Received: 29 November 2017 | Revised: 6 February 2018 | Accepted: 13 February 2018

DOI: 10.1111/cas.13549

REVIEW ARTICLE

WILEY Cancer Science

Frequency of unsatisfactory cervical cytology smears in cancer screening of Japanese women: A systematic review and meta-analysis

Satoyo Hosono[1,2] | Teruhiko Terasawa[3] | Takafumi Katayama[4] | Seiju Sasaki[5] | Keika Hoshi[6] | Chisato Hamashima[7]

[1]Department of Public Health, Nagoya City University Graduate School of Medical Sciences, Nagoya,Japan

[2]All Japan Labour Welfare Foundation, Tokai Clinic, Nagoya,Japan

[3]Section of General Internal Medicine, Department of Emergency and General Internal Medicine, Fujita Health University School of Medicine, Toyoake,Japan

[4]Department of Statistics and Computer Science, College of Nursing Art and Science, University of Hyogo, Hyogo,Japan

[5]St. Luke's International Hospital Affiliated Clinic, Center for Preventive Medicine, Tokyo,Japan

[6]Department of Hygiene, Kitasato University School of Medicine, Sagamihara, Japan

[7]Division of Cancer Screening Assessment and Management, Center for Public Health Sciences, National Cancer Center, Tokyo, Japan

Correspondence
Satoyo Hosono, Department of Public Health, Nagoya City University Graduate School of Medical Sciences, Mizuho-ku, Nagoya, Japan.
Email: hosochan@gmail.com

Funding Information
National Cancer Center, Tokyo, Japan

The Bethesda system (TBS) has been used for cervical cytological diagnosis in Japan since 2008. Evaluation of specimen adequacy is the most important aspect of quality assurance and for precise diagnosis in TBS. A systematic review and meta-analysis were carried out to assess the unsatisfactory specimen rate in the primary cervical cancer screening setting in Japan. Ovid Medline and Ichushi-Web databases were searched from inception through to May 2017. Prospective and retrospective studies that reported the proportion of unsatisfactory specimens in healthy asymptomatic Japanese women in a cervical cancer screening program were eligible for inclusion; 17 studies were included in the meta-analysis. The random-effects model meta-analysis calculated summary estimates of the unsatisfactory rate of 0.60% (95% confidence interval [CI], 0.18-1.96%; I^2 = 99%) for conventional cytology and 0.04% (95% CI, 0.00-0.35%; I^2 = 99%) for liquid-based cytology (LBC). However, comparative results between conventional and liquid-based cytology, based on four direct and nine comparative studies, showed no significant difference (summary odds ratio = 3.5×10^{-2} favoring LBC [95% CI, 6.9×10^{-4}-1.7]; I^2 = 98%). In the subgroup analyses and meta-regressions, use of non-cotton devices for conventional cytology and use of a particular platform for LBC were associated with lower unsatisfactory rates. Meta-regression also suggested chronological improvement in unsatisfactory rates for both tests. In Japanese cervical cancer screening programs, conventional cytology remains prevalent. Future research needs to focus on evaluating the impact of screening programs using LBC by comparing the accuracy, performance, and cost-effectiveness with conventional cytology in the Japanese population.

KEYWORDS
Bethesda system, cancer screening, cervical cancer, meta-analysis, unsatisfactory specimen rate

1 | INTRODUCTION

In Japan, cervical cytology has been used widely for primary screening of cervical cancer since 1983. Cytological diagnosis was initially carried out using the Nichibo classification created by the Japan Association of Obstetricians and Gynecologists in 1978.[1] The Nichibo classification categorizes cervical lesions into five stages from class I to class V according to cellular atypia. The Nichibo

classification had several issues; mainly, it did not define the management of human papillomavirus (HPV)-infected cells and the handling of cells due to small amounts in samples. To solve these problems, in 2008, the Japan Association of Obstetricians and Gynecologists formally adopted the Bethesda System (TBS) for reporting cervical cytological diagnoses.[2] The Bethesda System included specific statements about specimen adequacy, general categorization, and interpretation and results.[3] Unsatisfactory specimens may cause screening errors or interpretation errors. Therefore, evaluation of specimen adequacy is considered to be the most important aspect of the quality assurance component and precise diagnosis in TBS. Generally, the rate of unsatisfactory slides using liquid-based cytology (LBC) by other non-cotton swab collection devices was lower than in conventional cytology (CC).[4-7] Since adoption of TBS, Japanese gynecologists and cytologists have worked to reduce the proportion of unsatisfactory specimens. Although CC has been prevalent in Japanese primary cervical screening, LBC has been introduced to screening programs since the adoption of TBS in 2008, without definitive evidence of accuracy and performance.

In 2001, the research group for cancer screening founded by the Ministry of Health and Welfare of Japan recommended five cancer screening programs, including Pap smears for cervical cancer.[8] In 2010, our previous research group developed the guideline for cervical cancer screening and recommended CC and LBC for population-based and opportunistic cervical cancer screening. However, neither HPV testing alone nor a combination of two methods was recommended due to insufficient evidence.[9] The quality assurance system, target age group, screening interval, unsatisfactory samples, and test accuracy remained as important issues to consider.

Since publication of the previous guidelines, new studies regarding cervical cancer screening have been reported. In the framework of the Japanese research group for the development of cervical cancer screening guidelines, a systematic review and a meta-analysis were carried out to assess the unsatisfactory specimen rate in the primary cervical cancer screening programs in Japan. Furthermore, CC and LBC were compared by sample collection devices.

2 | MATERIALS AND METHODS

To update the evidence for the effectiveness of the Japanese screening programs and to confirm the best available method for cervical cancer screening in Japan, a systematic review and meta-analysis were carried out. Comparison of the unsatisfactory specimen rate between CC and LBC was listed as a clinical question by our research group. An unsatisfactory test is a quality assurance issue and associated with potential harm, because women with unsatisfactory results require repeat testing, which is inconvenient and may lead to anxiety. Therefore, this manuscript is a focused, in-depth systematic review and meta-analysis regarding unsatisfactory specimens in the Japanese primary cervical cancer screening program. There was no protocol specifically designed for this focused review.

2.1 | Data sources and searches

Ovid Medline and Ichushi-Web (the Japan Medical Abstracts Society) were searched from inception through to 31 October, 2016 (Ovid Medline) or 31 May, 2017 (Ichushi-Web) using search terms including "cervical cancer," "*Papanicolaou*," "cytology," and their synonyms. The languages of publications were restricted to English and Japanese. The search was supplemented by examining the references of eligible studies. Document S1 provides the exact search strategies.

