



Pessary placement in the prevention of preterm birth in multiple pregnancies: a propensity score analysis



Maurice J.M. Monfrance^{a,*}, Ewoud Schuit^{b,c,d}, Rolf H. Groenwold^b, Martijn A. Oudijk^e, Irene M. de Graaf^c, Caroline J. Bax^f, Dick J. Bekedam^g, Ben Willem Mol^{c,h}, Josje Langenveld^a

^a Department of Obstetrics and Gynaecology, Zuyderland Medical Center, Heerlen, The Netherlands

^b Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht, The Netherlands

^c Department of Obstetrics and Gynaecology, Academic Medical Center, Amsterdam, The Netherlands

^d Stanford Prevention Research Center, Stanford University, Stanford, CA, USA

^e Department of Obstetrics and Gynaecology, University Medical Center Utrecht, Utrecht, The Netherlands

^f Department of Obstetrics and Gynaecology, VU Medical Center, Amsterdam, The Netherlands

^g Department of Obstetrics and Gynaecology, Onze Lieve Vrouwe Gasthuis, Amsterdam, The Netherlands

^h The Robinson Institute, School of Pediatrics and Reproductive Health, University of Adelaide, Adelaide, Australia

ARTICLE INFO

Article history:

Received 19 August 2015

Accepted 12 November 2015

Keywords:

Pessary
Multiple pregnancy
Short cervix
Preterm birth
Propensity score

ABSTRACT

Objective: In asymptomatic women with a multiple pregnancy and short cervix prophylactic use of a cervical pessary might reduce preterm birth. We assessed the possible treatment effects of pessary use in pregnancy duration and for poor perinatal outcome.

Study design: This cohort study was performed between December 2012 and September 2014 in 44 hospitals in the Netherlands. Women with multiple pregnancy had a cervical length measurement between 16 and 22 weeks of gestation. When cervical length was below 38 mm, women were offered a cervical pessary. The course of pregnancy, including perinatal outcome in these women was compared to the outcome of women from the placebo group of the AMPHIA trial (ISRCTN40512715) (historical cohort). Propensity-score matching with replacement was used to create comparable baseline characteristics between both populations.

Results: We studied 63 women in the pessary group and 56 women as controls. Propensity-score matching generated 57 women in the intervention group matched to 57 women (31 unique) in the control group. Gestational age at delivery was comparable between both groups (HR 0.96, 95%-CI 0.46–1.46) as well as their delivery rates before 28, 32 and 37 weeks, RR 0.68 (95%-CI 0.21–2.18), RR 0.54 (95%-CI 0.21–1.41), and RR 1.22 (95%-CI 0.47–3.15), respectively. There was no difference in composite perinatal outcome (RR 1.36, 95%-CI 0.53–3.51) and perinatal mortality (RR 0.89, 95%-CI 0.24–3.38) either.

Conclusion: In this cohort study with propensity score analysis, pessary use did not prevent preterm birth in asymptomatic women with a multiple pregnancy and short cervix.

© 2015 Elsevier Ireland Ltd. All rights reserved.

Introduction

Preterm birth is the major contributor to perinatal mortality and serious neonatal morbidity. Approximately 30% of all patients admitted to a neonatal intensive care unit in the Netherlands are

born from a multiple pregnancy [1,2]. In women with a multiple pregnancy almost 50% deliver preterm (before the 37th week) in comparison to 6–10% of women with a singleton pregnancy [3]. These numbers indicate the need for interventions that will decrease the number of preterm births in multiple pregnancies.

Cervical length at mid-gestation is a sensitive predictor of spontaneous early preterm delivery [4] and may help identify women who might benefit from early intervention.

Several therapies including routine hospital admission for bed rest [5], 17-alpha-hydroxyprogesterone [4,6] and vaginal progesteron [7,8] were not effective in preventing early preterm birth in

* Corresponding author at: Department of Obstetrics and Gynaecology, Zuyderland Medical Center, Henri Dunantstraat 5, 6419 PC Heerlen, The Netherlands. Tel.: +31 06 55 82 55 03; fax: +31 045 576 66 25.

E-mail address: mjmmonfrance@gmail.com (Maurice J.M. Monfrance).

unselected multiple pregnancies. However, vaginal progesterone might improve perinatal outcome in women with a cervical length below 25 mm [9]. Ultrasonography-indicated use of a cervical cerclage in women with a twin pregnancy is not recommended outside the context of studies, since studies in a small number of women indicate harm [10,11].

