Received: 9 June 2023 | Revised: 28 July 2023 | Accepted: 18 September 2023

DOI: 10.1002/ijgo.15169

REVIEW ARTICLE

Obstetrics

Cervical pessary to prevent preterm birth and poor neonatal outcome: An integrity meta-analysis of randomized controlled trials focusing on adherence to the European Medical Device Regulation

Ioannis Kyvernitakis[1] | Ahmet A. Baschat[2] | Marcel Malan[1] | Werner Rath[3] | Richard Berger[4] | Wolfgang Henrich[5] | Ekkehard Schleussner[6] | Bahareh Yousefi[7] | Nina Timmesfeld[7] | Holger Maul[1]

[1]Department of Obstetrics and Prenatal Medicine, Asklepios Clinic Barmbek, Asklepios Medical School, Hamburg, Germany

[2]Center for Fetal Therapy, Department of Gynecology & Obstetrics, Johns Hopkins University, Baltimore, Maryland, USA

[3]Department of Obstetrics and Gynecology, University Hospital of Schleswig-Holstein, Kiel, Germany

[4]Department of Obstetrics and Gynecology, Marienhaus Klinikum St. Elisabeth, Neuwied, Germany

[5]Department of Obstetrics, Charité Universitätsmedizin Berlin, Berlin, Germany

[6]Department of Obstetrics and Prenatal Medicine, University of Jena, Jena, Germany

[7]Department of Medical Informatics, Biometry and Epidemiology, Ruhr-University Bochum, Bochum, Germany

Correspondence
Ioannis Kyvernitakis, Department of Obstetrics and Prenatal Medicine, Asklepios Clinic Barmbek, Ruebenkamp 220, Hamburg 22308, Germany.
Email: janniskyvernitakis@gmail.com

Abstract

Background: Findings from randomized trials (RCTs) on cervical pessary treatment to prevent spontaneous preterm birth are inconsistent.

Objectives: Our hypothesis suggests that adhering to the European Medical Device Regulation (MDR) and following the instructions for use are essential prerequisites for successful therapy. Conversely, the non-adherence to these guidelines will probably contribute to its failure.

Search Strategy and Selection Criteria: Based on validated criteria from integrity assessments we performed a systematic review identifying 14 RCTs evaluating the effect of cervical pessaries.

Data Collection and Analysis: We analyzed the implications of 14 criteria each accounting for 0–2 points of a score reflecting the clinical evaluation plan (CEP) as proposed by the MDR to evaluate the risk–benefit ratio of medical devices.

Main Results: Seven RCTs in each singleton and twin pregnancies (5193 "cases") were included, detecting a high heterogeneity within control groups ($I^2 = 85\%$ and 87%, respectively, $P < 0.01$). The CEP score varied from 11 to 26 points for all studies. The most common reasons for low scores and potential data compromise were poor recruitment rates, no (completed) power analysis, and no pre-registration, but mainly non-adherence to technical, biological, and clinical equivalence to the instructions for use as required by the MDR. All trials with score values greater than 20 had applied audit procedures. Within this group we found significantly reduced rates of spontaneous preterm birth at less than 34 weeks within the pessary group in singleton (odds ratio 0.28; 95% confidence interval 0.12–0.65) and twin pregnancies (odds ratio 0.30; 95% confidence interval 0.13–0.67). Similarly, there was a significant reduction in the composite poor neonatal outcome in singleton (odds ratio 0.25; 95% confidence interval 0.10–0.61) and twin pregnancies (odds ratio 0.54; 95% confidence interval 0.35–0.82) after a pessary as compared with controls.

Conclusion: Non-audited RCTs and meta-analyses mixing studies of different clinical quality as pre-defined by a CEP and the MDR pose the risk for erroneous conclusions.

KEYWORDS
audit, integrity, MDR, meta-analysis, pessary, singletons, twins

1 | INTRODUCTION

Preterm birth (PTB) is the leading cause of perinatal and infant mortality, accounting for approximately one-third of newborn deaths.[1] Among survivors, short-term complications and the risk for long-term neurodevelopmental, cardiovascular, and metabolic diseases are increased.[2] Spontaneous PTB (sPTB) is a syndrome with many causes.[3] During the last 30 years, premature cervical shortening diagnosed by transvaginal sonography[4] has become a pragmatic tool to identify asymptomatic patients at risk for PTB, although the sensitivity of a short cervical length for the prediction of sPTB varies.[5]

Cervical pessary placement in patients with a short cervical length and subsequent standardized management following the instructions for use (IFUs) reduced sPTB and adverse neonatal outcomes in the first randomized controlled trial (RCT) in singleton pregnancies.[6] However, subsequent trials and meta-analyses have shown inconsistent findings after pessary use. As a medical device, the IFU and the technical specification of a pessary are defined by the European Medical Device Regulation (MDR) or the US Food and Drug Administration.[7] Variable adherence to IFUs and a learning curve can impact the performance of the pessary and ultimately affect outcomes.[8] It has been demonstrated that clinical expertise, teaching, and continuous audit significantly influence outcomes and are therefore mandatory.[9] Although meta-analyses offer the theoretical advantage of greater statistical power by generating large sample sizes, important study details, such as clinical management or adherence to IFUs or audit procedures, are not necessarily considered for inclusion.[10,11] We hypothesize that this causes study heterogeneity in clinical standards and has an impact on outcomes in patients treated with a cervical pessary.

2 | METHODS

In May 2022, an electronic search of the databases PubMed and MEDLINE was finalized to identify RCTs of singleton and twin pregnancies that compared cervical pessary and standard care with standard care alone or alternative interventions for the prevention of PTB, with adverse neonatal outcomes as endpoints (Table 1). Primary search terms were cervical pessary AND preterm birth AND singleton OR twin pregnancy, including terms such as “no pessary” or “vaginal progesterone”. The language was restricted to English. Trials where the cervical pessary was indicated after an episode of contractions, in patients with placenta previa, or trials without any defined clinical outcome were excluded.

For each report, we extracted data on study design, ethics board approval, randomization, baseline characteristics, and outcomes. Studies were graded in 15 integrity domains that were modeled on MDR and CoRe Outcomes in WomeN's health (CROWN) initiative recommendations (Table 2). These integrity domains assigned three quality levels (0–2 points) for the device selection, choice of additional treatments, trial registration, criteria, conduct, audit, and reported outcome measures. We ascertained trial registration in clinicaltrials.gov and national databases to evaluate for retrospective bias[12] and performed a complementary analysis to assess the publication record of the study groups on cervical pessary, publication of RCTs, and meta-analyses in PubMed. For each group of investigators, we graded experience with pessary placement by the number of pessaries applied before the study start, during the study (both per center and per investigator), and estimated the clinical experience of the first author (supervisor). At least 30 applications was defined as an experience level where clinical success rates can be expected[8] (Table 3).

Additional information was extracted from study protocols, publications, personal communication, and online records regarding pessary type (including certified and non-certified products), selection criteria, data on ethical oversight, safety level, compliance, and referral to product instructions, power analyses, report on long-term outcomes according to CROWN, and finally audit.

An audit was defined as a supervised procedure prior to study initiation, a required prerequisite by the MDR for medical devices to verify conformance to quality standards and the true potential effectiveness of a product.[13] Study investigators who received supervised training for pessary insertion, follow-up, and removal AND showed adherence to the current IFUs were qualified as audited. Simply written guidance leaflets or video recordings regarding insertion, management, and removal alone without personal feedback were not considered as an audit procedure.

We further evaluated details on the gestational age at insertion and removal as well as reasons for “early pessary removal”, characteristics of additional interventions used in both the study and the control group, and clinical compliance. Finally, details of outcome data were specified such as the kind and number of poor outcome characteristics in each group to calculate effect sizes (Table 3).

To detect risks of bias, such as non-adherence to IFUs, missing outcome data, missing definition of outcome criteria, and selection of reported results, a meticulous analysis of compliance with the MDR requirements[13] was conducted for each trial, considering the Medical Device Coordination Group (MDCG) Documents, MDCG

TABLE 1 Baseline characteristics of the randomized trials.

