OPEN ACCESS

Effect of hands-on interprofessional simulation training for local emergencies in Scotland: the THISTLE stepped-wedge design randomised controlled trial

Erik Lenguerrand ![ORCID],[1] Cathy Winter,[2] Dimitrios Siassakos,[3] Graeme MacLennan,[4] Karen Innes,[4] Pauline Lynch,[5] Alan Cameron,[6] Joanna Crofts,[2] Alison McDonald,[4] Kirsty McCormack,[4] Mark Forrest,[4] John Norrie,[7] Siladitya Bhattacharya,[8] Tim Draycott[1,2]

Linked

Check for updates

ABSTRACT

Objective To assess whether the implementation of an intrapartum training package (PROMPT (PRactical Obstetric Multi-Professional Training)) across a health service reduced the proportion of term babies born with Apgar score <7 at 5 min ($<7^{5mins}$).

Design Stepped-wedge cluster randomised controlled trial.

Setting Twelve randomised maternity units with ≥900 births/year in Scotland. Three additional units were included in a supplementary analysis to assess the effect across Scotland. The intervention commenced in March 2014 with follow-up until September 2016.

Intervention The PROMPT training package (Second edition), with subsequent unit-level implementation of PROMPT courses for all maternity staff.

Main outcome measures The primary outcome was the proportion of term babies with Apgar$<7^{5mins}$.

Results 87 204 eligible births (99.2% with an Apgar score), of which 1291 infants had an Apgar$<7^{5mins}$ were delivered in the 12 randomised maternity units. Two units did not implement the intervention. The overall Apgar$<7^{5mins}$ rate observed in the 12 randomised units was 1.49%, increasing from 1.32% preintervention to 1.59% postintervention. Once adjusted for a secular time trend, the 'intention-to-treat' analysis indicated a moderate but non-significant reduction in the rate of term babies with an Apgar scores $<7^{5mins}$ following PROMPT training (OR=0.79 95%CI(0.63 to 1.01)). However, some units implemented the intervention earlier than their allocated step, whereas others delayed the intervention. The content and authenticity of the implemented intervention varied widely at unit level. When the actual date of implementation of the intervention in each unit was considered in the analysis, there was no evidence of improvement (OR=1.01 (0.84 to 1.22)). No intervention effect was detected by broadening the analysis to include all 15 large Scottish maternity units. Units with a history of higher rates of Apgar$<7^{5mins}$ maintained higher Apgar rates during the study (OR=2.09 (1.28 to 3.41)) compared with units with pre-study rates aligned to the national rate.

Conclusions PROMPT training, as implemented, had no effect on the rate of Apgar $<7^{5mins}$ in Scotland during the study period. Local implementation at scale was found to be more difficult than anticipated. Further research is required to understand why the positive effects observed in other single-unit studies have not been replicated in Scottish maternity units, and how units can be best supported to locally implement the intervention authentically and effectively.

Trial registration number ISRCTN11640515.

INTRODUCTION

Maternity care in the UK could be safer.[1–3] Perinatal outcomes are substantially worse than in countries of similar GDP[4] and more than 50% of adverse pregnancy outcomes are considered preventable.[5–7]

Training for intrapartum emergencies has been recommended to improve care[3–6 8–10] and mandated by litigation authorities[11 12] but data regarding effective intrapartum training are conflicting. Some training programmes have been associated with significant safety improvements,[13–18] whereas other initiatives have reported either no change[16 19] or a deterioration in outcomes.[15] Observational studies of a multiprofessional intrapartum emergencies training course for local maternity staff—PRactical Obstetric Multi-Professional Training—(PROMPT, Bristol, UK) have previously demonstrated improved compliance with clinical standards,[13 17] reductions in clinical error[14] and sustainable improvements in perinatal outcomes.[20] These positive effects were initially observed in a single maternity unit but have since been repeated across other

international settings.[21–23] Although these early data are extremely encouraging, it is important to determine whether these improvements could be replicated at scale across a national health system using a robust study design such as a randomised controlled trial (RCT).

With over 50 000 births annually,[24] Scotland provided an ideal setting to investigate the effect of implementing PROMPT training at national scale. Only 3 of the 15 Scottish maternity units (≥900 births annually—2012) had previously undertaken PROMPT training. Moreover, all NHS maternity outcomes are routinely collated centrally by the Information Statistics Division Scotland (ISD), providing a robust data set for analysis. Finally, the proportion of term babies with Apgar <7 at 5 min (Apgar<7^{5mins}),[25] an outcome that has previously been shown to be reduced following PROMPT training, was higher in Scotland than in other settings[13] (1.19% between 2004 and 2009, data from ISD Scottish Morbidity Record (SMR 02).