Recently, a nationwide multicenter randomized clinical trial, the ProTWIN-trial, showed potential benefit from the use of cervical pessaries in the prevention of preterm birth in women with a multiple pregnancy and a cervix <38 mm (25th percentile) at mid-gestation. Placement of a cervical pessary between 16 and 20 weeks gestation showed a significant reduction in very preterm birth rates (before 28 weeks of gestation) (RR 0.23, 95%-CI 0.06–0.87) and poor perinatal outcome (RR 0.43, 95%-CI 0.19–0.91) as compared to women without a pessary [12]. The current study aimed to determine if a cervical pessary reduces the number of preterm births and improves perinatal outcome in women with a multiple pregnancy and a cervical length below 38 mm at mid-gestation using a propensity score analysis.

Methods

Inclusion

This multicenter prospective cohort study was performed in 7 academic and 37 non-academic hospitals between December 2012 and September 2014. During the study period, these hospitals offered women, with a twin pregnancy between 16⁺⁰ and 22⁺⁰ weeks of gestation, a cervical length measurement as part of routine care and the option of a cervical pessary in case of a cervical length below 38 mm [12]. Henceforth, informed consent was not necessary.

Screening

Cervical length was measured between 16⁺⁰ and 22⁺⁰ weeks of gestation. The transvaginal probe was placed in the anterior fornix of the vagina, after which a sagittal view of the cervix was obtained, with the echogenic endocervical mucosa along the length of the canal. Calipers were used to measure the cervical length, the distance between the triangular area of echodensity at the external ostium and the V-shaped notch at the internal ostium.

Intervention and data collection

Asymptomatic women with a cervical length <38 mm were offered to have a pessary placed. An obstetrician inserted an Arabin pessary between 16⁺⁰ and 22⁺⁰ weeks of gestation in the outpatient clinic. The pessary (CE0482, MED/CERT ISO 9003/EN 46003; Dr. Arabin GmbH and Company, KG; Witten, Germany) is made of soft flexible silicone and available in different sizes.

The pessaries were removed at 36⁺⁰ weeks of gestation, or in case of premature rupture of the membranes, active vaginal bleeding, other signs of preterm labor, or severe patient discomfort.

Participating women were registered in a prospective registry at the start of the intervention. Additional data were obtained by means of case report forms, which were filled in by obstetricians, research nurses or the main researcher at least six weeks post-partum.

Control group

We chose a historic control group of women that had participated in the placebo arm of the AMPHIA trial (ISRCTN40512715), a randomized controlled trial comparing

placebo versus 17-alpha-hydroxyprogesterone in multiple pregnancies that was performed between 2006 and 2009, preceding the ProTWIN trial [17]. The AMPHIA trial randomized women with a multiple pregnancy at a gestational age between 15 and 19 weeks, to either weekly intramuscular injections of 250 mg 17alpha-hydroxyprogesterone caproate, or placebo injections. A transvaginal ultrasound examination for cervical length measurement was performed at randomization. Patients received their first injection between 16 and 20 weeks of gestation, after which they received weekly injections until 36 weeks of gestation or until delivery.

This group of women, the previously mentioned placebo group from the AMPHIA trial, is referred to as the control group. The prospective cohort as the pessary group.

In both the cohort study as well as in the AMPHIA trial, women with a previous spontaneous preterm birth before 34 weeks; serious congenital defects; or death of one or more fetuses; early signs of twin-to-twin transfusion syndrome; or a primary cerclage were excluded from participation.

Outcome measures

Primary outcome measure was time to delivery. Secondary outcome measures were preterm birth before 28, 32 and 37 weeks of gestation and a composite of poor perinatal outcome including severe respiratory distress syndrome (grade 2 or worse), bronchopulmonary dysplasia, intraventricular hemorrhage (grade 2b or worse), necrotizing enterocolitis, proven sepsis, stillbirth and neonatal death [13–16]. Furthermore we determined the number of NICU admissions and we investigated possible side-effects of pessary use.

Data analysis

Propensity score matching was used to control for potential covariate imbalances and to create maximally comparable groups (i.e. comparable regarding baseline characteristics) [18,19]. By using propensity score matching we intended to simulate a RCT setting. The propensity score was estimated by a logistic regression model and is the subject-specific probability of receiving an Arabin pessary conditional on baseline covariate values. Covariates included in the propensity score were selected based on their possible contribution to preterm birth risk and included chorionicity, cervical length at inclusion, history of preterm birth, patient age, parity, mode of conception, and twin or triplet pregnancy.