Author, year [reference]	Recruitment	Pessary vs control group		GA at recruitment	CL at recruitment (mm)	Progesterone?	Primary outcome	Pre-registration according to[12]
Singleton pregnancies								
Goya, 2012[6]	CL ≤ 25 mm	Arabin ($n=190$)	No pessary ($n=190$)	20–23, mean 22.3	19.0 ± 4.8	No	sPTB < 34 wks	Clin.Trials.gov NCT00706264
Hui, 2013[20]	CL < 25 mm	Arabin ($n=53$)	No pessary ($n=55$)	20–24, mean 21.9	20.1 ± 0.5	No	PTB < 34 wks	Clin.Trials.gov: not registered, ISRCTN18185477
Nicolaides, 2016[22]	CL ≤ 25 mm	Arabin ($n=465$)	No pessary ($n=467$)	20–24, mean 23.5	20.0 (14.0–22.0)	Pessary group 44%, controls 47%	sPTB < 34 wks	Clin.Trials.gov NCT00735137
Karbasian, 2016[17]	CL ≤ 25 mm	Arabin + vaginal progesterone ($n=71$)	Vaginal progesterone ($n=73$)	18–22, mean 19.6	22.0 ± 1.7	100%	PTB < 37 wks	Clin.Trials.gov: not registered
Saccone, 2017[37]	CL ≤ 25 mm, (pregnancies with previous PTB excluded)	Arabin ($n=150$)	No pessary ($n=150$)	18–23, mean 22.4	12.0 ± 5.8	Pessary group 89%, controls 83%	sPTB < 34 wks	Clin.Trials.gov NCT02716909
Dugoff, 2018[18]	CL ≤ 25 mm, (pregnancies with previous PTB excluded)	Bioteque Cup Pessary ($n=60$)	No pessary ($n=58$)	18–23, mean 21.1	Pessary group 17.6 (10.9–22.0); controls 19.0 (11.2–22.9)	Pessary 84%, controls 91%	PTB < 37 wks	Clin.Trials.gov NCT02056652
Cruz-Melguizo, 2018[19]	CL ≤ 25 mm	Arabin ($n=125$)	Vaginal progesterone ($n=118$)	19–22, mean 21.3	20.9 ± 4.2	Pessary 5%, progesterone group 100%	sPTB < 34 wks	Clin.Trials.gov NCT01643980
Multiple pregnancies								
Liem, 2013[28]	Multiple pregnancies (97.8% twins; 2.2% triplets)	Arabin ($n=401$)	No pessary ($n=407$)	12–20, mean 17.0	43.9 ± 8.3	No	Composite adverse perinatal outcome	Dutch trial registry: NTR1858
Nicolaides, 2016[23]	Twin pregnancies	Arabin ($n=588$)	No pessary ($n=589$)	20–24, mean 22.7	32.0 (27.0–37.0)	Pessary group 0%, controls 0.3%	sPTB < 34 wks	Clin.Trials.gov nicht registriert ISRCTN 01096902
Goya, 2016[25]	Twin pregnancies, CL ≤ 25 mm	Arabin ($n=68$)	No pessary ($n=66$)	20–23, mean 22.3	19.4 ± 3.6	No	sPTB < 34 wks	Clin.Trials.gov NCT01242410
Berghella, 2017[21]	Twin pregnancies, CL ≤ 30 mm	Bioteque Cup Pessary ($n=23$)	No pessary ($n=23$)	18–27, mean 21.1	Pessary group 16.7 (10.7–27.8), controls 22.9 (15.9–25.6)	Pessary group 4%, no pessary 9%	PTB < 34 wks	Clin.Trials.gov NCT02056639
Dang, 2019[27]	Twin pregnancies, CL < 38 mm	Arabin ($n=148$)	Vaginal progesterone 400 mg/day ($n=149$)	16–22, mean 17.8	31.3 ± 4.3	Pessary group 1%, progesterone group 100%	PTB < 34 wks	Clin.Trials.gov NCT02623881
Norman, 2020[24]	Twin pregnancies, CL ≤ 35 mm	Arabin ($n=230$)	No pessary ($n=250$)	18–21	Pessary group 28.8 (3.0–35.0), controls 29.5 (7.0–35)	No	sPTB < 34 wks, composite neonatal outcome	Clin.Trials.gov NCT02235181
Groussolles 2022[26]	Twin pregnancies, CL < 35 mm	Arabin ($n=157$)	No pessary ($n=158$)	16–24		No	Adverse neonatal outcome	Clin.Trials.gov NCT02328989

Abbreviations: CL, cervical length; GA, gestational age; PTB, preterm birth; sPTB, spontaneous preterm birth; wks, weeks.

TABLE 2 Summary of 15 criteria stratified according to a clinical evaluation plan based on criteria of the Medical Device Regulation and the integrity domain for the device selection, choice of additional treatments, trial registration, criteria, conduct, audit, and outcome, all stratified according to three quality levels (0–2) of each parameter.

Criteria	Description	Grade	Definition
Appropriate device (technical equivalence)	Were the data generated from the product in question (Arabin cervical pessary perforated)?	D2	Arabin cervical pessary perforated
		D1	Equivalent device (non-perforated)
		D0	Other device not certified for prevention of preterm birth
Further pharmacological treatment (biological equivalence)	Was additional treatment unequally distributed between treatment and control groups?	PT2	No other treatments (<10%)
		PT1	Tocolytics/progesterone (>10%)
		PT0	Antibiotics (>10%)
Pre-registration	Was the RCT pre-registered in clinicaltrials.org or others?[a]	PR2	Pre-registered in clinicaltrials.org
		PR1	Registered in local boards
		PR0	Not registered at all
Patient selection for recruitment	What were the inclusion criteria?	PS2	Centiles adapted to gestational age
		PS1	Strict cut-off values (e.g. 25 mm)
		PS0	No selection criteria as predefined outcome
Acceptable report	Does the report or the data pool contain sufficient information for conducting a rational objective assessment? (e.g. background risk)	R2	High quality (GA ≤ 24 weeks, history of PTB, conization or similar defined)
		R1	Minor deviation (GA > 24 weeks, risk history not evaluated)
		R0	Insufficient information (only CL)
Selection criteria for control groups	Were the data obtained from a patient group representative for the intended purpose and for the clinical condition?	P2	Patient selection according to cervical shortening centiles
		P1	Patient selection according to cervical shortening cut-off values
		P0	No predefined selection criteria
Data source	Were ethical criteria applied?	S2	Appropriate
		S1	Minor deviation
		S0	Major deviation
Safety level	Does the study refer to the clinical safety criteria?	L2	High (severance and incidence described)
		L1	Medium (severance or incidence described)
		L0	Low (no data)
Experience clinicians before RCT	How many patients had the clinicians already treated before the start of the RCT?	EC2	(on average) >30 patients
		EC1	5–30 patients
		EC0	0–4 patients[a]
Performance clinicians during RCT	What was the average recruitment of patients/center?	PC2	(on average) >30 patients
		PC1	5–30 patients
		PC0	0–4 patients[a]
Experience supervisor	What was the experience of the supervisor (first author) before the start of the RCT?	ES2	(on average) >30 patients
		ES1	5–30 patients
		ES0	0–4 patients[a]
Audits	Central control of diagnosis and therapy at the start of the trial?	AU2	Diagnosis and therapy
		AU1	Only diagnosis
		AU0	None
Compliance to instruction	Was current instruction for use of the device referred?	IFU2	Fully 80–100%
		IFU1	Partly 50–80%
		IFU0	Not recognizable < 50%

TABLE 2 (Continued)

Criteria	Description	Grade	Definition
Power analysis	Quality of adherence	PA2	Equivalent in patient number
		PA1	Underpowered
		PA0	Not indicated
Outcome	According to CROWN	O2	Short- and long-term follow-up (>2 years)
		O1	Short-term (composite) neonatal outcome
		O0	Only rates of PTB

Abbreviations: CL, cervical length; CROWN, CoRe Outcomes in WomeN's health; GA, gestational age; PTB, preterm birth; RCT, randomized controlled trial.
[a]Prior et al. 2017.[12]

2020-5 and 2020-6. These guidelines describe how to ensure technical, biological, and clinical equivalence to the investigated device under the directives 93/42/EEC and 90/385/EEC in clinical trials.[14,15]

Separate analyses were performed for screened and at-risk patients with a singleton pregnancy and for unselected or selected patients with twin pregnancies and a short cervix. The pooled relative risk for dichotomous data, mean difference, and 95% confidence interval for continuous data were calculated. Heterogeneity of the treatment effect was assessed with the I^2 statistic. Results from individual studies were pooled using a random-effects model. The number needed to treat was calculated with a 95% confidence interval where meta-analysis of dichotomous outcomes revealed a statistically significant beneficial or harmful effect of a cervical pessary. Furthermore, we explored potential sources of heterogeneity and assessed publication and related biases by examining the symmetry of funnel plots using the Egger test.

