sachi Nakagami, Makiko Suga, Keiko Miyamoto, Yoko Chiba, Chie Kouta and Masako Senda for their valuable insight and comments. All of them permit us to mention their name. We thank Melissa Leffler from Edanz Group (www.edanzediting.com/ac) for editing a draft of this manuscript.

Contributors KU, MK, SO and TN planned the study. KU and MK independently extracted evidence and TN supervised the process. KU, MK, SO and TN prepared and carried out the modified Delphi procedure. KU conducted the data analyses and prepared the manuscript. SO and TN critically commented on the manuscript. All authors read and approved the final manuscript.

Funding This work was supported by the Ministry of Education, Culture, Sports, Science and Technology, (Grant-in-Aid for Young Scientists B, 26870745, 2014-2017) and Health and Labor Sciences Research Grant.

Competing interests None declared.

Patient consent for publication Not required.

Ethics approval This study was approved by the Ethics Committee of Kyoto University Graduate School and Faculty of Medicine.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement No additional data are available.

REFERENCES

1. Institute of Medicine Committee. *Institute of medicine committee to design a strategy for quality, review assurance in medicare*. Washington, DC: National Academies Press, 1990.
2. Boyce MB, Browne JP, Greenhalgh J. The experiences of professionals with using information from patient-reported outcome measures to improve the quality of healthcare: a systematic review of qualitative research. *BMJ Qual Saf* 2014;23:508–18.
3. Higashi T, Shekelle PG, Adams JL, *et al*. Quality of care is associated with survival in vulnerable older patients. *Ann Intern Med* 2005;143:274–81.
4. Garner P, Hopewell S, Chandler J, *et al*. When and how to update systematic reviews: consensus and checklist. *BMJ* 2016;354:i3507.
5. Bonfill X, Roqué M, Aller MB, *et al*. Development of quality of care indicators from systematic reviews: the case of hospital delivery. *Implement Sci* 2013;8:42.
6. Kötter T, Blozik E, Scherer M. Methods for the guideline-based development of quality indicators--a systematic review. *Implement Sci* 2012;7:21.
7. Shekelle PG, Ortiz E, Rhodes S, *et al*. Validity of the agency for healthcare research and quality clinical practice guidelines: how quickly do guidelines become outdated? *JAMA* 2001;286:1461–7.
8. Brouwers MC, Kho ME, Browman GP, *et al*. AGREE II: advancing guideline development, reporting and evaluation in health care. *CMAJ* 2010;182:E839–42.
9. Venkatesh AK, Savage D, Sandefur B, *et al*. Systematic review of emergency medicine clinical practice guidelines: Implications for research and policy. *PLoS One* 2017;12:e0178456.
10. Antman EM, Lau J, Kupelnick B, *et al*. A comparison of results of meta-analyses of randomized control trials and recommendations