The pessary group was used as the reference group for matching to ensure as many matches as possible to women in the prospective cohort. One-to-*n* matching with replacement was performed with the nearest Mahalanobis metrics matching within calipers defined by the logit of the propensity score. For each woman who received a pessary, women in the control group with a similar propensity score were identified. The pair with the smallest Mahalanobis distance was then selected. The process was repeated until as many matches as possible could be made for the women in the pessary group.

The standardized mean difference was used to assess the balance of the covariates since it is a property of the sample, and unlike significance testing, does not depend upon the size of the sample [20]. A standardized mean difference greater than the absolute value of 0.1 was used to indicate that the samples were meaningfully different [21].

In order to achieve the best possible match between both groups different calipers were tested: 0.025, 0.05, and 0.1. Additionally, to account for the relative small sample size and thus the potential influence of different matching results due to matching with replacement the matching procedure was

performed five times for each caliper. For each matched set we assessed the number of characteristics for which the standardized mean difference was below 0.1 (lower is better) and the *c*-statistic of the propensity score model in the matched dataset 0.5 (closer to 0.5 is better) [22]. Based on these analyzes a caliper of 0.1 was used.

For the matched cohort, the primary outcome was compared between groups using a Cox proportional hazards model resulting in a hazard ratio (HR) with a 95% confidence interval (CI). Secondary outcomes on the maternal level were assessed using a log-binomial regression model, while outcomes on the child level were assessed using binomial Generalized Estimating Equations with a log-link function and an unstructured covariance matrix to

account for dependence of children within the same mother [23]. Both models will result in relative risks (RR) with 95% CI. The one-to-*n* matching was taken into account by applying more weight to those women from the control group that were matched more than once. Valid estimates of the treatment effect could be obtained without adjustment for the matching procedure or matching variables [24].

Statistical analysis was performed with IBM SPSS Statistics for Windows, Version 22.0 (IBM Corp., Armonk, NY) and R, version 3.1.1. (The R foundation for statistical computing, Vienna, Austria). For matching the MatchIt library was used. A nominal *p*-value <0.05 was considered to be statistically significant.