To account for dependence of outcomes for infants from multifetal gestations we carried out a sensitivity analysis where we treated infants from the same pregnancy as clusters and analyzed data using methods described for cluster-randomized trials.[16] The data were adjusted using an estimate of the intra-cluster correlation coefficient (ICC) derived from the trial (if possible), or from similar trials. For multifetal gestations where ICCs were not available, it was estimated, a sensitivity analysis performed and the effect of using two extremes of ICC was tested. Subgroup analyses were performed according to concomitant use of vaginal progesterone (yes vs no), cervical length and obstetric history (no previous PTB vs at least one previous PTB), and nulliparous versus parous women with no or at least one previous PTB. In addition, we investigated whether treatment effects differed between subgroups by an interaction test between treatment groups and specific subgroups. These analytic approaches were performed for cohorts stratified by study quality criteria to evaluate their impact on outcomes. An interaction *P*-value of 0.05 or more was considered to indicate that the effect of treatment did not differ significantly among subgroups. Subgroup and sensitivity analyses were performed for the primary outcome of sPTB before 34 weeks of gestation.

The meta-analysis was registered in PROSPERO on May 10, 2022 under the reference number CRD42022257456.

3 | RESULTS

Among 1149 publications identified during the initial search, 14 RCTs were eligible for inclusion in this meta-analysis, conducted over a period of 14 years, recruiting 5193 participants (Table 1 and Figure 1). Among this cohort, five studies all applying audit procedures reached 20 or more points according to the predefined clinical evaluation plan (CEP) and nine studies did not reach this cut-off level and had 19 points or less (Tables 2 and 3).

3.1 | Heterogeneity

Heterogeneity was determined for inclusion data of all participants in the control groups. Figure 2a shows the wide distribution of different characteristics within the trials for singleton controls ($n=1038$, $I^2=85\%$, $P<0.01$) suggesting variable selection criteria between trials. Similar differences ($n=1235$, $I^2=91\%$, $P<0.01$) were observed for controls in RCTs in twin pregnancies (Figure 2b).

3.2 | Productivity and recruitment rates

There were discrepancies between the estimated number of patients needed to treat for the power analyses and, even worse, the lack of any predefined power analysis in one trial.[17] Among trials that failed to achieve the primary outcomes in pessary treatment, enrollment rates ranged from 28.0% to 64.4%[18–21] and there was no explanation on the clinical aspects of pessary placement and management. Furthermore, many pessaries were prematurely removed, resulting in a relevant discrepancy between the intention to treat and the per protocol analyses. In the trial of Nicolaides et al.,[22] 24.5% of pessaries (114 from 465) were removed too early after placement in singleton pregnancies and 22.3% (131 from 588) in twin pregnancies[23]

TABLE 3 Application of the clinical evaluation plan involving 14 randomized controlled trials (RCTs) on pessary treatment to prevent preterm birth; the cut-off value of 20 points separates studies without ($n = 9$) and with ($n = 5$) an appropriate clinical audit.

Criteria/study	Goya, 2012[6]	Hui, 2013[20]	Nicolaides, 2016[22]	Karbasian, 2016[17]	Saccone, 2017[37]	Dugoff, 2018[18]	Cruz-Melguizo, 2018[19]	Liem, 2013[28]	Nicolaides, 2016[23]	Goya, 2016[25]	Berghella, 2017[21]	Dang, 2019[27]	Norman, 2020[24]	Groussolles, 2022[26]
Appropriate device (technical equivalence)	2	2	1	2	2	0	1	1	1	2	0	2	2	2
Further pharmacological treatment (biological equivalence)	2	2	0	1	1	1	1	2	2	2	2	1	2	2
Pre-registration	2	1	2	0	2	2	2	1	1	2	2	2	2	2
Patient selection for recruitment	1	1	1	1	1	1	1	2	1	1	1	2	1	1
Acceptable report	2	0	2	0	2	2	2	2	2	2	2	2	2	2
Data source	2	0	2	2	2	2	2	2	2	2	2	2	2	2
Safety level	2	1	2	2	2	2	2	2	2	2	2	2	2	2
Experience clinicians before RCT	2	0	0	0	2	1	0	2	0	2	0	0	0	0
Performance clinicians during RCT	2	2	1	2	2	2	2	1	1	2	0	2	0	1
Experience supervisor	2	0	0	0	2	0	1	2	0	2	0	0	0	0
Audits	2	0	0	0	2	0	0	1[a]	0	2	0	2	0	0
Referral to instruction criteria	2	0	0	0	0	0	0	1	0	1	0	2	1	1
Power analysis	2	1	2	0	2	1	2	2	2	2	1	2	2	2
Outcome	1	1	1	1	1	1	1	2	1	1	1	2	1	1
Total grading	26	11	14	11	23	15	17	23	15	25	13	23	17	18

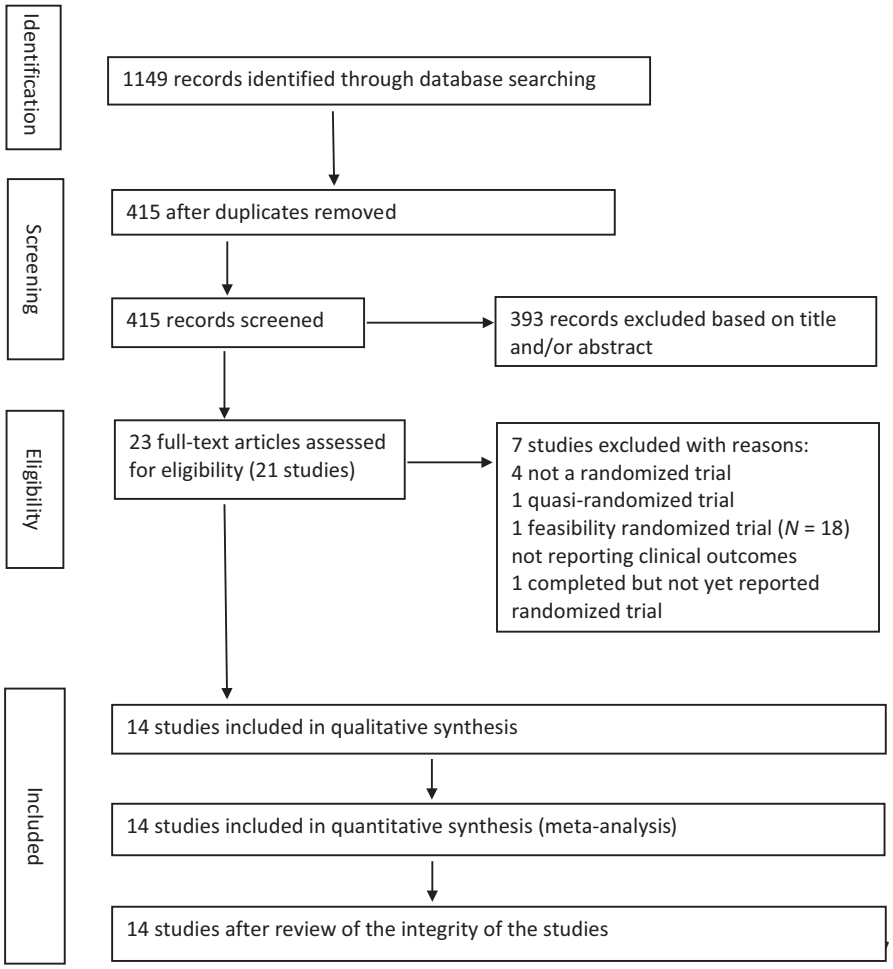

FIGURE 1 Flowchart summarizing the inclusion of studies. Among 1149 publications identified during the initial search, 14 randomized controlled trials were eligible for inclusion in this meta-analysis conducted over a period of 14 years, recruiting 5193 participants.

because of complaints of vaginal discharge, which is a recognized benign adverse effect of the pessary. In the trial of Berghella et al.,[21] 16 from 23 (70%) devices were removed prior to 36 weeks of pregnancy. Norman et al.[24] reported on an 11.3% rate (26/230) of pessary removal after patient requests and a 5.7% (13/230) rate of spontaneous expulsion.

3.3 | Effect of teaching and audit procedures before inclusion

The primary obstetric outcome within the total group of singleton pregnancies did not reveal significant differences between the pessary and the standard care group (odds ratio 0.73, 95% confidence interval 0.37–1.44). However, when only studies from centers with an audit procedure were included, the risk reduction was in favor of the pessary group (odds ratio 0.28, 95% confidence interval 0.12–0.65). In contrast, centers without any practical experience or audit did not reach any significant difference between the intervention and control groups (odds ratio 1.15, 95% confidence interval 0.84–1.58), see Figure 3a. There were no differences between the intervention and control groups in gestational age at birth in the total group (odds ratio 0.69, 95% confidence interval −0.25 to 1.63; Figure 3b). In centers with an audit, the rate of the composite adverse outcome was significantly lower in the pessary groups as compared with controls (odds ratio 0.25, 95% confidence interval 0.10–0.61; Figure 3c,d).