2.2 | Study selection

Six single reviewers (TT, CH, SH, TK, SS, and KH) screened non-overlapping sets of abstracts. Two investigators (TT and SH) independently examined the full text of each retrieved publication for eligibility. Eligible studies included prospective and retrospective studies that reported the proportion of unsatisfactory specimens in healthy asymptomatic Japanese women (aged ≥20 years) who participated in a cervical cancer screening program. Only studies that evaluated cervical cytology smears based on the 2001 TBS were included.[2,3] Full details of the inclusion criteria and their operational definitions are available in Document S2. Discrepant results were resolved by consensus between two investigators (TT and SH). Unresolved discrepancies were adjudicated in a research meeting involving six reviewers.

2.3 | Data extraction

One reviewer (TT) extracted descriptive data from each eligible paper; another (SH) verified all of the data. Information on study, subject, and test characteristics were extracted. Studies that had a paired design, in which samples from all participants were assessed with both CC and LBC tests, were operationally defined as direct comparative studies, and the proportion of unsatisfactory cervical smears of both tests was directly compared. In contrast, studies were categorized as indirect comparative studies if a subject underwent one cytology test only (i.e. either CC or LBC), and the proportions of samples with unsatisfactory results were compared between the two tests. Document S2 provides details of the extracted items.

The primary outcome of interest was the proportion of unsatisfactory cervical cytology smears using TBS 2001.[3] The total number of examined samples and the number of samples categorized as unsatisfactory were extracted. A single reviewer (CH or TT) extracted the pertinent quantitative data; one other investigator (TT or SH) confirmed the data.

2.4 | Assessment of study validity

No established assessment tools were appropriate for assessing the risk of bias and the applicability of studies that focused on the frequency of non-contributory test results. Therefore, it was planned to assess the implementation and reporting of the selection of participants, collection and preparation of smear samples, and evaluation of results. However, poor reporting precluded detailed analyses of

these items. Therefore, study validity regarding comparisons between CC and LBC was only assessed descriptively.

2.5 | Data synthesis

The proportion of unsatisfactory cervical smears for each study was first calculated as the number of unsatisfactory results divided by the total number of examined samples, along with their exact 95% confidence intervals (CIs). The summary proportions were then estimated by undertaken a random-effects binomial logistic regression when there were three or more studies.[10] For studies that reported the results of both CC and LBC, the proportions of unsatisfactory cervical smears were compared by estimating odds ratios (ORs) as the relative effect measure, because unsatisfactory smears were rare. A random-effects conditional logistic regression was used to calculate the summary ORs and their 95% CIs without zero-event corrections.[10] We calculated the I^2 statistics with their 95% CIs, and considered I^2 of >50% or >75% to be suggestive of intermediate or high heterogeneity, respectively.[11,12] Subgroup analyses and univariate meta-regression were carried out for sample collection devices (cotton swab vs other non-cotton swab collection devices combined) and commercial platforms for LBC (SurePath [Becton Dickinson, Franklin Lakes, NJ, USA] vs all other platforms). Meta-regressions were also performed to model the first year of the research period for each study as an effect modifier covariate. Tests for funnel plot asymmetry were not applied because these methods have not been established for studies reporting proportions as the outcome measures. Complete details on the statistical methods used are described in Document S2. All analyses were carried out using Stata SE, version 14.1 (Stata, College Station, TX, USA). *P*-values for all comparisons were two-tailed, and significance was defined as $P < .05$.

3 | RESULTS

3.1 | Literature flow and eligible studies

The searches of published work identified 28 potentially eligible publications (Fig. 1). After full-text screening, 11 publications were excluded: five studies that did not adopt TBS, two studies that evaluated partially overlapping participants, one that assessed data from unclear settings, one that assessed data from hospital settings, one that assessed data from laboratory settings, and one review without original data. Finally, 17 studies (three direct,[13-15] eight indirect,[16-23] and one both[24] comparison studies between CC and LBC, and five studies assessing CC only[25-29]) including 1 497 451 unique healthy women were eligible.

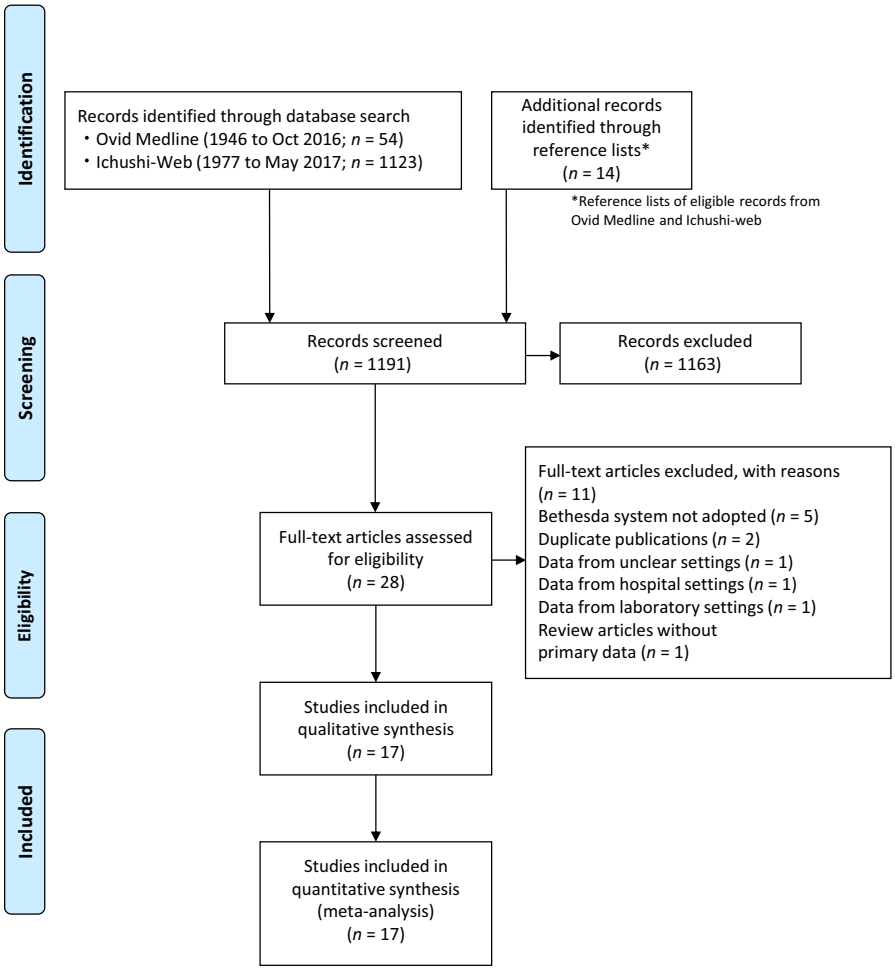

FIGURE 1 Flow diagram of this systematic review of published reports and meta-analysis of the frequency of unsatisfactory cervical cytology smears in Japanese women