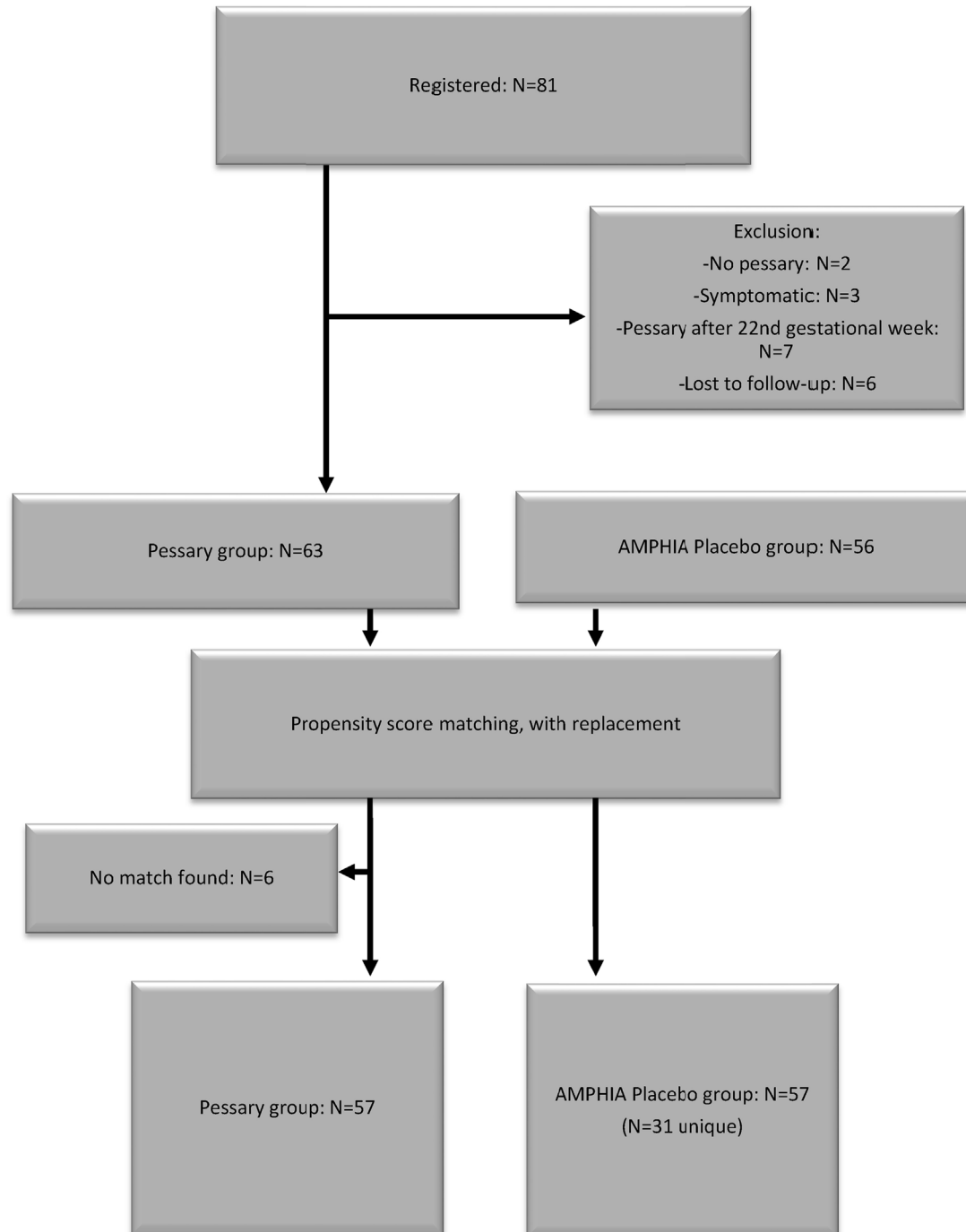


Fig. 1. Trial profile.

Table 1
Baseline characteristics.

	Before matching			After matching		
	Pessary group (N=63)	Control group (N=56)	SMD	Pessary group (N=57)	Control group (N=31)	SMD
Maternal characteristics						
Age (years) ^b	32 (5.6)	31.6 (5.2)	0.076	32 (5.7)	32.1 (6.3)	0.052
Body mass index (kg/m ²) ^a	22.8 (20.2–26.8)	23.1 (21.3–26.2)	0.227	22.8 (20.2–27.5)	23.7 (22.0–24.8)	0.312
Caucasian	38 (63%)	46 (82%)	0.432	35 (65%)	48 (84%)	0.642
Higher education	10 (40%)	18 (44%)	0.079	9 (43%)	13 (43%)	0.039
Smoking during pregnancy	4 (7%)	12 (21%)	0.417	4 (8%)	17 (30%)	0.495
Nulliparous	20 (32%)	21 (38%)	0.121	18 (32%)	28 (49%)	0.083
Previous preterm delivery	5 (8%)	5 (9%)	0.036	4 (7%)	9 (16%)	0.096
Pregnancy characteristics						
Pregnancy after fertility treatment ^c	24 (38%)	22 (39%)	0.024	23 (40%)	16 (28%)	0.169
Triplets	4 (6%)	3 (5%)	0.042	4 (7%)	3 (5%)	0.096
Monochorionic pregnancy	14 (22%)	10 (18%)	0.109	13 (23%)	13 (23%)	0.085
Cervical length at inclusion (mm) ^a	30 (23–34)	34 (29–36)	0.256	30 (24–35)	27 (22–34)	0.001
Gestational age at pessary placement (weeks, days) ^a	20.3 (19.5–20.6)	19.6 (18.9–21.3)	0.350	20.3 (19.5–20.6)	19.7 (19.1–21.7)	0.502

Data are presented as N (%), ^a median (IQR), or ^b median (SD).

^c Ovarian hyperstimulation, in vitro fertilization, intracytoplasmic sperm injection, or intrauterine insemination. Standardized mean difference (SMD).

Results

Participants

Between December 2012 and September 2014, 81 women with a twin pregnancy and a cervical length <38 mm were registered as potential participants. Twelve women were excluded from final analysis: two women never had a pessary placed, three women were symptomatic and had contractions at the time of cervical length measurement or pessary placement and in seven women pessary placement was performed after the 22nd week of gestation. For six women data were irretrievable. The remaining 63 patients were eligible for inclusion and analysis (Fig. 1).

In the AMPHIA-trial, 335 women received placebo, of which 298 underwent cervical length measurement and 56 had a cervical length <38 mm, thereby confining the control group.

Baseline characteristics are presented in Table 1. Before matching baseline characteristics were different as indicated by the different distributions of the propensity score in both groups (Fig. 2A) and by the standardized mean difference criterion, which

revealed that 5 out of 12 baseline covariates had a standardized mean difference ≤ 0.10 (Table 1).