Among twin trials there were three studies[23,25,26] that evaluated rates of spontaneous preterm deliveries as compared with other trials[21,26–28] considering all (preterm) deliveries, including inductions of labor and any form of cesarean delivery. There was a significantly lower frequency of sPTB in the pessary group as compared with controls within centers after an audit procedure (odds ratio 0.30, 95% confidence interval 0.13–0.67). Similarly, there were no significant differences between the study arms (odds ratio 0.71, 95% confidence interval 0.35–1.43) when audit procedures were not performed (Figure 4a). However, audit influenced the composite adverse outcome, with a significant reduction in pessary groups as compared with controls (odds ratio 0.54, 95% confidence interval 0.35–0.82) (Figure 4b–d).

The frequency of tocolysis, vaginal discharge, antenatal corticosteroids, cesarean deliveries, and perinatal mortality in both singleton and twin pregnancies is demonstrated in Table 4.

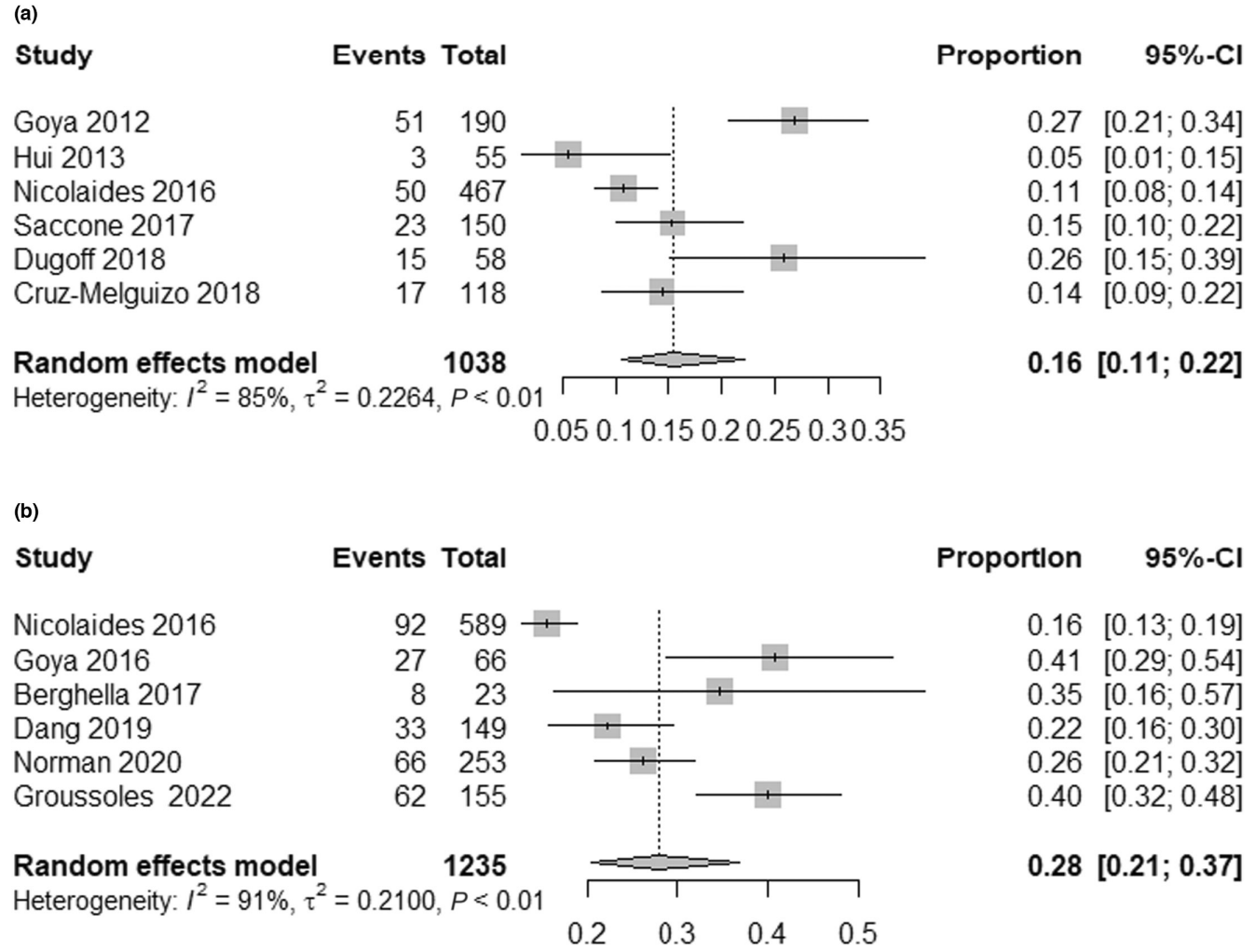

FIGURE 2 Heterogeneity between control groups in randomized controlled trials in (a) singleton pregnancies and (b) twin pregnancies.

3.4 | Applying the CEP to all 14 trials

The risk of bias is potentially attributable to non-compliance with technical, biological, and clinical criteria as defined by IFUs and MDCG guidelines.[13–15] Dugoff et al.[18] and Berghella et al.[21] used Bioteque pessaries for the treatment of genital prolapse, which unlike the Arabin pessary are not Conformite Europeenne (CE) certified for the prevention of PTB (Tables 2 and 3).[29] The concurrent use of antibiotics and vaginal progesterone differed between trials, potentially affecting the biological effects. A learning curve of 30 patients, as already published by Franca et al.,[8] had not been completed by investigators in many trials.[17–24,26,27]

3.5 | Study oversight

Within the 14 RCTs, ethic committees that had approved the study design were usually listed as the institutional or local ethics committee. Nevertheless, three studies retrospectively evoked ethical concerns, which were finally all "underpowered". Specifically, they stopped recruitment prematurely. In addition, Hui et al.[20] selected only low-risk cases, whereas the risk patients of the same department were presented in a cohort study reaching different conclusions.[30] Although the authors stated that the study was "blinded", it was not specified how this was achieved, as healthcare providers and patients must be informed of the risks and the need to remove the device in case of labor. The studies by Berghella et al.[21] and Dugoff et al.[18] used products that are not certified for the prevention of PTB.

4 | DISCUSSION AND CONCLUSIONS

Our study applied established quality metrics to 14 RCTs determining the benefits and risks of cervical pessaries in the prevention of PTB. We identified significant variations in study design, quality metrics, and management and demonstrated that neither the compliance to IFUs nor clinical experience were considered in most trial protocols. Stratifying our meta-analysis by criteria of the European MDR and adherence to IFUs we confirmed that placement and supervision of an Arabin pessary by experienced practitioners reduced the risk for PTB in singleton and twin pregnancies.[31,32] Our findings highlight important factors contributing to the benefit of