The aim of this study was to conduct an RCT to investigate whether the implementation of an intrapartum emergencies training package across all large maternity units within an entire health service could reduce the proportion of term babies with an Apgar score <7^{5mins}.

METHODS
The protocol for this trial has been published previously.[26]

Intervention
The intervention was the PROMPT training package (Second edition), which included a 2 day PROMPT Train the Trainers (T3) programme (TIDieR checklist, online supplementary appendix 1), with subsequent unit-level implementation of local PROMPT courses for all maternity staff. PROMPT courses are 'in-house', adaptable multiprofessional training days which cover the management of core obstetric emergencies, including postpartum haemorrhage, sepsis and shoulder dystocia as well as fetal monitoring in labour and team-working. Each participating maternity unit identified a multiprofessional 'in-house' training team (recommended: two midwives, one obstetrician and one anaesthetist) to attend the T3 programme at The Scottish Clinical Simulation Centre in Larbert, before the start of their allocated intervention step. The T3 programme included a demonstration 'PROMPT Course' and a T3 day to guide 'in house' trainers in how to use the PROMPT Course-in-a-Box (Second Edition) and set up local PROMPT courses.[27 28] The 'in-house' trainers were asked to implement local PROMPT courses, aiming to train all their maternity staff within 12 months of commencing local training.

Study design, setting and participants
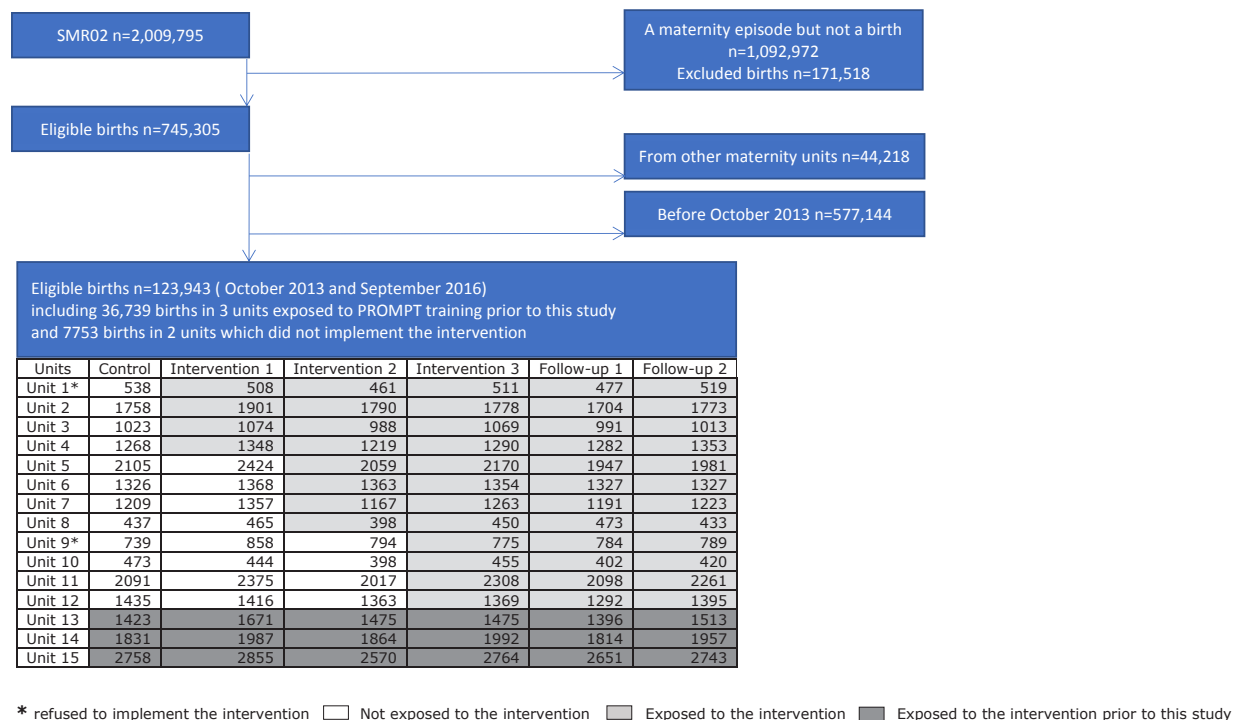We planned a cross-sectional design stepped-wedge cluster randomised controlled superiority trial (SW-RCT). This was viewed as the safest, most pragmatic and ethical option, with the additional rigour of randomisation.

The study comprised a control period when none of the units received training, three intervention periods, and at least two follow-up periods (figure 1). Four maternity units were randomised to each intervention step. Each period lasted 6 months to allow all 12 units to receive their intervention within 18 months. This would reduce potential contamination between trained and untrained units (rotation of trained staff between units) within the pragmatic funding and time constraints.