Study, clinical, and test characteristics. Overall, the study designs were not reported in sufficient detail (Table 1). No studies explicitly reported whether participants were prospectively enrolled or whether data were retrospectively assessed. Only three studies reported that the investigators consecutively evaluated the smear samples from a group of eligible subjects;[14,15,25] the other 14 studies did not describe the consecutiveness of the assessed smear results. Eight studies were based on a single center's experience, whereas the other seven studies collected and assessed the pertinent data from multiple institutions (two studies not described). The sample size varied and ranged from 384 to 615 231. Only five studies reported the average age of study participants,[13-15,20,22] which ranged from 39 to 53 years. The reported proportion of screening positives, defined as "atypical squamous cells of undetermined significance " or worse results, ranged between 1.5% and 5.3%; invasive cancer was in general reported in 0.1% or fewer subjects.

Test characteristics and diagnostic criteria are presented in Table 2. Regarding the collection devices used, most studies that assessed conventional smears used cotton swabs, whereas the CervexBrush (Rovers Medical Devices, Oss, Netherlands) was the most commonly used device in studies of LBC. SurePath was the most commonly assessed LBC platform; two other studies from a single institution evaluated TACAS Pro (Medical & Biological Laboratories, Nagoya, Japan), and only one study used ThinPrep (Hologic Inc., Marlborough, MA, USA). No

TABLE 1 Characteristics of studies reporting unsatisfactory specimens in healthy asymptomatic Japanese women (aged ≥20 years) who participated in a cervical cancer screening program

Study ID (location)[reference no.]	Study years	No. of institutions	Sample size	Average age, years (range)	≥ASCUS / ≥HSIL / cancer, %
Direct comparison studies of LBC vs CC					
Ashikawa 2014 (Tokyo)[13]	2011	Single center	363	48 (26-82)	4.7/2.5/0 (CC); 4.9/2.6/0 (LBC)
Kuramoto 2015A (Kanagawa)[14]	2013-2014	Single center	3483	50 (SD = 13)	1.8/0.1/0.1 (CC); 1.6/0.1/0.0 (LBC)
Kuramoto 2015B (Kanagawa)[15]	2014	Single center	516	39 (SD = 23)	2.3/0.4/0 (CC); 2.1/0.4/0 (LBC)
Furutate 2016 (Saitama)[24]	2014 (both); 2010-2014 (CC)[a]; 2015 (LBC) [a]	Single center	207 (both); 57 790 (CC) [a]; 747 (LBC) [a]	ND	5.3/1.0/0 (CC); 4.3/1.0/0 (LBC)
Indirect comparison studies of LBC vs CC					
Akamatsu 2005 (Niigata)[16]	2003	Single center	923 (CC); 17 049 (LBC)	ND	ND/0.1/0 (CC); ND/0.2/0 (LBC)
Akamatsu 2008 (Niigata)[17]	2005-2006	3 centers	26 644 (CC); 50 032 (LBC)	ND	ND/0.6/0.2 (CC)[b]; ND/0.7/0.3 (LBC) [b]
Fujii 2012 (Tottori)[18]	2010	31 centers and mobile medical vehicles	12 973 (CC); 8094 (LBC)	ND	1.5/0.4/0.1 (total)
Kamei 2012 (Kanagawa)[19]	2010 (CC); 2010-2011 (LBC)	Single center	1835 (CC); 1796 (LBC)	ND	1.8/0.4/0 (CC); 3.9/0.5/0 (LBC)
Tachibana 2013 (Chiba)[20]	2006-2012 (CC); 2011-2012 (LBC)	Single center	609 297 (CC); 5934 (LBC)	53 (20-88)	ND/ND/ND (CC); 3.4/0.7/0.1 (LBC)
Kato 2016 (Ibaraki)[21]	2011-2012 (CC); 2013-2014 (LBC)	ND	191 796 (CC); 191 741 (LBC)	ND	2.1/ND/0.0 (CC); 2.5/ND/0.0 (LBC)
Kuroshima 2016 (Okinawa)[22]	2011-2012 (CC); 2013-2014 (LBC)	ND	45 621 (CC); 45 129 (LBC)	50 (17-100)	1.6/0.3/0.0 (CC); 3.3/0.6/0.0 (LBC)
Kuwakubo 2016 (Tochigi)[23]	2012 (CC); 2013 (LBC)	Multicenter	12 063 (CC); 12 486 (LBC)	ND	2.5/0.3/0.1 (CC); 3.1/0.4/0.1 (LBC)
Non-comparative studies of CC					
Kuramoto 2009 (Kanagawa)[25]	2009	Single center	1967	ND	4.0/0/0
Shirayama 2011 (Tokyo)[26]	2008	46 centers	1273	ND	ND
Takano 2011 (Chiba)[27]	2009-2010	18 centers and mobile medical vehicles	15 514	ND	ND
Morimura 2012 (Fukushima)[28]	2008	Multicenter	69 584	ND	ND
Morimura 2013 (Fukushima)[29]	2009-2011	114 centers	108 025	45-49 (ND)	ND

ASCUS, atypical squamous cells of undetermined significance; CC, conventional cytology; HSIL, high-grade squamous intraepithelial lesion; LBC, liquid-based cytology; ND, no data; SD, standard deviation.

[a]Data pertaining to indirect comparison also reported.

[b]Based on the whole study population, not from the subjects from whose date unsatisfactory rates were calculated.

TABLE 2 Test characteristics and diagnostic criteria included in studies of cervical cancer screening programs in Japan