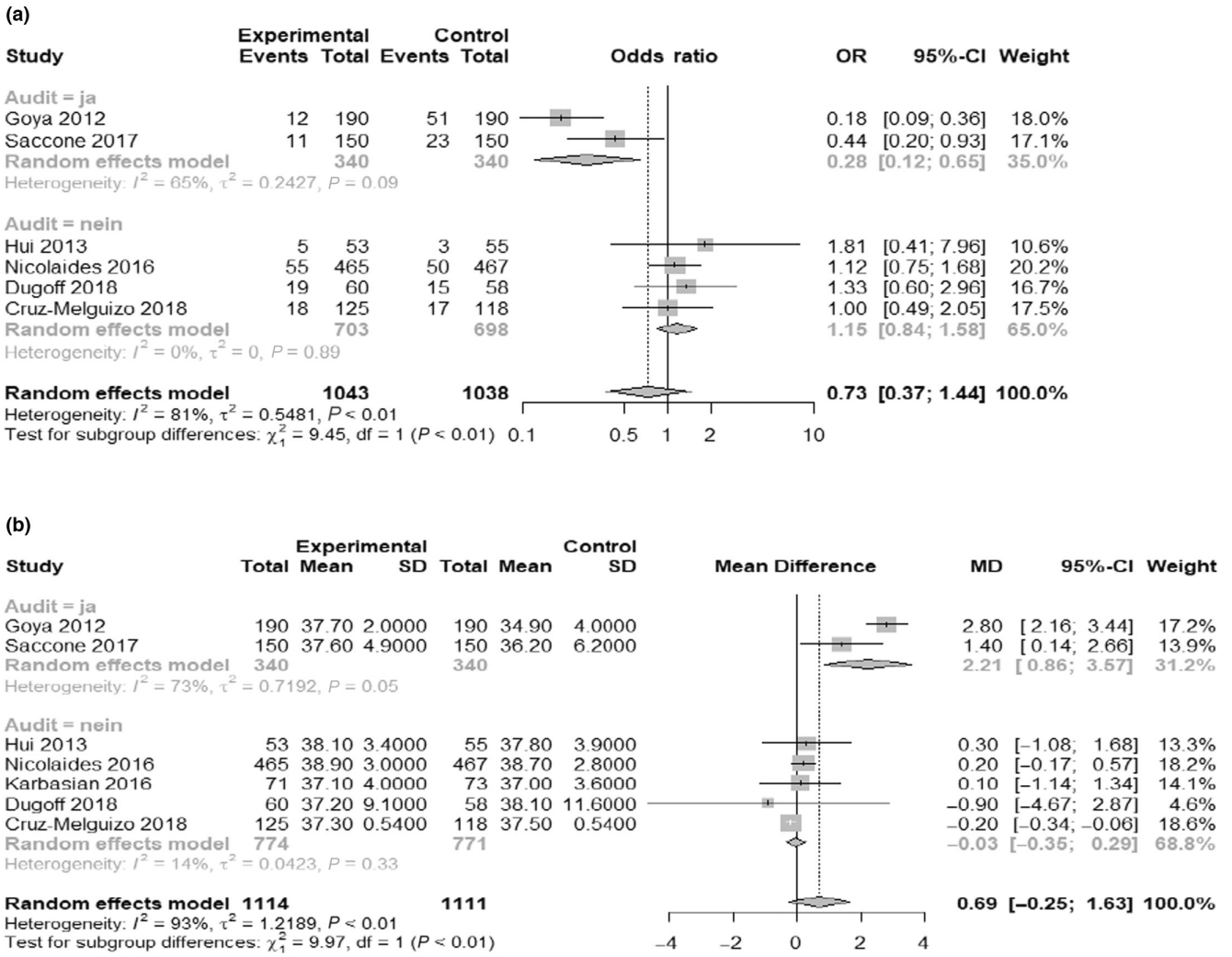

FIGURE 3 (a) Spontaneous preterm birth before 34 weeks of pregnancy in singleton pregnancies according to audit procedure. (b) Gestational age at delivery in singleton pregnancies with and without pessary, according to audit procedure. (c) Composite neonatal outcome in singletons with and without pessary, according to audit procedure. (d) Perinatal mortality in singleton pregnancies, according to audit procedure.

cervical pessary placement and show the potential disadvantage of systematic reviews and meta-analyses that do not account for quality metrics between studies. Our findings indicate that this may be a considerable confounder for meta-analyses evaluating cervical pessary use for the prevention of PTB.

Statistical criteria for systematic reviews were originally established for pharmacologic interventions. However, clinical trials may require more stringent audit related to indication and performance, especially when mechanical devices are evaluated. We have witnessed systematic failures of patient selection, pessary insertion, surveillance, and removal, suggesting potential clinical bias. We have observed significant heterogeneity among trials because of differences in selection criteria. Trials with applied audits showed a significant reduction in sPTB and composite adverse neonatal outcome. The importance of audits is illustrated by the Trial of Umbilical and Fetal Flow in Europe (TRUFFLE), where Doppler waveforms, even if obtained by experienced examiners, were independently audited.[33] Accordingly, audits appear prudent for the indication, management strategy, and level of management expertise for trials on high-risk patients with a PTB syndrome.[34]

The authors of the first pessary RCT in pregnancies where the patients had a short cervix[6] only involved qualified and audited specialists, concluding that pessary treatment prevented PTB and neonatal morbidity. The same study group demonstrated a significant reduction in PTB in twin pregnancies after pessary placement. The Dutch ProTwin trial[28] reported similar results in a subgroup of twin pregnancies in patients with a short cervical length.

In 2016, the importance of teaching and adherence to protocols was documented by a secondary analysis of the ProTwin trial.[35] In the same year, two multi-continental RCTs were published.[22,23] Regrettably, clinicians who have systematically proclaimed audits for prenatal research[36] have not applied teaching and audits for the complex indication and treatment with cervical pessaries. One possible explanation might be the lack of clinical experience of study coordinators themselves. The additional lack of a protocol based on

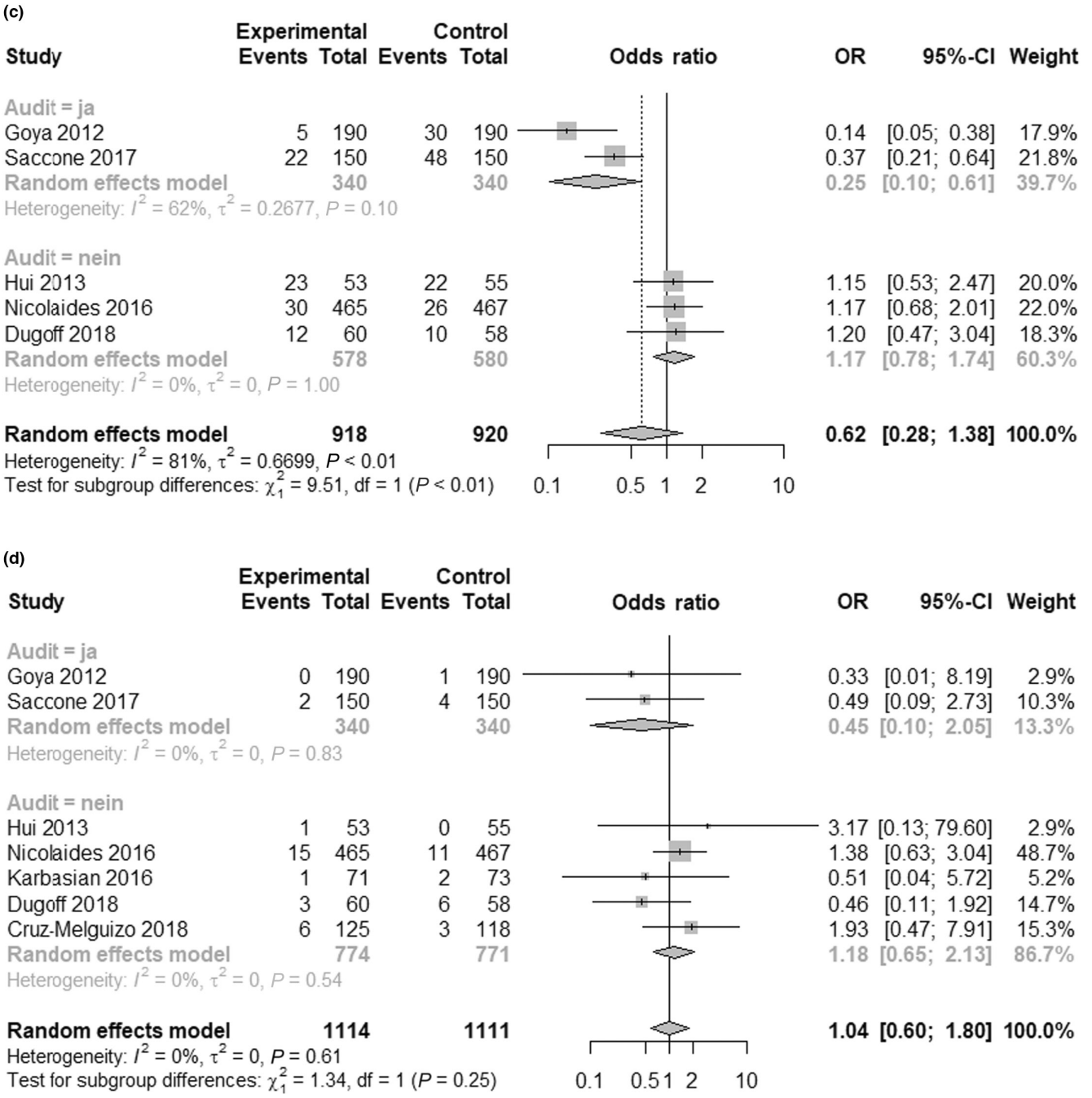

FIGURE 3 (Continued)

the IFUs explains the high rate of early pessary removal and unnecessary antibiotics. Even though these papers were published in high-impact journals they did not reach the standards as required by the MDR (Table 3).