Out of 63 women in the pessary group 57 were matched with 31 women in the control group. Because of the matching with replacement some women in the control group were used multiple times (range 1–4). Matching showed to be successful with similar propensity score distributions (Fig. 2B) and a higher number of baseline covariates with a standardized difference ≤ 0.10 (7 out of 12; Table 1). The c-statistic when fitting the propensity score in the matched dataset was 0.54.

Demography

After matching, women in the pessary group had a mean age of 32.0 (SD 5.7) years, 73% had a dichorionic diamniotic twins, 32% was nulliparous and 7% had a previous preterm birth <37 weeks. Mean cervical length at inclusion was 30 mm (IQR 24–35).

Women in our control group differed mainly from the pessary group in percentage of Caucasian women (84% versus 65%) and the number of smokers (30% versus 8%).

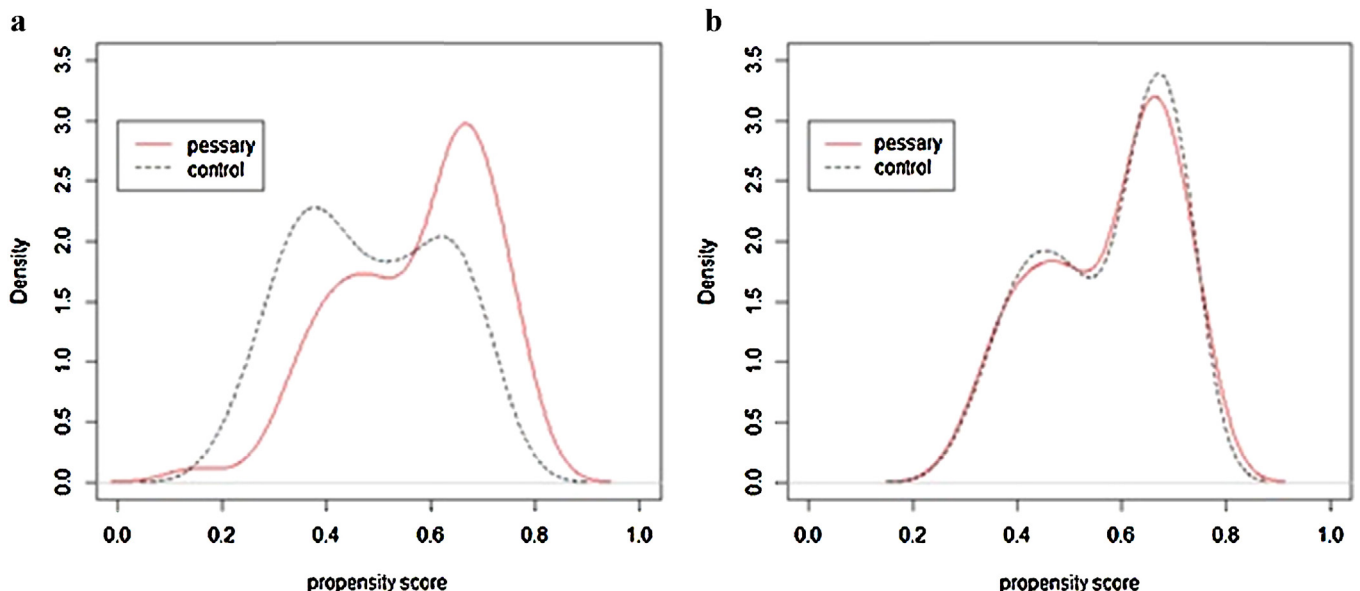


Fig. 2. Propensity scores, subject-specific probability of receiving a cervical pessary conditional on baseline covariate values (A) before matching and (B) after matching.

Table 2
Pregnancy and neonatal outcome measures.