Study ID[reference no.]	Sample collection device (CC)	Sample collection device (LBC)	LBC platform	Cytological classification system
Direct comparison studies of LBC vs CC				
Ashikawa 2014[13]	Cervex-Brush	Cervex-Brush	SurePath	"Bethesda 2001"
Kuramoto 2015A[14]	Cervex-Brush	Cervex-Brush	TACAS Pro	"Bethesda 2001"
Kuramoto 2015B[15]	Cytopick α	Cytopick α	TACAS Pro	"Bethesda 2001"
Furutate 2016[24]	Cervex-Brush (ND for CC only)	Cervex-Brush (ND for LBC only)	ThinPrep	"Bethesda 2001"
Indirect comparison studies of LBC vs CC				
Akamatsu 2005[16]	Cotton swab	Cotton swab or Cervex-Brush	SurePath	Squamous cellularity, % of obscure field, and endocervical transformation zone component by Bethesda 2001 and Bethesda 1991[a]
Akamatsu 2008[17]	Cotton swab or spatula	Cervex-Brush, cotton swab, spatula, or others	SurePath	Squamous cellularity, % of obscure field, and endocervical transformation zone component by Bethesda 2001 and Bethesda 1991 [a]
Fujii 2012[18]	Cotton swab	Cervex-Brush	SurePath	Squamous cellularity and % of obscure field by Bethesda 2001
Kamei 2012[19]	Cotton swab	Cervex-Brush	SurePath	Squamous cellularity by Bethesda 2001
Tachibana 2013[20]	ND	Cervex-Brush	SurePath	"Bethesda system"
Kato 2016[21]	Cotton swab, brush, spatula, or others	Cotton swab, brush, spatula, or others	SurePath	"Bethesda system"
Kuroshima 2016[22]	Mostly cotton swab	Cervex-Brush, EndoCervex-Brush, or Cervex-Brush Combi	SurePath	"Bethesda 2001"
Kuwakubo 2016[23]	ND	ND	SurePath	"Bethesda 2001"
Non-comparative studies of CC				
Kuramoto 2009[25]	Cotton swab and cytobrush	NA	NA	"Bethesda 2001"
Shirayama 2011[26]	Cotton swab, brush, or cytopick	NA	NA	Squamous cellularity and % of obscure field by Bethesda 2001 [b]
Takano 2011[27]	ND	NA	NA	"Bethesda 2001"
Morimura 2012[28]	ND	NA	NA	Squamous cellularity by Bethesda 2001
Morimura 2013[29]	Cotton swab, spatula, or brush	NA	NA	"Bethesda 2001"

CC, conventional cytology; LBC, liquid-based cytology; NA, not applicable; ND, no data.
[a]Subgroup data based on squamous cellularity and percentage of obscure field by Bethesda 2001 only are also reported.
[b]No observed atypical cells were also required. Collection devices: CervexBrush (Rovers Medical Devices, Oss, Netherlands); SurePath (Becton Dickinson, Franklin Lakes, NJ, USA); TACAS pro (Medical & Biological Laboratories, Nagoya, Japan); ThinPrep (Hologic Inc., Marlborough, MA, USA).

studies reported in sufficient detail on the collectors of the cervical smear samples or who screened or evaluated the unsatisfactory results.

3.2 | Assessment of study validity in comparative studies

In Table 2, four direct comparison studies compared CC and LBC in the same participants, implying that the subjects evaluated by the two tests were completely comparable. Thus, this design should have prevented differences in the distribution of other factors that may affect the unsatisfactory smear results. However, after a cervical sample collection, the investigators first smeared a portion of the sample onto a microscope slide for the conventional test, and they then rinsed the rest of the collected sample on the collection device into the fluid medium for the LBC. This "two-step" procedure might introduce bias, because the first step (smearing) could affect the quality and/or quantity of the rest of the sample still on the collection device.

In contrast, nine indirect comparison studies evaluated unsatisfactory smears in two completely independent subpopulations, thus, the comparability of other factors in the subjects tested was not assured. For example, cervical samples for conventional tests were typically collected with cotton swabs, whereas most samples (308 755/337 577; 91.5%) for LBC were collected with the Cervex-Brush and then assessed with the SurePath platform.

3.3 | Conventional cytology

For CC, a total of 17 studies involving 1 159 874 healthy women reported a wide ranging proportion of unsatisfactory smear results, from 0.00% to 18.09% (Fig. 2). The random-effects model meta-analysis

calculated a summary estimate of the unsatisfactory specimen rate of 0.60% (95% CI, 0.18-1.96%; I^2 = 99%). In the subgroup analyses for specific collection devices (Fig. 2), although the summary proportion was higher for cotton swabs (summary estimate = 2.43% [95% CI, 0.55-10.06%; I^2 = 100%]) than for collection devices other than non-cotton swabs combined (summary estimate = 0.31% [95% CI, 0.05-2.03%; I^2 = 100%]), there was no significant difference between the two collection device groups (OR = 0.13 [95% CI, 0.01-1.45]; P = .10). Meta-regression showed that the first year of the study period was the only factor significantly associated with decreased proportions of unsatisfactory results (per-year decrease = 0.49% [95% CI, 0.40-0.58%]; P < .001).

3.4 | Liquid-based cytology

A total of 12 studies including 337 577 healthy women reported the proportion of unsatisfactory smear results, which ranged from 0.00% to 8.3% (Fig. 3). Of 12 studies, five reported no unsatisfactory results. The summary proportion was 0.04% (95% CI, 0.00-0.35%; I^2 = 99%). A subgroup analysis for the largest group assessed with a consistent test method (i.e. the samples were collected with the CervexBrush and assessed with the SurePath platform; 91.5% of all LBC tests) had a summary estimate of 0.02% (95% CI, 0.00-0.18%; I^2 = 100%) (Fig. 3). Although no specific collection device on average was associated with a higher or lower proportion of unsatisfactory results (OR = 0.97 [95% CI, 0.00-6.12]; P = .27 for cotton swabs vs non-cotton swab devices combined), use of the SurePath platform, compared with the two other platforms combined, was significantly associated with lower proportions of unsatisfactory smears (OR = 0.012 [95% CI, 0.00-0.87]; P = .043). Again, the first year of the study period was significantly associated with a lower proportion (per-year decrease = 0.47% [95% CI, 0.26-0.68%]; P < .001).