Even the French PESSARONE trial[26] failed to reach the MDR-based quality metrics, as 10.8% of the participants had missing values for the primary endpoint, 79% of parents refused neonatal data collection, and recruitment rates were extremely low.

Saccone et al.[37] found a significant reduction of PTB in singleton pregnancies when pessaries were additionally applied with vaginal progesterone. A two-fold reduction in PTB less than 34 weeks and up to a four-fold reduction in adverse neonatal outcome were published in twin pregnancies in patients with a cervical length less than the 25th centile.[27]

The MDR has introduced regulations[13] to avoid the application of non-equivalent studies for guidelines and recommendations. Recent meta-analyses have not considered the MDR criteria.[24,38] Cannie et al.[39] demonstrated that the pessary can reduce cervical funneling and elongate the cervix by cervical "sacralization", as shown in magnetic resonance imaging, and warned that during follow-up device

dislocations should be excluded. Technical properties of different pessary types, (low) experience, and (insufficient) learning curves as reflected by the number of recruitments per center for the complete management were not reported in trials, a trend that has continued in a current "network meta-analysis" (Mol et al. submitted). It is not surprising that authors of trials with recruitment below five patients per center[24] reported on high rates of patient discomfort, suggesting suboptimal application or reassurance.[40] Nicolaides and coworkers[22,23] determined that vaginal discharge following pessary placement was not due to bacterial pathogens as more than 40% of patients received essentially unindicated antibiotic therapy. Notably, antibiotics may have an impact on the vaginal microbiome and even induce PTB.[22]

Even within the meta-analysis by Conde-Agudelo et al.[38] the criteria of technical, biological, or clinical equivalence or adherence to IFUs were not even debated. Instead, the "blinded" underpowered study of Hui et al.[20] carried out by unexperienced investigators was suggested to be a perfect trial within their Cochrane tool.

The criteria we applied to objectify the compliance with MDR and IFUs (Tables 2 and 3) demonstrated the clinical inconsistency of 14 RCTs. The term "tsunami of meta-analyses" by Alfirevic[41] addressed hyperprofilic analyses of heterogeneous findings. In addition, inadequate selection of low-risk control groups and mixed populations with different background risks from any healthcare system were common but not debated.

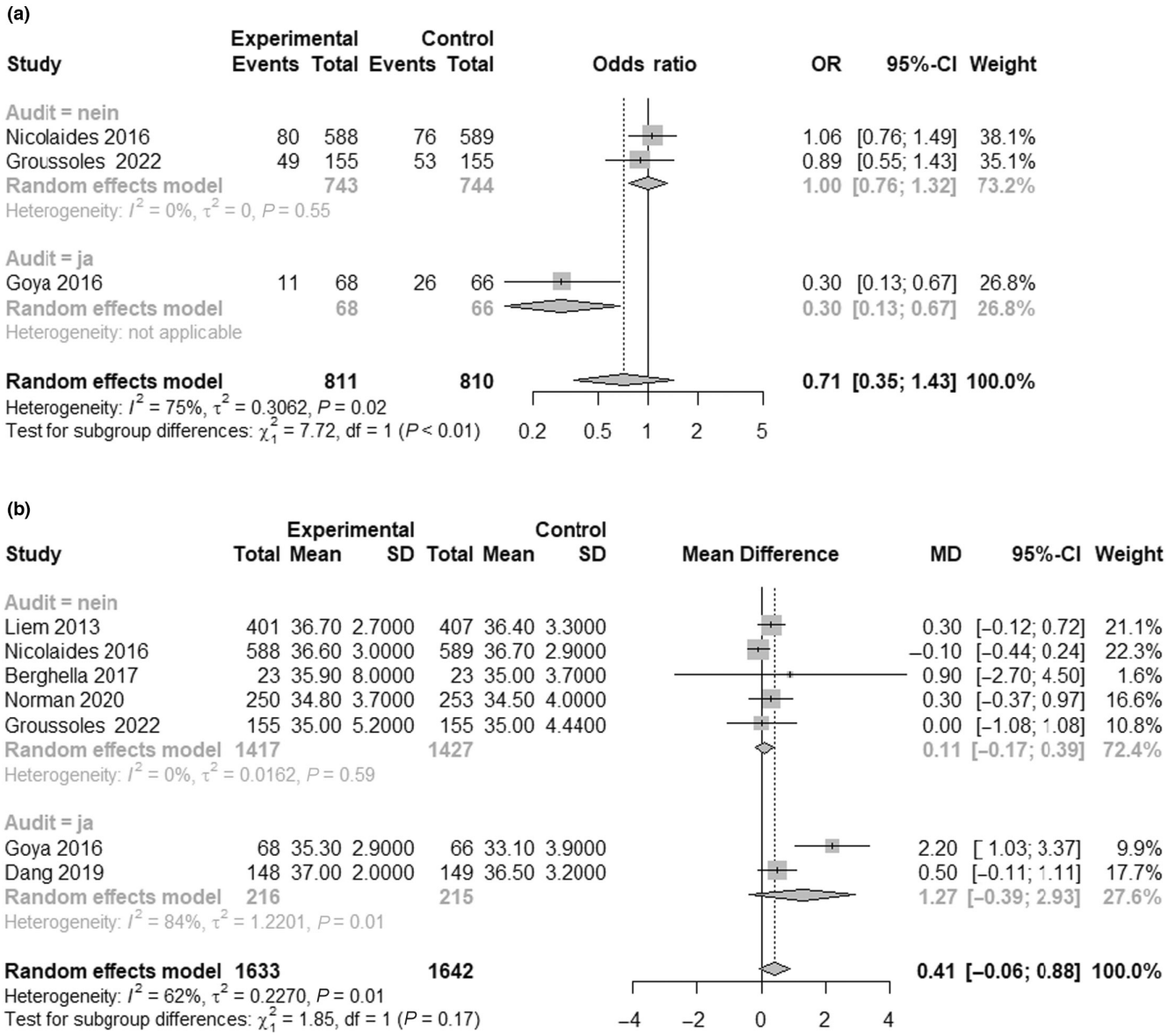

FIGURE 4 (a) Spontaneous preterm birth before 34 weeks of pregnancy in twin pregnancies according to audit procedure. (b) Gestational age at delivery in twin pregnancies with and without pessary, according to audit procedure. (c) Composite neonatal outcome in twin pregnancies with and without pessary, according to audit procedure. (d) Perinatal mortality in twin pregnancies, according to audit procedure.