Maternity units in NHS Scotland with ≥900 births/year (n=15) that had not previously participated in the PROMPT T3 programme were eligible (n=12) and invited to participate through the Scottish Committee of the Royal College of Obstetricians and Gynaecologists. The three units that had already been conducting PROMPT training prior to this study were included in a supplementary preplanned analysis to assess the intervention at national scale. Units with <900 births/year were excluded (<6% of births between 01/01/2000 and 30/09/2013), as their staffing structures may have differed from the larger units. Babies born preterm, at home, at other hospitals and then transferred to a participating unit were all excluded, as were those confirmed to have an intrauterine death prior to labour. Babies born by elective caesarean section (CS) were also excluded as their care was not impacted by the intrapartum training provided by PROMPT. The first four participating maternity units attended a T3 programme in March 2014. The 6-month period preceding the start of the first intervention period ('Intervention 1') (01/10/2013–01/04/2014) was identified as the control period ('Control'). The study ended with the completion of the last follow-up period ('Follow-up period 2') on the 30/09/2016.

Randomisation, concealment and blinding
One eligible maternity unit that had attended PROMPT T3 training before the start of the study but had not commenced 'in-house' local PROMPT courses, was allocated to period 1 for inclusion in the study. The remaining participating units were allocated to an intervention period using the 'imbalance statistic' method[29 30] to balance units by size (small, medium, large annual births). An independent statistician from the Centre for Healthcare Randomised Trials (CHaRT) randomly selected an allocation sequence list from a subset with the most desirable balance properties using computer-generated random numbers. CHaRT then informed the nominated in-house training coordinator from each unit of the date that their team had been allocated to attend the T3 programme. Participating units that were unable to start training at their allocated time slot were manually reassigned to another

| SMR02 n=2,009,795 | | A maternity episode but not a birth n=1,092,972 Excluded births n=171,518 |

| Eligible births n=745,305 |

| From other maternity units n=44,218 |

| Before October 2013 n=577,144 |

Eligible births n=123,943 (October 2013 and September 2016)
including 36,739 births in 3 units exposed to PROMPT training prior to this study
and 7753 births in 2 units which did not implement the intervention

Units	Control	Intervention 1	Intervention 2	Intervention 3	Follow-up 1	Follow-up 2
Unit 1*	538	508	461	511	477	519
Unit 2	1758	1901	1790	1778	1704	1773
Unit 3	1023	1074	988	1069	991	1013
Unit 4	1268	1348	1219	1290	1282	1353
Unit 5	2105	2424	2059	2170	1947	1981
Unit 6	1326	1368	1363	1354	1327	1327
Unit 7	1209	1357	1167	1263	1191	1223
Unit 8	437	465	398	450	473	433
Unit 9*	739	858	794	775	784	789
Unit 10	473	444	398	455	402	420
Unit 11	2091	2375	2017	2308	2098	2261
Unit 12	1435	1416	1363	1369	1292	1395
Unit 13	1423	1671	1475	1475	1396	1513
Unit 14	1831	1987	1864	1992	1814	1957
Unit 15	2758	2855	2570	2764	2651	2743

* refused to implement the intervention ☐ Not exposed to the intervention ☐ Exposed to the intervention ☐ Exposed to the intervention prior to this study

Figure 1: **Consort flow diagram of the Thistle Stepped-Wedge trial**
A more detailed Consort flow diagram is provided in appendix 4

Figure 1 Consort flow diagram of the Thistle Stepped-Wedge trial. A more detailed Consort flow diagram is provided in online supplementary appendix 4.

period, maintaining balance in the size of the maternity units in each period.

Pregnant women were not made aware of maternity unit participation in the study, and any additional training that each unit would normally undertake continued during the study period. Because of the nature of the intervention and study design, the staff and research teams could not be blinded.

Primary outcome

The primary outcome was the Apgar score of eligible term infants measured at 5 min after birth. An Apgar score $<7^{5mins}$ is associated with poor long-term outcomes,[31] [32] can be improved by best care[13] and is independent of maternal demographics.[31] No secondary outcomes were investigated.

Data sources

Birth information was obtained from Information Services Division (ISD) Maternity Inpatient and Day Case dataset (SMR 02). Data were accessed remotely using the National Services Scotland Safe Haven (Access permission: XRB13180). Clinical and individual information were extracted to obtain the date of birth (used to identify the study period to which a birth should be linked), Apgar score at 5 min and variables required to identify the relevant population of term infants such as the type of birth (vaginal, emergency or elective CS), gestation (<37,≥37 weeks), place of birth (at home, transfer from a non-NHS unit,

transfer from a NHS unit, in the unit of interest) and intrauterine deaths.