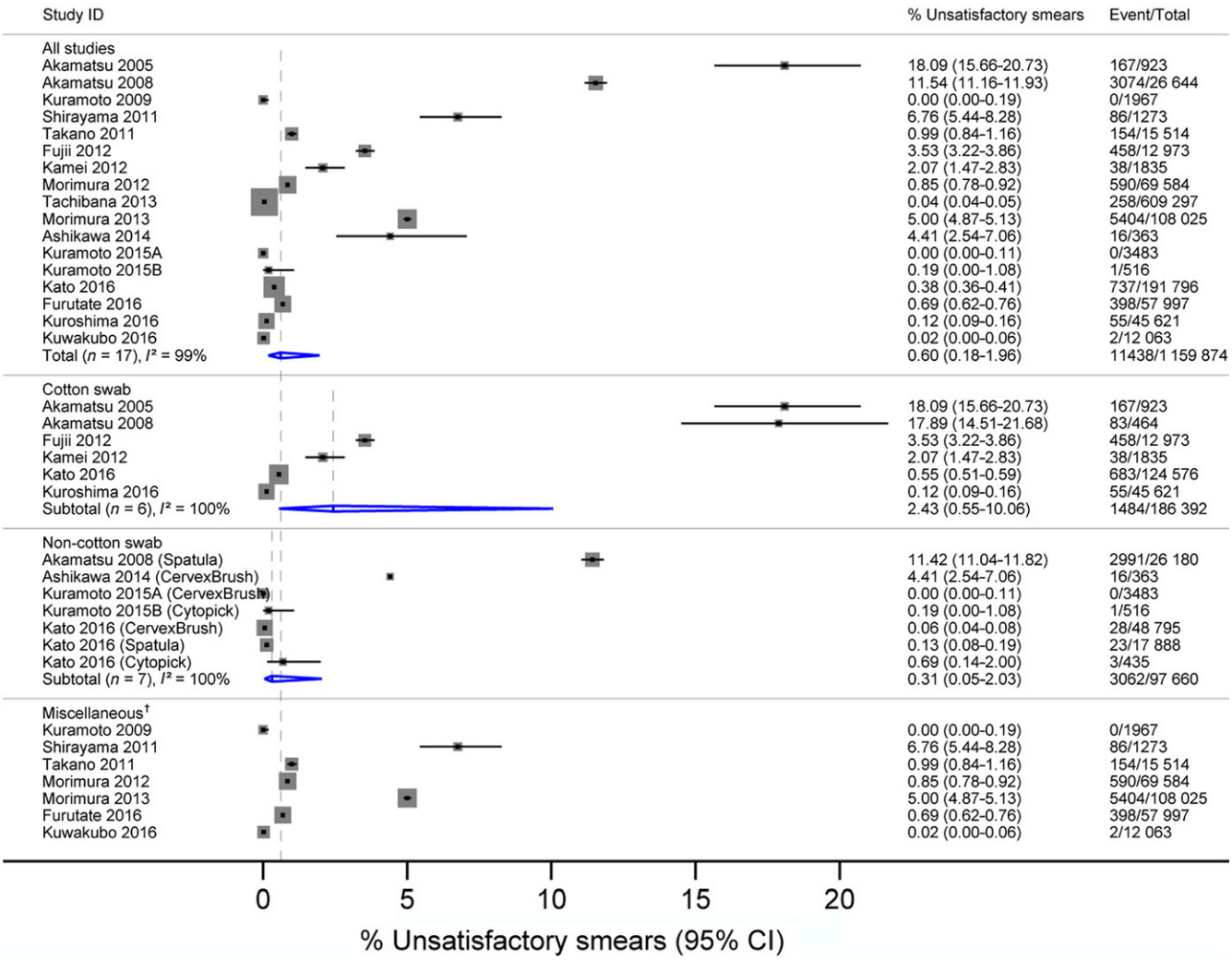

FIGURE 2 Meta-analysis of the unsatisfactory specimen rate using conventional cytology in the primary cervical cancer screening setting in Japan, according to sampling device. Diamonds depict the summary unsatisfactory specimen rate with the 95% confidence interval (CI). Each square and horizontal line indicates the unsatisfactory rate and the corresponding 95% CI, respectively, for each study. †Studies that jointly analyzed multiple different collection devices (no separate data reported). Collection devices: CervexBrush (Rovers Medical Devices, Oss, Netherlands); Cytopick (Matsunami Glass Ind., Ltd., Osaka, Japan)