	Before matching		After matching		RR (95%-CI)
	Pessary group (N=63)	Control group (N=56)	Pessary group (N=57)	Control group (N=57) ^b	
Maternal outcome measures					
Gestational age at delivery (weeks, days)	34.6 (32.8–37.0)	35.4 (33.0–37.4)	35.0 (32.5–37.0)	33.6 (31.9–37.0)	0.96 (0.46–1.46) ^a
<28 weeks	8 (13)	7 (12)	8 (14)	7 (12)	0.68 (0.21–2.18)
<32 weeks	14 (22)	12 (21)	13 (23)	20 (35)	0.54 (0.21–1.41)
<37 weeks	46 (73)	37 (66)	41 (72)	42 (74)	1.22 (0.47–3.15)
	Pessary group (N=130)	Control group (N=115)	Pessary group (N=118)	Control group (N=117) ^b	RR (95%-CI)
Neonatal outcome measures					
Birth weight					
<2500 g	96 (74)	78 (68)	86 (73)	91 (78)	0.73 (0.28–1.89)
<1500 g	31 (24)	24 (21)	27 (23)	20 (17)	1.64 (0.58–4.67)
Composite adverse perinatal outcome	34 (26)	27 (24)	33 (28)	27 (24)	1.36 (0.53–3.51)
Respiratory distress syndrome	18 (14)	13 (11)	18 (15)	16 (14)	1.20 (0.40–3.59)
Bronchopulmonary dysplasia	4 (3)	2 (2)	4 (3)	3 (3)	1.10 (0.15–8.06)
Intraventricular hemorrhage	2 (2)	1 (1)	2 (2)	2 (2)	0.99 (0.06–16.21)
Necrotizing enterocolitis	3 (2)	1 (1)	2 (2)	0 (0)	NC
Sepsis	7 (6)	0 (0)	7 (6)	0 (0)	NC
Mortality	9 (7)	13 (11)	9 (8)	11 (9)	0.49 (0.14–1.74)
Stillbirth	1 (1)	1 (1)	1 (1)	1 (1)	0.93 (0.06–15.00)
Death before discharge	8 (6)	12 (11)	8 (7)	10 (9)	0.85 (0.22–3.23)
Admission to neonatal intensive care unit	31 (24)	24 (21)	29 (25)	39 (33)	0.66 (0.20–2.14)

Data are presented as N (%) or median (IQR). NC = not calculated.

^a Hazard ratio instead of RR.

^b Women and children that were matched more than once are included more than once for the calculation of median and totals.

Primary outcome

Median gestational age at delivery in the pessary group was 35⁺⁰ (IQR 32⁺⁴–37⁺⁰) versus 33⁺⁴ (IQR 31⁺⁶–37⁺⁰) in the control group (HR 0.96, 95% CI 0.46–1.46). (Table 2 and Fig. 3).

Secondary outcomes

No differences were found between groups for delivery before 28, 32 and 37 weeks of gestation (Table 2). Adverse perinatal outcome occurred in 33 out of 118 children (28%) in the pessary group and 27 out of 117 (24%) in our controls (RR 1.36, 95%-CI 0.53–3.51). Perinatal mortality was comparable in the pessary group and control group with 9 out of 118 (8%) versus 11 out of 117 (9%), respectively (RR 0.89, 95%-CI 0.24–3.38). The number of NICU admissions was not different between the two groups either.

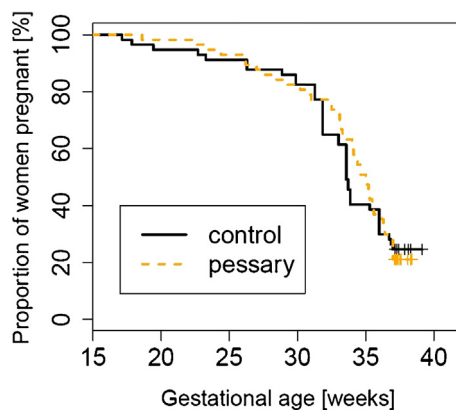


Fig. 3. Kaplan–Meier curve illustrating the proportion of continued pregnancies in women with a cervical length below 38 mm for the intervention (cervical pessary) as well as the non-intervention group.

Side-effects and pessary removal

The main complaints in the total cohort of women treated with a cervical pessary were pain or discomfort (8.3%) and vaginal discharge (50%). Other reported side-effects which might possibly be contributed to pessary use were: more frequent voiding, unwanted loss of urine and sacral pressure.

Forty-two (67%) women treated with a pessary had their pessaries removed before the 36th week of gestation. Median gestational age at pessary removal was 34 weeks and 4 days (IQR 32⁺⁰–36⁺⁰). Of these, 14% had vaginal blood loss, 33% had ruptured membranes, 26% had contractions and 2.3% had their pessary removed before the 36th week because of pain (not otherwise specified) or other discomforts such as excessive discharge (Table 3).

Discussion

This cohort study with propensity score analysis showed that pessary use did not effectively prevent preterm birth in asymptomatic women with a multiple pregnancy and a cervical length <38 mm. Time to delivery, as well as poor perinatal outcome, did not differ compared to the control population. Pessary use did not appear to have negative effects either.

Table 3
Reasons for pessary removal.^a

	Reason for pessary removal (N=42)
Pain/discharge	1 (2)
Vaginal blood loss	6 (14)
Other	16 (38)
PPROM	14 (33)
Contractions	11 (26)

Data are presented as N (%).