Each maternity unit provided process information about their local PROMPT courses using a standardised data collection sheet (online supplementary appendix 2). Each unit's training coordinator supplied the dates of their local courses, and anonymised data on numbers of attendees by grade and clinical role. They also submitted course programmes that provided information on the actual course content and degree of PROMPT implementation, and therefore the fidelity of the intervention at unit level. The training coordinators were also requested to provide some baseline data on local staff composition and volume, with the aim of determining the cumulative proportion of staff trained in each unit during the study period.

Statistical analysis

Assuming four maternity units were randomised at each intervention period, with three intervention periods (steps 1–3) and two follow-up periods (steps 4 and 5) each comprising 6 months, an average cluster size of 1200 births and an intracluster correlation coefficient of 0.1, we estimated we could detect a relative reduction of 35% or more in the Apgar$<7^{5mins}$ (1.18%–0.77%) with 80% power and two-sided alpha of 5%.[33] The 12 eligible Scottish maternity units had a rate of Apgar$<7^{5mins}$ at term of 1.18% in 2010, and a monthly average of 200 births per maternity unit, that is, around 1200 births per period duration.

We compiled the prevalence rates of Apgar$<7^{5mins}$ by maternity unit and/or time periods as the number of eligible babies with Apgar$<7^{5mins}$ divided by the total number of eligible births.

Births recorded between 01/01/2000 and 30/09/2013 were used to identify the pre-study context, that is, unit preintervention Apgar scores (online supplementary appendix 3) and birth volumes. An initial intention-to-treat (ITT) analysis was undertaken across the 12 randomised units. Exposure to the intervention was defined using the randomisation schedule. Units that were unable to attend the T3 programme, comply with the randomisation schedule or were unable to implement any local training during the study period, were analysed according to their allocated intervention step. We investigated the intervention effect (exposed/unexposed to PROMPT) using logistic mixed regression (with random effects for clustering of births within unit), adjusted for time-modelled as a categorical variable (five dummy variables for the six periods (table 1) with the control period as the reference) to account for a potential confounding effect of an underlying time-trend in the outcome.[34 35] This model provided the time-averaged intervention effect.[33 35 36] To account for heterogeneity in the pre-study Apgar rates between units and the unit size, a factor used in the minimisation randomisation, this main analysis was also adjusted for pre-study maternity unit size (small <2000, medium 2000–5000 (reference), large >5000 births/year), as well as status of pre-study Apgar score$<7^{5mins}$ rate (below, above, at national average (reference), online supplementary appendix 3). This analysis was conducted on complete-cases but was then repeated to account for the 866 infants (0.7%) with missing Apgar scores. All births with missing score were first considered to be comparable to those with a 'high' Apgar score (good to excellent baby condition) and assigned into the Apgar$\geq7^{5mins}$ group (Imputation 1). The 164 babies with missing scores admitted to a neonatal care unit and/or who received some form of resuscitation other than facial oxygen were then assigned to the Apgar score$<7^{5mins}$ group; the remaining 702 births with missing data were assigned to the Apgar score$\geq 7^{5mins}$ group (Imputation 2).

To account for departures from the randomisation schedule, an 'as-implemented' analysis was conducted replicating the above analyses but using the date of the first local training to define exposure to the intervention, that is, the time periods during which births were actually 'exposed and unexposed' to the intervention (figure 2). Data from randomised units that did not implement the intervention were considered as unexposed throughout the study.

To provide an assessment of the intervention effect across the health service, all 15 Scottish maternity units with ≥900 births/year, including the three non-randomised units that had implemented PROMPT training prior to this study, were analysed in an exploratory analysis using the 'as-implemented' approach. Data from these three units were considered as exposed for all time periods. We finally modelled a categorical variable (not exposed, exposed ≤6 months, 6–12, 12–18, 18–24 or ≥24 months) in the 'as-implemented' analyses to explore the effect of the duration of exposure to the intervention.[36]

Log-likelihood ratio (LR) test was performed to assess the comparisons (using two-sided p value).

RESULTS

Between 2000 and September 2013, the period preceding the study, the overall prevalence rate of Apgar$<7^{5mins}$ across all 15 participating and non-participating large maternity units ≥900 births/year was 1.16%, increasing from 1.06% in 2000 to 1.23% in 2013 (risk ratio: 1.17 95% CI (1.02 to 1.34), p=0.022) (online supplementary appendix 3-figure 1). Four maternity units (U4, U5, U6, U15) had consistently lower rates of Apgar$<7^{5mins}$ compared with the overall 2000–2013 rate observed across all 15 units (0.92% (0.88 to 0.96)); five units (U2, U3, U7, U8, U9) had higher rates of Apgar$<7^{5mins}$(1.51% (1.44 to 1.57)) (online supplementary appendix 3-figure 2). The rates observed in the remaining units were not different from the overall rate. Based on this pre-study information, units were classified for the following analyses as having a status of 'low', 'normal' or 'high' pre-study prevalence rate of Apgar$<7^{5mins}$.