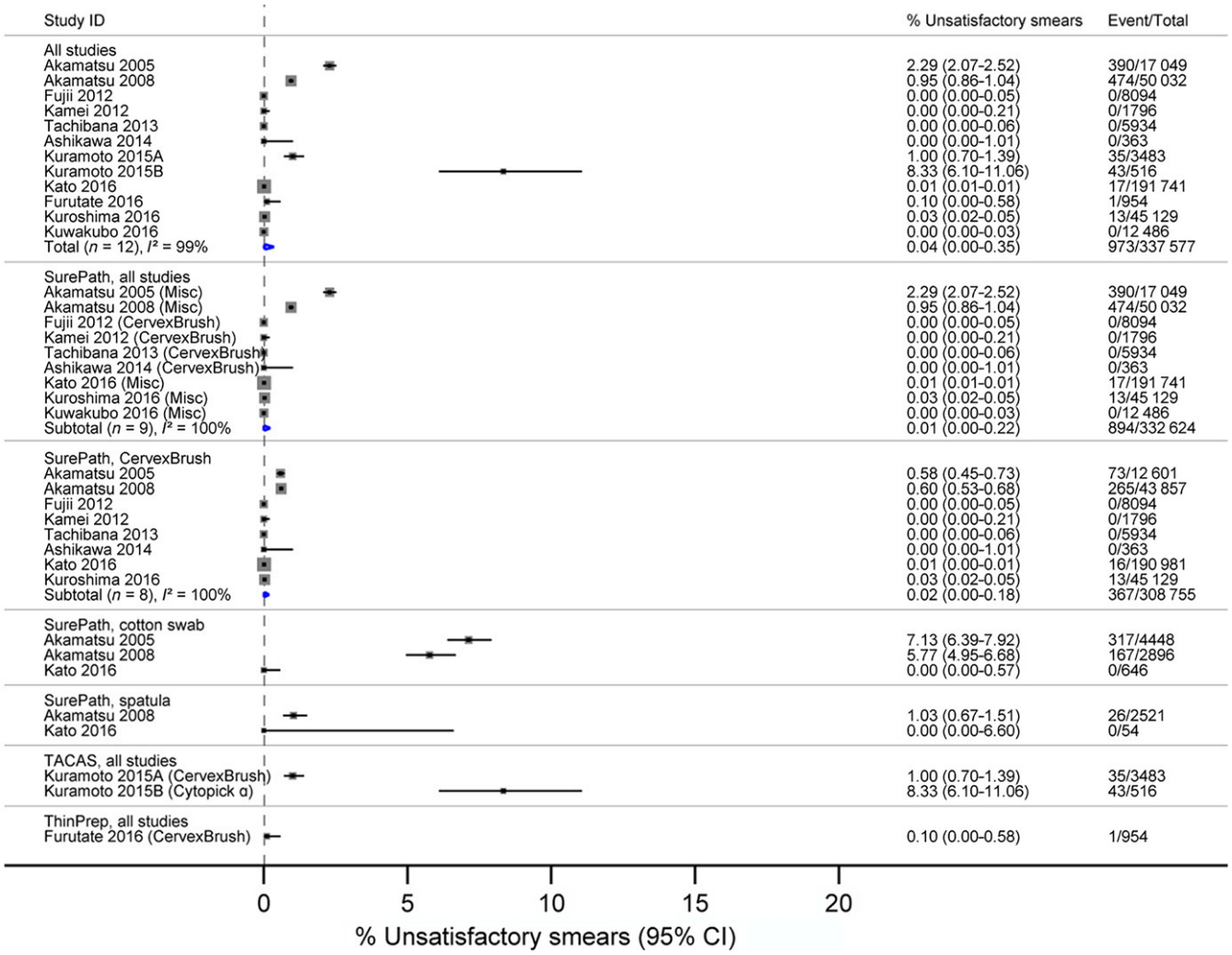

FIGURE 3 Meta-analysis of the unsatisfactory specimen rate using liquid-based cytology in the primary cervical cancer screening setting in Japan, according to cytology platform and sampling device. Diamonds depict the summary unsatisfactory rate with the 95% confidence interval (CI). Each square and horizontal line indicates the unsatisfactory rate and the corresponding 95% CI, respectively, for each study. Collection devices: CervexBrush (Rovers Medical Devices, Oss, Netherlands); SurePath (Becton Dickinson, Franklin Lakes, NJ, USA); TACAS Pro (Medical & Biological Laboratories, Nagoya, Japan); and ThinPrep (Hologic Inc., Marlborough, MA, USA). Misc, data based on joint analysis of multiple different collection devices

3.5 | Comparisons between CC and LBC

Four direct comparative studies (n = 4569) and nine indirect comparative studies (n = 1 291 950) contributed to the comparative evidence (Fig. 4). Overall, the reported comparative results were heterogeneous, and no evidence was found that either CC or LBC had a higher (or lower) proportion of unsatisfactory smear results than the other (OR = 3.5×10^{-2} favoring LBC for lower proportions of unsatisfactory results [95% CI, 6.9×10^{-4}-1.7]; I^2 = 98%). However, an evident difference was suggested between direct and indirect comparative evidence (relative OR = 8.7×10^{-4} [95% CI, 5.4×10^{-6}-0.14]; P for interaction = 0.007). Although indirect evidence suggested that LBC had a lower unsatisfactory smear proportion than CC (OR = 2.4×10^{-2} (95% CI, 4.5×10^{-3}-0.13; I^2 = 85%), direct evidence was insufficient as to whether either cytology test had a lower proportion than the other (OR = 1.9×10^{2} favoring CC for lower proportions of unsatisfactory results [95% CI, 4.9×10^{-4}-7.5×10^{7}]; I^2 = 99%).

4 | DISCUSSION

In this study, the summary unsatisfactory specimen rate in Japanese primary cervical cancer screening was estimated. The unsatisfactory rate was more likely to be lower for LBC than for CC. However, a direct comparison between CC and LBC did not show a significant difference. The proportions of unsatisfactory results for both tests continuously improved throughout the study period. In terms of the collection devices, the summary unsatisfactory proportion for CC was significantly higher for cotton swab than for non-cotton swab devices; however, a similar finding was not observed for LBC. Limited evidence suggested that use of the SurePath platform was

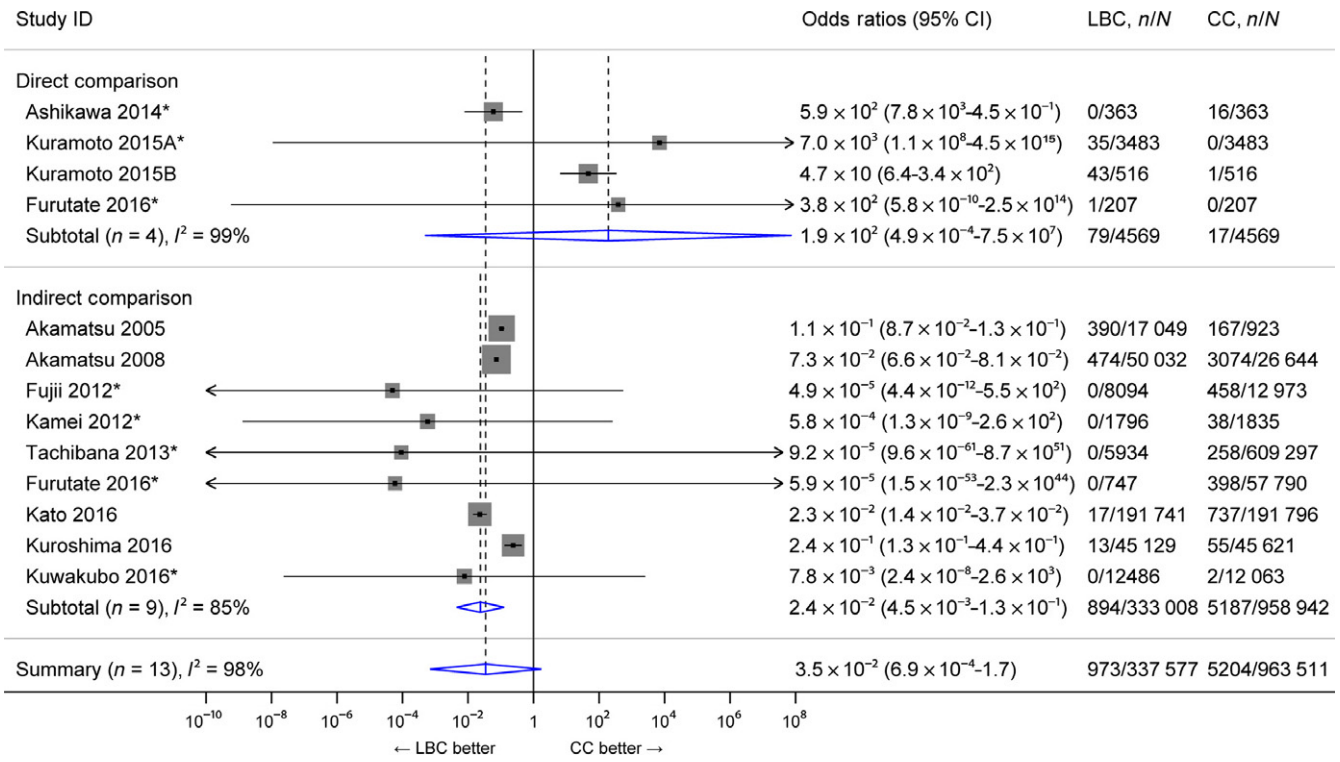

FIGURE 4 Meta-analysis of studies that compared unsatisfactory specimen rates between conventional cytology (CC) and liquid-based cytology (LBC) in the primary cervical cancer screening setting in Japan. Diamonds depict the summary odds ratio with the 95% confidence interval (CI). Each square and horizontal line indicates the odds ratio and the corresponding 95% CI, respectively, for each study. Point estimates and their 95% CIs are based on the "empirical" continuity zero-event correction

associated with lower proportions of unsatisfactory smears than non-SurePath platforms.