^a Components may add up to more than the group total since women may have had more than one reason for removal.

This study was the first published cohort study on pessaries in the prevention of preterm birth. Randomized controlled trials showed alternating, but promising results in the effectiveness of pessary use in the prevention of preterm birth. The ProTWIN-trial showed that use of a cervical pessary in multiple pregnancies with a cervical length <38 mm significantly reduced perinatal mortality (RR 0.13, 95%-CI 0.03–0.60), morbidity (RR 0.43, 95%-CI 0.19–0.91) and extreme preterm birth (<28 weeks of gestation) (RR 0.23, 95%-CI 0.06–0.87) [12]. The PECEP-TWIN-trial also showed significant reductions in preterm birth, but mainly before the 34th week of gestation in women with a cervical length below 25 mm (16.2% versus 25.7%, $P=0.0001$). There appeared to be no differences in mortality or morbidity rates [25].

We decided to perform this cohort study, hoping it would give us valuable information about the effectiveness of pessary therapy after implementation in daily practice. Despite substantial global investment in the commissioning of health service research and clinical guidelines to support decision-making, evidence suggests that whilst the transfer of research to practice is possible, its success can be variable. Mainly due to unwarranted variations in practice and in the resulting outcomes which cannot be explained by characteristics of the patients. Examples are a patients context, as well as interactions between multiple, interconnected factors at the level of individuals, groups, organizations and wider health systems [26]. These factors are probably less dominant in randomized clinical trials.

We intended to resemble a randomized controlled clinical trial by using propensity score matching. However, unlike a randomized trial, propensity score techniques can balance observed covariates but cannot control for unmeasured covariates [19]. Henceforth, this might have influenced outcome measures in a positive or negative way. Additionally, we had a relatively small sample size. As a result, this study was underpowered to detect differences in perinatal outcomes between both groups.

However, this study was possibly large enough to detect a treatment effect as seen in subgroup analysis of the previous ProTWIN study (78 multiple pregnancies).

Nevertheless, we have to state that the effectiveness of pessary use in multiple pregnancies with a short cervix remains unclear. We cannot properly explain the difference in results of this study considering gestational age at delivery compared to the previously mentioned RCTs, especially the ProTWIN. Possible explanations might be the study's design and the previously mentioned difference in sample size. Furthermore there might have been a different number of hidden symptomatic women at inclusion. Another explanation may lie in different baseline characteristics between the population used in this study and the ProTWIN, especially the high percentage of non-Caucasian women in our pessary group (37% versus 9%) and the higher number of smokers in this study's control group (30% versus 6%). It remains unclear whether or not these differences are significant.

This further emphasizes the need for proper evaluation of therapies before implementation, and the need for further research to determine the effects of pessary use itself, especially since pessaries are easily applicable in everyday practice, cheap (€38 per pessary) and no negative effects on perinatal outcome or time to delivery have been reported so far. However, pessary use may increase vaginal discharge and there is a risk for cervical necrosis if the pessary is not removed at a proper time before delivery [27].

In summary, we could not confirm the positive effect of pessary therapy in the prevention of preterm birth in asymptomatic multiple pregnancies with a short cervix in this propensity score analysis. With ongoing studies on this subject the definite answer whether pessaries are effective in the prevention of preterm birth will come. If effective, evaluation of implementation, similar to this study, is still necessary.

Acknowledgements

BWM and JL designed the study. MJMM and JL coordinated the main trial. MJMM collected data. RHG advised on data analysis. ES analyzed the data. MJMM, JL, BWM and ES interpreted the data. MJMM and JL wrote the first drafts of the manuscript. All authors critically revised the first drafts, and approved the final version.

This article was not funded, nor sponsored.