A CONSORT diagram of the SW-RCT is presented in figure 1 and online supplementary appendix 4.

There were 123 943 eligible births in the 15 main Scottish maternity units over the study duration; 123 077 were assigned an Apgar^{5mins} score (99.3%) and 1765 infants were scored an Apgar$<7^{5mins}$.

Three of the 15 maternity units had already commenced PROMPT training prior to the study. Ten of the 12 randomised units implemented the intervention; the remaining 2 declined to participate during the study period. There was a total of 87 204 eligible births (including 99.2% with an assigned Apgar score) and 1291 infants with an Apgar$<7^{5mins}$ in the 12 randomised maternity units.

Figure 2 describes the delivery of the intervention as initially planned and also as actually implemented, including the starting point, length of time to implement local courses and timings for each training session. Units 1–4, 5–8 and 9–12 were randomised to intervention period 1, 2 and 3, respectively. There was considerable variation in the implementation of the local PROMPT courses at unit level in terms of volume of training sessions delivered (online supplementary appendix 5 table 1) and adherence to the randomisation schedule (figure 2).

Data on local training submitted by the participating units demonstrated significant differences in local implementation, including the content of courses. All units included elements of team working in their

Table 1 Prevalence rate of Apgar score$<7^{5\text{mins}}$ during the study period*

Period	Control	Intervention 1	Intervention 2	Intervention 3	Follow-up 1	Follow-up 2
Overall						
Apgar$<7^{5}$ (n)	189	215	208	224	227	228
Births (n)	14 342	15 425	13 905	14 664	13 857	14 321
Prevalence, 95% CI	1.32 (1.14 to 1.52)	1.39 (1.22 to 1.59)	1.50 (1.31 to 1.71)	1.53 (1.34 to 1.74)	1.64 (1.44 to 1.86)	1.59 (1.40 to 1.81)
Intention-to-treat: Intervention as per randomisation						
Unexposed						
Apgar$<7^{5\text{mins}}$ (n)	189	147	52			
Births (n)	14 342	10 619	4540			
Prevalence, 95% CI	1.32 (1.14 to 1.52)	1.38 (1.12 to 1.62)	1.14 (0.87 to 1.50)	—	—	—
Exposed						
Apgar$<7^{5\text{mins}}$ (n)		68	156	224	227	228
Births (n)		4806	9365	14 664	13 857	14 321
Prevalence, 95% CI	—	1.41 (1.11 to 1.79)	1.67 (1.42 to 1.95)	1.53 (1.34 to 1.74)	1.64 (1.44 to 1.86)	1.59 (1.40 to 1.81)
'As-implemented': Intervention as per first training						
Unexposed						
Apgar$<7^{5\text{mins}}$ (n)	189	184	124	87	51	21
Births (n)	14 342	14 293	9040	6641	2839	1306
Prevalence, 95% CI	1.32 (1.14 to 1.52)	1.29 (1.11 to 1.49)	1.37 (1.15 to 1.63)	1.31 (1.06 to 1.61)	1.80 (1.37 to 2.36)	1.61 (1.05 to 2.45)
Exposed						
Apgar$<7^{5\text{mins}}$ (n)		31	84	137	176	207
Births (n)		1 132	4 865	8 023	11 018	13 015
Prevalence, 95% CI	—	2.74 (1.93 to 3.87)	1.73 (1.40 to 2.13)	1.71 (1.45 to 2.01)	1.60 (1.38 to 1.85)	1.59 (1.38 to 1.82)