Although the preceding evidence indicated the advantages of LBC, CC screening is still prevalent in Japan. In Japanese screening programs, gynecologists themselves have taken cervical samples, and well-trained cytologists have interpreted the samples carefully; therefore, a lower unsatisfactory specimen rate was expected. Despite the high heterogeneity, the summary unsatisfactory rate for CC in this study was 0.60% (95% CI, 0.18-1.96), which was similar to previous studies.[5,30] Among selected studies, Akamatsu et al. reported the unsatisfactory rate for CC as 18.09% (95% CI, 15.66-20.73%)[16] and 11.54% (95% CI, 11.16-11.93%)[17] before the adoption of TBS in 2008. Although the range of the unsatisfactory rate of CC varied widely, between 0.00%[14,20,25] and 6.76%,[26] studies published after 2008 showed an improved unsatisfactory rate. Regarding the sampling device for CC, cotton swabs might cause a significantly higher unsatisfactory rate (summary rate = 2.43%, 95% CI, 0.55-10.06%) than non-cotton swab devices (summary rate = 0.31%, 95% CI, 0.05-2.03%). However, the unsatisfactory rate for cotton swabs improved from 18.09%[16] to 0.12%[22] during the study period. Morimura et al. informed the unsatisfactory rate and the reason to each medical institute, followed by the decreased unsatisfactory rate.[29] After adoption of TBS in 2008, educational programs in obtaining appropriate specimens were provided for gynecologists all over Japan. The awareness and advanced skills of Japanese gynecologists might have contributed to the improvements in the rate of unsatisfactory smears throughout the study period, regardless of the type of cytological test or sampling device. Therefore, we could observe a significant association between the first year of the study, as a proxy for the level of experience of the sample takers, and a lower unsatisfactory rate in the meta-regression analysis, but no significant association between collection device groups.

However, the range in the unsatisfactory rate for LBC varied between 0.00% and 2.29%, except for the study by Kuramoto et al.[15] The summary rate of unsatisfactory slides for LBC in this study was 0.04% (95% CI, 0.00-0.35%), which was significantly lower than for CC. In this meta-analysis, SurePath was the most popular LBC platform and had a significantly lower unsatisfactory rate. A replication study for other platforms might be required. In the near future, LBC in combination with HPV testing will be more common in Japan.

Generally, the advantages of LBC are fewer unsatisfactory results due to a monolayer of cells on a prepared slide, reduced reading times, and the residual materials can be used for HPV DNA testing. Liquid-based cytology has already been adopted in screening programs in some countries, including the UK. In 2002, Moss et al. reported that the introduction of LBC resulted in a clear reduction in the reported rate of inadequate smears (from 9% to 1-2%), with the need for fewer repeat samples, across three pilot sites in England.[31] Considering this evidence and the technique's cost-effectiveness, the National Institute for Health and Care Excellence published guidance

on the use of LBC for cervical screening in 2003, which recommended that LBC be used as the primary means of processing samples in England and Wales.[32] Rollout of LBC across the screening program in England was completed in 2008.[33] Many studies also reported that the proportion of unsatisfactory specimens was lower for LBC than for CC,[4-7,30] whereas a systematic review by Davey et al. found no evidence that LBC reduced the proportion of unsatisfactory slides[34] (Table S1). In the present study, the OR based on four direct comparative studies was not significant, whereas LBC was significantly associated with a reduced unsatisfactory rate in the indirect comparison design (Fig 4). The discrepancy between these studies might result from different study designs, study settings, age distribution of study cohorts, sample takers, cytologists, collection devices, or the LBC platform.

The strength of this study is that it is the first meta-analysis that combined 17 studies from the primary screening programs in Japan. There was a reduced unsatisfactory rate after adoption of TBS in 2008. A limitation of the study is that information on possible determinants of unsatisfactory results, such as age, study design, sampling methods, collectors of samples, or type of laboratory used, was not collected. These factors might cause high heterogeneity in this study. Second, the rate difference in the unsatisfactory specimens between CC and LBC can be an alternative statistical metric, which may be clinically more easily to interpret than the OR. However, we selected OR; the rate differences in this case were too heterogeneous to allow for statistical pooling (data not shown). Finally, it was difficult to compare the unsatisfactory rate in detail using the aggregate data approach applied in this study. An individual participant data approach could provide useful evidence, including concordance analysis (e.g., kappa statistics) or clinical information, if possible.

The evidence in this study indicated that LBC might reduce the unsatisfactory rate of cervical cytology smears compared with CC. However, comparative designs did not provide significant evidence, although evidence for the adequacy of specimens could be added as a comparison between CC and LBC. Future comparative research is required to investigate the accuracy, performance, and cost-effectiveness in the Japanese health-care setting before the adoption of LBC as the primary screening program in Japan.

ACKNOWLEDGMENTS

This study was supported by the National Cancer Center, Tokyo, Japan (grant nos. 26-A-30 and 29-A-16). The authors would like to thank Ms. Kanoko Matsushima, Ms. Noriko Hida, Mr. Tomoki Kishi, and Ms. Junko Asai for research assistance. The authors also appreciate the informative suggestions from Dr. Mika Mizuno at Aichi Cancer Center Hospital (Aichi, Japan).

DISCLOSURE STATEMENT

The authors have no conflict of interest.

ORCID

Satoyo Hosono http://orcid.org/0000-0003-2569-0127
Chisato Hamashima http://orcid.org/0000-0003-2585-8479