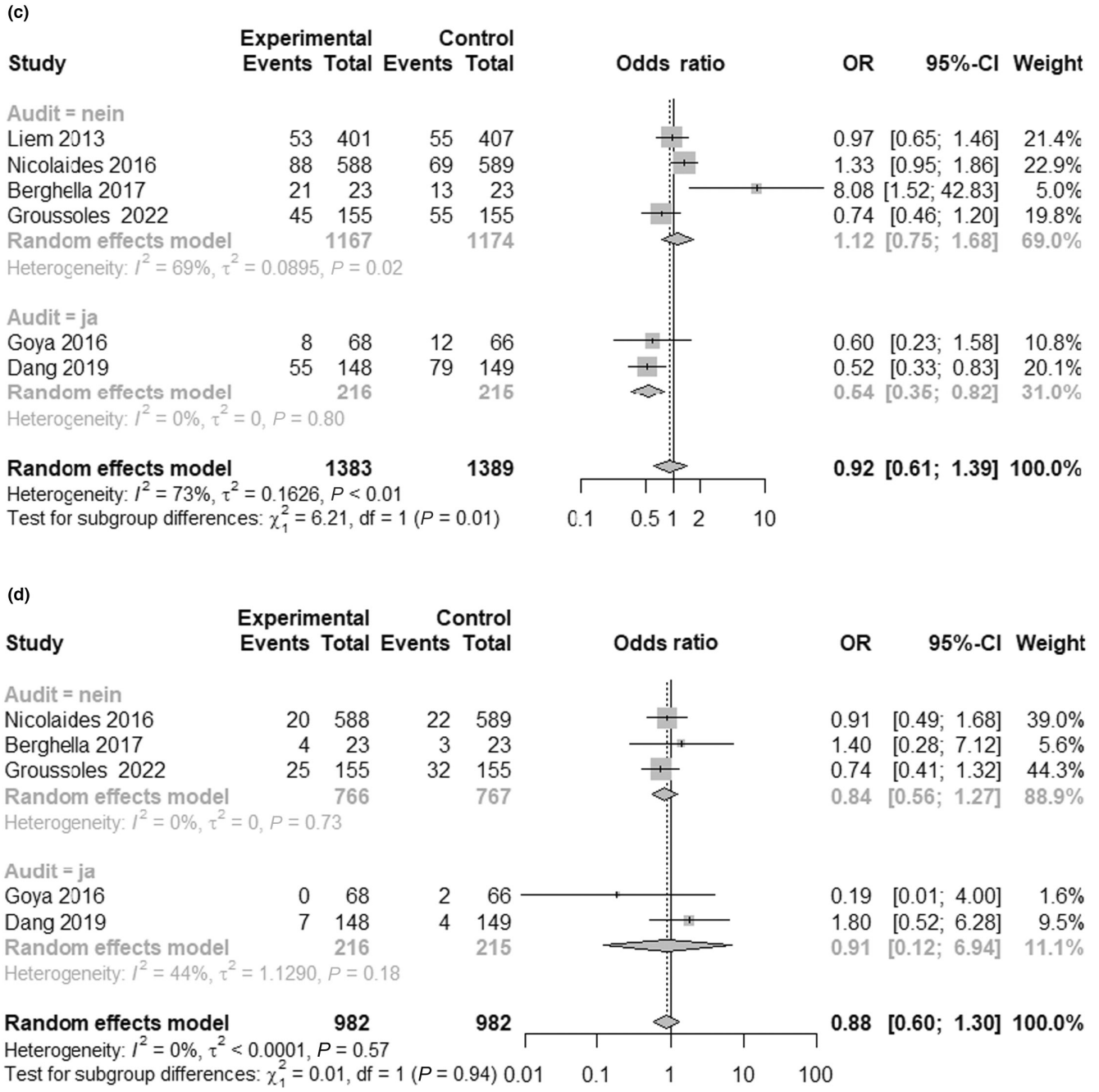

FIGURE 4 (Continued)

This meta-analysis shares similar limitations as previous ones due to its retrospective nature. The pre-defined selection bias by the relatively late delivery rates in control groups can only be narratively analyzed and is suggestive for heterogeneity.

The strength of this analysis is that we sought to compare trials by also categorizing the experience of the participating clinicians, compliance to instructions of medical devices, and audit procedures. Notably, the latter correlated with a high CEP score value, leading to a significant reduction in sPTB and composite neonatal outcome for both singletons and twins.

In conclusion, we scrutinized the importance of MDR criteria, the adherence to the IFUs of cervical pessaries, and criteria of equivalence before performing RCTs or including trials into meta-analyses.

Even in the hands of unexperienced clinicians and no adherence to IFUs, the Arabin cervical pessary demonstrated no safety risks, within these trials or within any corresponding data banks. However, this should not lead to the wrong conclusion that conservative treatment can be performed more liberally than other clinical procedures. Like other procedures, the performance of this therapy finally depends on complex measures and specifications. As a syndrome with multiple causes, threatening PTB demands healthcare providers who possess a profound understanding and genuine empathy for their patients' experiences and needs.

TABLE 4 Secondary obstetric outcomes within audited trials.

Outcome or subgroup	Pessary group	Control group	Odds ratio (95% confidence interval)	Weight
Singleton pregnancy audited trials				
sPTB before 28 + 0 weeks	10/340	15/340	0.40 (0.15–1.08)	40.2%
sPTB before 34 + 0 weeks	23/340	74/340	0.28 (0.12–0.65)	35%
sPTB before 37 + 0 weeks	71/340	162/340	0.31 (0.11–0.83)	43.9%
Tocolysis	64/190	101/190	0.45 (0.30–068)	80.9%
Antenatal corticosteroids	80/190	121/190	0.41 (027–0.63)	58.9%
Cesarean	86/340	97/340	0.85 (0.58–1.24)	88.7%
Vaginal discharge	130/150	69/150	7.63 (4.32–13.49)	53.9%
Multiple pregnancy audited trials				
sPTB before 28 + 0 weeks	9/148	7/149	1.31 (0.48–3.62)	11.9%
sPTB before 34 + 0 weeks	11/68	26/66	0.30 (0.13–0.67)	54.4%
sPTB before 37 + 0 weeks	73/148	91/149	0.62 (0.39–0.98)	35.3%
Tocolysis	22/68	29/66	0.61 (0.30–1.23)	19.2%
Antenatal corticosteroids	25/68	31/66	0.66 (0.33–1.31)	17.4%
Cesarean	156/216	152/215	1.12 (0.70–1.78)	18.1%
Vaginal discharge	172/216	71/215	20.98 (1.48–296.60)	52.2%

Abbreviation: sPTB, spontaneous preterm birth.

Our findings indicate that the comprehensive supervision of a meta-analysis requires stringent application of quality metrics to the evaluated studies, especially for medical devices with defined application criteria. The evaluation of statistical quality metrics alone poses the risk of erroneous conclusions that are not applicable to clinical practice.

AUTHOR CONTRIBUTIONS

Ioannis Kyvernitakis, Ahmet A. Baschat, Werner Rath, Richard Berger, and Holger Maul were responsible for conceptualization, acquisition, analysis, interpretation of data, and writing the manuscript. Marcel Malan, Wolfgang Henrich, and Ekkehard Schleussner were responsible for acquisition, analysis, interpretation of data, and writing the manuscript. Bahareh Yousefi and Nina Timmesfeld were responsible for acquisition, analysis, study design, statistical analyses, interpretation of data, and writing the manuscript.

CONFLICT OF INTEREST STATEMENT

The authors have no conflicts of interest.

DATA AVAILABILITY STATEMENT

The data supporting the findings of this study are available upon reasonable request from the corresponding author.

ORCID

Ioannis Kyvernitakis https://orcid.org/0000-0002-8721-5818

REFERENCES

1. Chawanpaiboon S, Vogel JP, Moller AB, et al. Global, regional, and national estimates of levels of preterm birth in 2014: a systematic review and modelling analysis. *Lancet Glob Health*. 2019;7(1):e37-e46. doi:10.1016/S2214-109X(18)30451-0
2. Patel RM. Short-and long-term outcomes for extremely preterm infants. *Am J Perinatol*. 2016;33(3):318-328. doi:10.1055/s-0035-1571202
3. Romero R, Dey SK, Fisher SJ. Preterm labor: one syndrome, many causes. *Science*. 2014;345(6198):760-765. doi:10.1126/science.1251816
4. Iams JD, Goldenberg RL, Meis PJ, et al. The length of the cervix and the risk of spontaneous premature delivery. National Institute of Child Health and Human Development Maternal Fetal Medicine Unit Network. *N Engl J Med*. 1996;334(9):567-572. doi:10.1056/NEJM199602293340904
5. Salomon LJ, Diaz-Garcia C, Bernard JP, Ville Y. Reference range for cervical length throughout pregnancy: non-parametric LMS-based model applied to a large sample. *Ultrasound Obstet Gynecol*. 2009;33(4):459-464. doi:10.1002/uog.6332
6. Goya M, Pratcorona L, Merced C, et al. Cervical pessary in pregnant women with a short cervix (PECEP): an open-label randomised controlled trial. *Lancet*. 2012;379(9828):1800-1806. doi:10.1016/S0140-6736(12)60030-0
7. Kyvernitakis I, Berger R, Maul H. Letter to the editor: FIGO good practice recommendations on the use of pessary for reducing the frequency and improving outcomes of preterm birth. *Int J Gynaecol Obstet*. 2022;157(1):216-217. doi:10.1002/ijgo.14099
8. Franca MS, Hatanaka AR, Cruz JJ, et al. Cervical pessary plus vaginal progesterone in a singleton pregnancy with a short cervix: an experience-based analysis of cervical pessary's efficacy. *J Matern Fetal Neonatal Med*. 2022;35(25):6670-6680. doi:10.1080/14767058.2021.1919076
9. Nicolaides KH. Nuchal translucency and other first-trimester sonographic markers of chromosomal abnormalities. *Am J Obstet Gynecol*. 2004;191(1):45-67. doi:10.1016/j.ajog.2004.03.090
10. Grey A, Bolland MJ, Avenell A, Klein AA, Gunsalus CK. Check for publication integrity before misconduct. *Nature*. 2020;577(7789):167-169. doi:10.1038/d41586-019-03959-6
11. Bolland MJ, Avenell A, Gamble GD, Grey A. Systematic review and statistical analysis of the integrity of 33 randomized