*Based on the 86 514 births with an assigned Apgar score observed in the 12 maternity units which had attended the THISTLE Study Train the Trainer (T3) programme and had not implemented the intervention prior to the study (two units declined to participate during the study period).
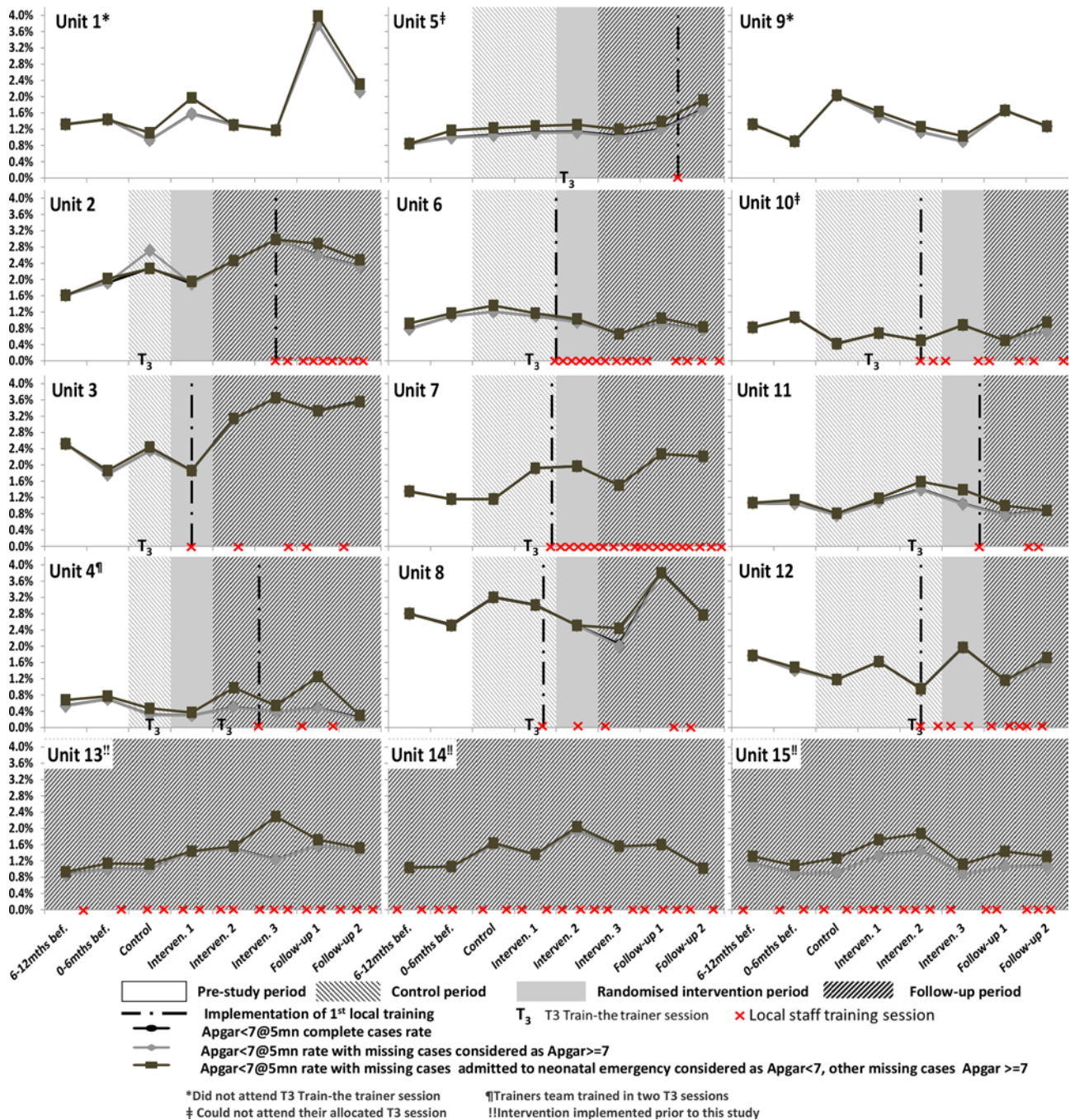
Figure 2 Apgar<7^{5mins} prevalence rates by maternity units and study time periods, randomisation schedule and actual intervention implementation.

courses, but only three included fetal monitoring sessions, which is considered a core component of PROMPT training and related to Apgar scores. Shoulder dystocia training, another core element of PROMPT, was also not included in some programmes (online supplementary appendix 5 table 2). Participating units provided information on numbers of staff trained during their local courses but were unable to provide baseline unit-level staffing numbers, and therefore, we were not able to determine the cumulative proportion of staff trained by unit during the study period.[26] The overall Apgar<7^{5mins} prevalence rate observed during the study period in the 12 randomised maternity units

was 1.49%: increasing from 1.32% pre-intervention to 1.59% post-intervention (unadjusted test for trend of odds, $\chi^2=6.4$, p=0.01) (table 1). This underlying time trend was adjusted for in the different analyses.