of clinical experts. Treatments for myocardial infarction. *JAMA* 1992;268:240–8.
11. Wright JD, Pawar N, Gonzalez JS, *et al*. Scientific evidence underlying the American College of Obstetricians and Gynecologists' practice bulletins. *Obstet Gynecol* 2011;118:505–12.
12. Jacobs AK, Anderson JL, Halperin JL, *et al*. The evolution and future of ACC/AHA clinical practice guidelines: a 30-year journey: a report of the American College of Cardiology/American Heart Association Task Force on practice guidelines. *Circulation* 2014;130:1208–17.
13. Saver BG, Martin SA, Adler RN, *et al*. Care that matters: quality measurement and health care. *PLoS Med* 2015;12:e1001902.
14. Thomas RJ, King M, Lui K, *et al*. AACVPR/ACCF/AHA 2010 Update: Performance Measures on Cardiac Rehabilitation for Referral to Cardiac Rehabilitation/Secondary Prevention Services Endorsed by the American College of Chest Physicians, the American College of Sports Medicine, the American Physical Therapy Association, the Canadian Association of Cardiac Rehabilitation, the Clinical Exercise Physiology Association, the European Association for Cardiovascular Prevention and Rehabilitation, the Inter-American Heart Foundation, the National Association of Clinical Nurse Specialists, the Preventive Cardiovascular Nurses Association, and the Society of Thoracic Surgeons. *J Am Coll Cardiol* 2010;56:1159–67.
15. Drozda JP, Ferguson TB, Jneid H, *et al*. 2015 ACC/AHA Focused Update of Secondary Prevention Lipid Performance Measures: A Report of the American College of Cardiology/American Heart Association Task Force on Performance Measures. *J Am Coll Cardiol* 2016;67:558–87.
16. Sanders AE, Nininger J, Absher J, *et al*. Quality improvement in neurology: Dementia management quality measurement set update. *Neurology* 2017;88:1951–7.
17. Weled BJ, Adzhigirey LA, Hodgman TM, *et al*. Critical care delivery: the importance of process of care and icu structure to improved outcomes: an update from the american college of critical care medicine task force on models of critical care. *Crit Care Med* 2015;43:1520–5.
18. Wiler JL, Welch S, Pines J, *et al*. Emergency department performance measures updates: proceedings of the 2014 emergency department benchmarking alliance consensus summit. *Acad Emerg Med* 2015;22:542–53.
19. Carinci F, Van Gool K, Mainz J, *et al*. Towards actionable international comparisons of health system performance: expert revision of the OECD framework and quality indicators. *Int J Qual Health Care* 2015;27:137–46.
20. Kleefstra SM, Zandbelt LC, de Haes HJ, *et al*. Trends in patient satisfaction in Dutch university medical centers: room for improvement for all. *BMC Health Serv Res* 2015;15:112.
21. Grol RBR, Moss F. Quality improvement research: understanding the science of change in health care. *BMJ Publish Group* 2004.
22. Ministry of Health, Labour and Welfare, Japan. *Vital statistics of Japan 2015*: National Statistics Center, Statistics Japan, 2016.
23. Japanese Nurse Association. Chapter 5 Risk scores and medical institutions for parturient. In: *Survey on temporary relocation system of midwife and acceptability of midwifery practice etc in 2012*, 2014.
24. 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.
25. Japan Society of Obstetrics and Gynecology and Japan Association of Obstetricians and Gynecologists. Guidelines for obstetrical practice in Japan. 2014. http://www.jsog.or.jp/activity/pdf/gl_sanka_2014.pdf (Accessed 21 Jan 2018).
26. Japanese Midwives Assosiation. Midwife Clinical Practice Guidelines. 2014. http://www.midwife.or.jp/pdf/guideline/guideline.pdf (Accessed 21 Jan 2018).
27. General Incorporated Assosiation Japan Acadeny of Midwifery. 2012 Evidence-based guidelines for midwifery care during childbirth. 2012. http://square.umin.ac.jp/~jam/docs/ebm_guideline_childbirth2012.pdf (Accessed 21 Jan 2018).
28. 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.
29. 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.
30. Higashi T. Lessons learned in the development of process quality indicators for cancer care in Japan. *Biopsychosoc Med* 2010;4:14.
31. Mukai H, Higashi T, Sasaki M, *et al*. Quality evaluation of medical care for breast cancer in Japan. *Int J Qual Health Care* 2016;28:110–3.
32. Nothacker M, Stokes T, Shaw B, *et al*. Reporting standards for guideline-based performance measures. *Implement Sci* 2016;11:6.
33. Murphy MK, Black NA, Lamping DL, *et al*. Consensus development methods, and their use in clinical guideline development. *Health Technol Assess* 1998;2:1–88.
34. Nakabayashi M. Study on safety steps in obstetrics department. Health Labour Science Research Grant Research on Health Technology Assessment, The summary and sharing documentation in 2004. 2004:1–16.
35. Japan Society of Obstetrics and Gynecology. Social insurance ABC for obstetricians. 2016.
36. AHRQ National Guideline Clearinghouse. Inclusion criteria. 2014. https://www.guideline.gov/help-and-about/summaries/inclusion-criteria (Accessed 21 Jan 2018).
37. AHRQ National Quality Measures Clearinghouse. Inclusion Criteria. 2017. https://qualitymeasures.ahrq.gov/help-and-about/summaries/inclusion-criteria (Accessed 21 Jan 2018).
38. World Health Organization. *WHO Handbook for Guideline Development*. 2nd edn, 2014.
39. Brouwers MC, Kho ME, Browman GP, *et al*. AGREE II: advancing guideline development, reporting and evaluation in health care. *J Clin Epidemiol* 2010;63:1308–11.
40. Andrews JC, Schünemann HJ, Oxman AD, *et al*. GRADE guidelines: 15. Going from evidence to recommendation-determinants of a recommendation's direction and strength. *J Clin Epidemiol* 2013;66:726–35.
41. Guyatt GH, Oxman AD, Vist GE, *et al*. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ* 2008;336:924–6.
42. National Quality Forum. Measure evaluation criteria and guidance for evaluating measures for endorsement. 2016. http://www.qualityforum.org/Measuring_Performance/Submitting_Standards/Measure_Evaluation_Criteria.aspx (Accessed 21 Jan 2018).
43. 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. *Circulation* 2005;111:1703–12.
44. Fitch K, Burnand B, LaCalle JR, *et al*. *The RAND/UCLA appropriateness method user's manual*: RAND Corporation, 2001.
45. World Health Organization. Guidelines on maternal, newborn, child and adolescent health Listed by year of publication. http://www.who.int/maternal_child_adolescent/guidelines/guidelines-year/en/ (Accessed 21 Jan 2018).
46. Joint Commission. *Perinatal care: percentage of patients with elective vaginal deliveries or elective cesarean births at greater than or equal to 37 and less than 39 weeks of gestation completed*: United States Agency for Healthcare Research and Quality (AHRQ) National Guideline Clearinghouse, 2017.
47. Kerr EA, Hayward RA. Patient-centered performance management: enhancing value for patients and health care systems. *JAMA* 2013;310:137–8.
48. den Breejen EM, Nelen WL, Schol SF, *et al*. Development of guideline-based indicators for patient-centredness in fertility care: what patients add. *Hum Reprod* 2013;28:987–96.
49. Medical Information Network Distribution Service. The method of evaluation and selection for clinical practice guideline. 2016. http://minds.jcqhc.or.jp/s/about_us_evaluation_method (Accessed 10 Jan 2018).