REFERENCES

1. Japan Asssociation of Obstetricians and Gynecologists. *Guideline for Utreine Cervix Cancer Screening*. Tokyo, Japan: Japan Asssociation of Obstetricians and Gynecologists; 1989. (in Japanese)
2. Suzuki M, Konno R, Hirai Y, Ohmura M. *For an Understanding of a Bethesda System 2001 Conformity Uterine Cervix Cytology Narrative*. Tokyo, Japan: Japan Asssociation of Obstetricians and Gynecologists; 2008. (in Japanese)
3. Solomon D, Davey D, Kurman R, et al. The 2001 Bethesda System: terminology for reporting results of cervical cytology. *JAMA*. 2002;287:2114-2119.
4. Davey E, d'Assuncao J, Irwig L, et al. Accuracy of reading liquid based cytology slides using the ThinPrep Imager compared with conventional cytology: prospective study. *BMJ*. 2007;335:31.
5. Castle PE, Bulten J, Confortini M, et al. Age-specific patterns of unsatisfactory results for conventional Pap smears and liquid-based cytology: data from two randomised clinical trials. *BJOG*. 2010;117:1067-1073.
6. Ronco G, Cuzick J, Pierotti P, et al. Accuracy of liquid based versus conventional cytology: overall results of new technologies for cervical cancer screening: randomised controlled trial. *BMJ*. 2007;335:28.
7. Beerman H, van Dorst EB, Kuenen-Boumeester V, Hogendoorn PC. Superior performance of liquid-based versus conventional cytology in a population-based cervical cancer screening program. *Gynecol Oncol*. 2009;112:572-576.
8. Hisamichi S. *Guideline for Cancer Screening Programs*. Tokyo, Japan: Japan Public Health Association; 2001. (in Japanese).
9. Hamashima C, Aoki D, Miyagi E, et al. The Japanese guideline for cervical cancer screening. *Jpn J Clin Oncol*. 2010;40:485-502.
10. Stijnen T, Hamza TH, Ozdemir P. Random effects meta-analysis of event outcome in the framework of the generalized linear mixed model with applications in sparse data. *Stat Med*. 2010;29:3046-3067.
11. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stat Med*. 2002;21:1539-1558.
12. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ*. 2003;327:557-560.
13. Ashikawa T, Umezawa T, Nomura K, et al. Specimen adequancy by BD SurePath™ liquid based cytology and conventional direct-smear cytology compared by using split-samples obtained with a Cervex-Brush™. *J Japan Soc Clin Cytol*. 2014;53:89-93. (in Japanese).
14. Kuramoto H, Sugimoto N, Iwami Y, Kato C, Hori M, Iida M. Evaluation of a new automated processing system (TACAS Pro) for liquid-based procedures. *Acta Cytol*. 2015;59:207-212.
15. Kuramoto H, Sugimoto N, Iwami Y, et al. Testing of Cytopick as a sampling device for LBP using a fully automated processing system (TACAS Pro). *J Kanagawa Soc Clin Cytol*. 2015;20:9-14. (in Japanese).
16. Akamatsu S, Himeji Y, Nagasawa Y, et al. The propriety of the Gynecological mass screening: Comparison of the Thinlayer method and the conventional method. *J Japan Soc Clin Cytol*. 2005;44:63-68. (in Japanese).
17. Akamatsu S, Himeji Y, Ikuta N, Shimagaki N, Maruoka H, Kodama S. The propriety of the gynecological mass screening: Comparison of the liquid based method and the conventional method. *J Japan Soc Clin Cytol*. 2008;47:420-424. (in Japanese).

18. Fujii K, Kakita K, Kuroda K, Tomita Y, Hasegawa T, Tomiyama M. Effects and Problems by introducin The 2001 Bethesda System. *J Prevent Med*. 2012;465:26-29 (in Japanese)
19. Kamei K, Yashiro T, Miyakita M, Ishikawa H. Comparison of liquid based cytology and conventional smears in cervical cancer screening in our health check-up clinic. *Official J Japan Soc Ningen Dock*. 2012;27:17-22. (in Japanese).
20. Tachibana M, Kasai T, Kurokawa Y, et al. Examination of the effectiveness of combining liquid based cytology and HPV-DNA testing for mass screening of uterine cervical cancer. *J Japan Soc Clin Cytol*. 2013;52:545-551. (in Japanese).
21. Kato N, Nagamine N, Miyaoka N, et al. Efficiency of the incorporated liquid-based cytology method in cervical cancer screening. *J Prevent Med*. 2016;489:24-28. (in Japanese)
22. Kuroshima Y, Ohtake K, Akamine N, et al. Effectiveness of introduction of liquid-based cytology (LBC) in municipality cervical cancer screening in the Chubu District in the main island of Okinawa—Comparison of the conventional and LBC methods. *J Japan Soc Clin Cytol*. 2016;55:137-141. (in Japanese)
23. Kuwakubo O, Abe C, Mashiko K, et al. Effectiveness of liquid based cytology and human papillomavirus testing for cervical cancer screening. *J Prevent Med*. 2016;489:20-23. (in Japanese)
24. Furutate S, Fukushima S, Takahashi M. Report on effects of liquid based cytology (LBC) for uterine cervical cancer screening. *J Prevent Med*. 2016;489:30-33. (in Japanese)
25. Kuramoto H, Iida M, Sugimoto N. Health care for females using cytology, colposcopy and pelvic examination-their efficacy-. *Yobouigaku*. 2009;192:77-83. (in Japanese)
26. Shirayama T, Seino S, Kobayashi N, Banno K, Fujii T, Aoki D. A current applicability of The Bethesda System 2001 in the uterine cervical cancer screening. *J Japan Soc Clin Cytol*. 2011;50:1-5. (in Japanese).
27. Takano H, Kasai T, Soda A, et al. Actual frequency of unsatisfactory specimens and its changes by introducing The Bethesda System 2001 to cervical cancer screening. *J Japan Soc Clin Cytol*. 2011;50:158-162. (in Japanese).
28. Morimura Y, Chiba S, Araki Y, et al. Adequacy of cervical cytology by The Bethesda System cell number in specimens in cervical cancer mass screening. *J Japan Soc Clin Cytol*. 2012;51:110-115. (in Japanese).
29. Morimura Y, Chiba S, Araki Y, et al. Information of unsatisfactory rate to individual clinician in uterine cervical mass screening decreased unsatisfactory specimens. *J Japan Soc Clin Cytol*. 2013;52:330-334. (in Japanese).
30. Fremont-Smith M, Marino J, Griffin B, Spencer L, Bolick D. Comparison of the Surepath™ liquid-based Papanicolaou smear with the conventional Papanicolaou smear in a multisite direct-to-vial study. *Cancer Cytopathol*. 2004;102:269-279.
31. Moss SMGA, Legood R, Henstock E. Evaluation of HPV/LBC cervical screening pilot studies. First report to the Department of Health on evaluation of LBC (December 2002), 2002.
32. The National Institute for Health and Care Excellence. Guidance on the use of liquid-based cytology for cervical screening. Available from URL: https://www.nice.org.uk/guidance/ta69/resources/guidance-on-the-use-of-liquidbased-cytology-for-cervical-screening-pdf-2294706450373. Accessed October 19, 2017.
33. Albrow R, Kitchener H, Gupta N, Desai M. Cervical screening in England: the past, present, and future. *Cancer Cytopathol*. 2012;120:87-96.
34. Davey E, Barratt A, Irwig L, et al. Effect of study design and quality on unsatisfactory rates, cytology classifications, and accuracy in liquid-based versus conventional cervical cytology: a systematic review. *Lancet*. 2006;367:122-132.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: Hosono S, Terasawa T, Katayama T, Sasaki S, Hoshi K, Hamashima C. Frequency of unsatisfactory cervical cytology smears in cancer screening of Japanese women: A systematic review and meta-analysis. *Cancer Sci*. 2018;109:934–943. https://doi.org/10.1111/cas.13549