Effect of the intervention

The ITT analysis showed a moderate but non-significant reduction in the proportion of babies with an Apgar score<7^{5mins} after the implementation of the intervention (OR=0.79 95%CI(0.63 to 1.01), table 2). Consistent results were observed in the sensitivity analyses. No intervention effect was found when exposure to the intervention was defined according

Table 2 PROMPT training effectiveness (OR and associated 95% CI); main analysis and sensitivity analyses*

| | Across participating units† | | | | | | Across health system‡ | | |
| | Intention-to-treat | | | 'As-implemented' | | | 'As-implemented' | | |
	OR	95% CI	P value	OR	95% CI	P value	OR	95% CI	P value
Intervention effect									
Unexposed	Ref			Ref			Ref		
Exposed									
Complete-cases	0.79	(0.63 to 1.01)	0.06	1.01	(0.84 to 1.22)	0.92	1.09	(0.93 to 1.28)	0.27
Imputation 1	0.80	(0.63 to 1.01)	0.06	1.01	(0.84 to 1.22)	0.90	1.09	(0.93 to 1.28)	0.26
Imputation 2	0.81	(0.64 to 1.01)	0.06	1.00	(0.83 to 1.19)	0.99	1.11	(0.95 to 1.28)	0.18
Adjustment factors									
Time period									
Control period	Ref			Ref			Ref		
Intervention period 1									
Complete-cases	1.16	(0.93 to 1.43)	0.19	1.06	(0.87 to 1.29)	0.58	1.09	(0.92 to 1.29)	0.32
Imputation 1	1.15	(0.93 to 1.43)	0.19	1.05	(0.86 to 1.29)	0.60	1.09	(0.92 to 1.28)	0.33
Imputation 2	1.14	(0.93 to 1.41)	0.21	1.05	(0.87 to 1.28)	0.60	1.07	(0.91 to 1.26)	0.42
Intervention period 2									
Complete-cases	1.34	(1.03 to 1.74)	0.03	1.13	(0.91 to 1.40)	0.26	1.18	(0.99 to 1.40)	0.06
Imputation 1	1.34	(1.03 to 1.73)	0.03	1.13	(0.91 to 1.39)	0.28	1.18	(0.99 to 1.40)	0.07
Imputation 2	1.35	(1.05 to 1.73)	0.02	1.15	(0.94 to 1.41)	0.17	1.18	(1.00 to 1.39)	0.05
Intervention period 3									
Complete-cases	1.46	(1.08 to 1.99)	0.01	1.16	(0.92 to 1.44)	0.20	1.08	(0.90 to 1.30)	0.40
Imputation 1	1.46	(1.07 to 1.98)	0.02	1.15	(0.92 to 1.44)	0.22	1.08	(0.90 to 1.30)	0.41
Imputation 2	1.46	(1.09 to 1.97)	0.01	1.18	(0.95 to 1.46)	0.13	1.07	(0.90 to 1.27)	0.44
Follow-up period 1									
Complete-cases	1.57	(1.15 to 2.13)	<0.01	1.23	(0.96 to 1.58)	0.09	1.17	(0.96 to 1.41)	0.11
Imputation 1	1.56	(1.15 to 2.12)	<0.01	1.23	(0.96 to 1.57)	0.10	1.16	(0.96 to 1.41)	0.12
Imputation 2	1.63	(1.21 to 2.19)	<0.01	1.32	(1.04 to 1.67)	0.02	1.19	(1.00 to 1.43)	0.05
Follow-up period 2									
Complete-cases	1.53	(1.12 to 2.07)	<0.01	1.20	(0.93 to 1.56)	0.16	1.09	(0.89 to 1.33)	0.40
Imputation 1	1.52	(1.12 to 2.06)	<0.01	1.23	(0.96 to 1.57)	0.10	1.09	(0.89 to 1.32)	0.41
Imputation 2	1.50	(1.12 to 2.02)	<0.01	1.21	(0.95 to 1.56)	0.13	1.06	(0.88 to 1.28)	0.55
Pre-study maternity unit size									
Small									
Complete-cases	1.27	(0.72 to 2.23)	0.41	1.24	(0.72 to 2.16)	0.44	1.18	(0.75 to 1.86)	0.46