controlled trials. *Neurology*. 2016;87(23):2391-2402. doi:10.1212/WNL.0000000000003387
12. Prior M, Hibberd R, Asemota N, Thornton JG. Inadvertent P-hacking among trials and systematic reviews of the effect of progestogens in pregnancy? A systematic review and meta-analysis. *BJOG*. 2017;124(7):1008-1015. doi:10.1111/1471-0528.14506
13. UNION TEPATCOTE. REGULATION (EU) 2017/745 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on medical devices, amending directive 2001/83/EC, regulation (EC) No 178/2002 and regulation (EC) No 1223/2009 and repealing Council directives 90/385/EEC and 93/42/EEC. *Off J Eur Union*. 2017;5.5(2017):117-186.
14. Clinical evaluation—Equivalence. 2020.
15. Group MD. Guidance on sampling of MDR Class IIa/Class IIb and IVDR Class B/Class C devices for the assessment of the technical documentation. 2019.
16. Ciolino JD, Diebold A, Jensen JK, Rouleau GW, Koloms KK, Tandon D. Choosing an imbalance metric for covariate-constrained randomization in multiple-arm cluster-randomized trials. *Trials*. 2019;20:293. doi:10.1186/s13063-019-3324-5
17. Karbasian N, Sheikh M, Pirjani R, Hazrati S, Tara F, Hantoushzadeh S. Combined treatment with cervical pessary and vaginal progesterone for the prevention of preterm birth: a randomized clinical trial. *J Obstet Gynaecol Res*. 2016;42(12):1673-1679. doi:10.1111/jog.13138
18. Dugoff L, Berghella V, Sehdev H, Mackeen AD, Goetzl L, Ludmir J. Prevention of preterm birth with pessary in singletons (PoPPS): randomized controlled trial. *Ultrasound Obstet Gynecol*. 2018;51(5):573-579. doi:10.1002/uog.18908
19. Cruz-Melguizo S, San-Frutos L, Martinez-Payo C, et al. Cervical pessary compared with vaginal progesterone for preventing early preterm birth: a randomized controlled trial. *Obstet Gynecol*. 2018;132(4):907-915. doi:10.1097/AOG.0000000000002884
20. Hui SY, Chor CM, Lau TK, Lao TT, Leung TY. Cerclage pessary for preventing preterm birth in women with a singleton pregnancy and a short cervix at 20 to 24 weeks: a randomized controlled trial. *Am J Perinatol*. 2013;30(4):283-288. doi:10.1055/s-0032-1322550
21. Berghella V, Dugoff L, Ludmir J. Prevention of preterm birth with pessary in twins (PoPPT): a randomized controlled trial. *Ultrasound Obstet Gynecol*. 2017;49(5):567-572. doi:10.1002/uog.17430
22. Nicolaides KH, Syngelaki A, Poon LC, et al. A randomized trial of a cervical pessary to prevent preterm singleton birth. *N Engl J Med*. 2016;374(11):1044-1052. doi:10.1056/NEJMoa1511014
23. Nicolaides KH, Syngelaki A, Poon LC, et al. Cervical pessary placement for prevention of preterm birth in unselected twin pregnancies: a randomized controlled trial. *Am J Obstet Gynecol*. 2016;214(1):3.e1-3.e9. doi:10.1016/j.ajog.2015.08.051
24. Norman JE, Norrie J, MacLennan G, et al. Evaluation of the Arabin cervical pessary for prevention of preterm birth in women with a twin pregnancy and short cervix (STOPPIT-2): an open-label randomised trial and updated meta-analysis. *PLoS Med*. 2021;18(3):e1003506. doi:10.1371/journal.pmed.1003506
25. Goya M, de la Calle M, Pratcorona L, et al. Cervical pessary to prevent preterm birth in women with twin gestation and sonographic short cervix: a multicenter randomized controlled trial (PECEP-twins). *Am J Obstet Gynecol*. 2016;214(2):145-152. doi:10.1016/j.ajog.2015.11.012
26. Groussolles M, Winer N, Sentilhes L, et al. Arabin pessary to prevent adverse perinatal outcomes in twin pregnancies with a short cervix: a multicenter randomized controlled trial (PESSARONE). *Am J Obstet Gynecol*. 2022;227(2):271.e1-271.e13. doi:10.1016/j.ajog.2022.01.038
27. Dang VQ, Nguyen LK, Pham TD, et al. Pessary compared with vaginal progesterone for the prevention of preterm birth in women with twin pregnancies and cervical length less than 38 mm: a randomized controlled trial. *Obstet Gynecol*. 2019;133(3):459-467. doi:10.1097/AOG.0000000000003136
28. Liem S, Schuit E, Hegeman M, et al. Cervical pessaries for prevention of preterm birth in women with a multiple pregnancy (ProTWIN): a multicentre, open-label randomised controlled trial. *Lancet*. 2013;382(9901):1341-1349. doi:10.1016/S0140-6736(13)61408-7
29. Kyvernitakis I, Arabin B. Re: prevention of preterm birth with pessary in twins (PoPPT): a randomized controlled trial. *Ultrasound Obstet Gynecol*. 2017;50(3):408-409. doi:10.1002/uog.18808
30. Ting YH, Lao TT, Wa Law L, et al. Arabin cerclage pessary in the management of cervical insufficiency. *J Matern Fetal Neonatal Med*. 2012;25(12):2693-2695. doi:10.3109/14767058.2012.712559
31. Arabin B, Halbesma JR, Vork F, Hubener M, van Eyck J. Is treatment with vaginal pessaries an option in patients with a sonographically detected short cervix? *J Perinat Med*. 2003;31(2):122-133. doi:10.1515/JPM.2003.017
32. Arabin B, Alfirevic Z. Cervical pessaries for prevention of spontaneous preterm birth: past, present and future. *Ultrasound Obstet Gynecol*. 2013;42(4):390-399. doi:10.1002/uog.12540
33. Lees C, Marlow N, Arabin B, et al. Perinatal morbidity and mortality in early-onset fetal growth restriction: cohort outcomes of the trial of randomized umbilical and fetal flow in Europe (TRUFFLE). *Ultrasound Obstet Gynecol*. 2013;42(4):400-408. doi:10.1002/uog.13190
34. Barbone AS, Li X, Arabin B, Kira Y, Jani JC, Cannie MM. Preliminary modeling of effective positioning of Arabin cerclage pessary in women at high risk of preterm birth. *Ultrasound Obstet Gynecol*. 2020;55(4):557-558. doi:10.1002/uog.20375
35. Tajik P, Monfrance M, van't Hooft J, et al. A multivariable model to guide the decision for pessary placement to prevent preterm birth in women with a multiple pregnancy: a secondary analysis of the ProTWIN trial. *Ultrasound Obstet Gynecol*. 2016;48(1):48-55. doi:10.1002/uog.15855
36. Nicolaides KN. Cervical pessary and preterm singleton birth. *N Engl J Med*. 2016;375(6):e10. doi:10.1056/NEJMc1605536
37. Saccone G, Maruotti GM, Giudicepietro A, Martinelli P, Italian Preterm Birth Prevention Working G. Effect of cervical pessary on spontaneous preterm birth in women with singleton pregnancies and short cervical length: a randomized clinical trial. *JAMA*. 2017;318(23):2317-2324. doi:10.1001/jama.2017.18956
38. Conde-Agudelo A, Romero R, Nicolaides KH. Cervical pessary to prevent preterm birth in asymptomatic high-risk women: a systematic review and meta-analysis. *Am J Obstet Gynecol*. 2020;223(1):42-65 e2. doi:10.1016/j.ajog.2019.12.266
39. Cannie MM, Dobrescu O, Gucciardo L, et al. Arabin cervical pessary in women at high risk of preterm birth: a magnetic resonance imaging observational follow-up study. *Ultrasound Obstet Gynecol*. 2013;42(4):426-433. doi:10.1002/uog.12507
40. Seravalli V, Strambi N, D'Arienzo A, et al. Patient's experience with the Arabin cervical pessary during pregnancy: a questionnaire survey. *PloS One*. 2022;17(1):e0261830. doi:10.1371/journal.pone.0261830
41. Kyvernitakis I, Maul H, Bahlmann F. Controversies about the secondary prevention of spontaneous preterm birth. *Geburtshilfe Frauenheilkd*. 2018;78(6):585-595. doi:10.1055/a-0611-5337

How to cite this article: Kyvernitakis I, Baschat AA, Malan M, et al. Cervical pessary to prevent preterm birth and poor neonatal outcome: An integrity meta-analysis of randomized controlled trials focusing on adherence to the European Medical Device Regulation. *Int J Gynecol Obstet*. 2023;00:1-14. doi:10.1002/ijgo.15169