References

- [1] Dutch National Obstetric Registration 2002 Prismaant; 2002.
- [2] Dutch National Neonatal Registration 2002 Prismaant; 2002.
- [3] Schaaf JM, Mol BW, Abu-Hanna A, Ravelli AC. Trends in preterm birth: singleton and multiple pregnancies in the Netherlands, 2000–2007. *BJOG* 2011;118:1196–204.
- [4] Lim AC, Schuit E, Papatsonis D, et al. Effect of 17-alpha hydroxyprogesterone caproate on cervical length in twin pregnancies. *Ultrasound Obstet Gynecol* 2012;40:426–30.
- [5] Crowther CA. Hospitalisation and bed rest for multiple pregnancy. *Cochrane Database Syst Rev* 2000;2:CD000110.
- [6] Senat MV, Porcher R, Winer N, et al. *Am J Obstet Gynecol* 2013;208:194E1–E8.
- [7] Norman JE, Mackenzie F, Owen P, et al. Progesteron for prevention of preterm birth in twin pregnancy (STOPPIT): a randomised, double-blind, placebo-controlled study and meta-analysis. *Lancet* 2009;373:2034–40.
- [8] Rode L, Klein K, Nicolaides KH, Krampfl-Bettelheim E, Tabor A. Prevention of preterm delivery in twin gestations (PREDICT): a multicenter, randomized, placebo-controlled trial on the effect of vaginal micronized progesterone. *Ultrasound Obstet Gynecol* 2011;38:272–80.
- [9] Schuit E, Stock S, Rode L, et al. Effectiveness of progestogens to improve perinatal outcome in twin pregnancies: an individual participant data meta-analysis. *BJOG* 2015;122(1):27–37.
- [10] Zork N, Biggio J, Tita A, Rouse D, Gyamfi-Bannerman C. Decreasing prematurity in twin gestations: predicaments and possibilities. *Obstet Gynecol* 2013;122:375–9.
- [11] Berghella V, Odibo AO, To MS, Rust OA, Althuisius SM. Cerclage for short cervix on ultrasonography: meta-analysis of trials using individual patient-level data. *Obstet Gynaecol* 2005;106:181–9.
- [12] 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;382:1341–9.
- [13] Papile LA, Burstein J, Burstein R, Koffler H. Incidence and evolution of subependymal and intraventricular hemorrhage: a study of infants with birth weights less than 1500 g. *J Pediatr* 1978;92:529–34.
- [14] Jobe AH, Bancalari E. Bronchopulmonary dysplasia. *Am J Respir Crit Care Med* 2001;163:1723–9.
- [15] Giedion A, Haelfliger H, Dangel P. Acute pulmonary X-ray in hyaline membrane disease treated with artificial ventilation and positive end-expiratory pressure (PEP). *Pediatr Radiol* 1973;1:145–52.
- [16] Bell MJ, Ternberg JL, Geigin RD, et al. Neonatal necrotizing enterocolitis: therapeutic decisions based upon clinical staging. *Ann Surg* 1978;187:1–7.
- [17] Lim AC, Schuit E, Bloemenkamp K, et al. 17 α -Hydroxyprogesterone caproate for the prevention of adverse neonatal outcome in multiple pregnancies: a randomized controlled trial. *Obstet Gynecol* 2011;118(3):513–20.
- [18] Rosenbaum PR, Rubin DB. The bias due to incomplete matching. *Biometrics* 1985;41(1):103–16.
- [19] Gilbert SA, Grobman WA, Landon MB, et al. Elective repeat cesarean delivery compared with spontaneous trial of labor after a prior cesarean delivery: a propensity score analysis. *Am J Obstet Gynecol* 2012;206(4):311 e1–3119e.
- [20] Austin PC. A critical appraisal of propensity-score matching in the medical literature between 1996 and 2003. *Stat Med* 2008;27(12):2037–49.
- [21] Normand ST, Landrum MB, Guadagnoli E, et al. Validating recommendations for coronary angiography following acute myocardial infarction in the elderly: a matched analysis using propensity scores. *J Clin Epidemiol* 2001;54(4):387–98.
- [22] Franklin JM, Rassen JA, Ackermann D, Bartels DB, Schneeweiss S. Metrics for covariate balance in cohort studies of causal effects. *Stat Med* 2014;10(33(1)):1685–99.
- [23] Gates S, Brocklehurst P. How should randomised trials including multiple pregnancies be analysed? *BJOG* 2004;111(3):213–9.
- [24] Sjölander A, Greenland S. Ignoring the matching variables in cohort studies – when is it valid and why? *Stat Med* 2013;32(27):4696–708.
- [25] Goya M, Rodo C, Muñoz B, et al. Cervical pessary to prevent preterm birth in twin's pregnant women with a short cervix: a multicentre randomised controlled trial (PECEP-Twins). Barcelona: Paper presented at: 13th World congress in fetal medicine; September 14–17; 2014.
- [26] Hadlington S, Makarow M, Højgaard L, et al. (European Science Foundation). Forward look – implementation of medical research in clinical practice.
- [27] Arabin B, Helbesma JR, Vork F, Hübener M, van Eyck J. Is treatment with vaginal pessaries an option in patients with a sonographically detected short cervix? *J Perinat Med* 2003;31:122–33.