Continued

Table 2 Continued

| | Across participating units† | | | | | | Across health system‡ | | |
| | Intention-to-treat | | | 'As-implemented' | | | 'As-implemented' | | |
	OR	95% CI	P value	OR	95% CI	P value	OR	95% CI	P value
Imputation 1	1.26	(0.71 to 2.23)	0.42	1.24	(0.71 to 2.16)	0.45	1.18	(0.75 to 1.86)	0.47
Imputation 2	1.27	(0.79 to 2.04)	0.33	1.25	(0.79, 1.97)	0.34	1.21	(0.83 to 1.77)	0.32
Medium	Ref			Ref			Ref		
Large									
Complete-cases	1.28	(0.78 to 2.10)	0.33	1.27	(0.78 to 2.06)	0.33	1.29	(0.90 to 1.84)	0.16
Imputation 1	1.27	(0.77 to 2.10)	0.34	1.27	(0.78 to 2.06)	0.34	1.29	(0.90 to 1.84)	0.17
Imputation 2	1.26	(0.83 to 1.89)	0.27	1.25	(0.84 to 1.85)	0.27	1.27	(0.95 to 1.70)	0.10
Pre-study maternity Apgar score									
Below national rate									
Complete-cases	0.75	(0.41 to 1.37)	0.35	0.72	(0.40 to 1.29)	0.27	0.69	(0.46 to 1.04)	0.07
Imputation 1	0.74	(0.40 to 1.35)	0.33	0.71	(0.40 to 1.28)	0.26	0.68	(0.45 to 1.03)	0.07
Imputation 2	0.88	(0.53 to 1.44)	0.61	0.85	(0.53 to 1.36)	0.49	0.83	(0.59 to 1.16)	0.26
Similar to national rate	Ref			Ref			Ref		
Above national rate									
Complete-cases	2.17	(1.31 to 3.58)	<0.01	2.09	(1.28 to 3.41)	<0.01	1.91	(1.32 to 2.77)	<0.01
Imputation 1	2.17	(1.31 to 3.60)	<0.01	2.10	(1.28 to 3.43)	<0.01	1.91	(1.32 to 2.78)	<0.01
Imputation 2	2.03	(1.33 to 3.09)	<0.01	1.96	(1.32 to 2.93)	<0.01	1.82	(1.34 to 2.48)	<0.01

Intention-to-treat: Intervention periods defined according to randomisation schedule. 'As-implemented': Intervention periods defined according to actual start of first local intervention training session. Complete-cases: Analysis includes only births with assigned Apgar score considered (n=86 514 for the 12 units or n=123 077 for the 15 units). Imputation 1: All births included (n=87 204 for the 12 units or n=123 943 for the 15 units). Births with missing Apgar scores assigned to Apgar<7^{5mins} group. Imputation 2: All births included. Births with missing Apgar scores considered as Apgar<7^{5mins}. Babies with no Apgar score but recorded as admitted to a neonatal intensive care unit and/or received a form of resuscitation other than facial oxygen assigned to Apgar<7^{5mins} group; remaining babies with missing scores assigned to the Apgar≥7^{5mins} group.

*Mixed effects logistic regression (with random effects for clustering of births within unit) assessing intervention and adjusted for time trend, pre-study maternity unit size and status of pre-study prevalence rate of Apgar score<7^{5mins}.

†12 Scottish maternity units not exposed to the intervention prior to this study including two randomised units which declined to participate during the study period.

‡All 15 Scottish maternity units with ≥900 births/year.

to its actual implementation rather than the randomisation schedule (OR=1.01 (0.84 to 1.22), table 2). The duration of exposure to the intervention did not affect the outcome ($OR_{exposed \leq 6months\ vs\ not\ exposed}$=1.08 (0.88 to 1.32), $OR_{exposed 6-12months\ vs\ not\ exposed}$=0.93 (0.73, 1.19), $OR_{exposed 12-18months\ vs\ not\ exposed}$=1.02 (0.75 to 1.37), $OR_{exposed 18-24months\ vs\ not\ exposed}$=1.18 (0.82 to 1.71), $OR_{exposed \geq 24months\ vs\ not\ exposed}$=1.12 (0.66 to 1.87); overall LR test χ^2=3.7, p=0.86). There was also no evidence of an intervention effect when the analyses were extended across all 15 large maternity units (table 2).

Effect of time-trend, pre-study Apgar<7^{5mins} rate and maternity unit size

A time-trend was observed in the ITT analyses, with higher rates of Apgar<7^{5mins} towards the end of the study, compared with the beginning (table 2). This trend was no longer evident in the 'as-implemented' analyses (overall LR test χ^2=3.1, p=0.69), with numerous births in the same time period being reclassified as unexposed rather exposed (and vice-versa), including births which occurred in units 1 and 9 being fully reclassified as unexposed during all time periods.

The rate of term babies with an Apgar score<7^{5mins} remained higher in units with a pre-study Apgar rate higher than the national rate compared with units with a pre-existing rate comparable to the national rate (OR=2.09 (1.28 to 3.41), table 2). Units with a low pre-study rate had comparable Apgar score<7^{5mins} rates during the study period compared with units with an average pre-study rate (OR=0.72 (0.40 to 1.29)). There was no evidence of interaction between the intervention and the pre-study Apgar rate (overall LR test: In ITT χ^2=0.9, p=0.64, 'as-implemented'$_{12units}$ χ^2=4.2, p=0.12, 'as-implemented'$_{15units}$ χ^2=3.4, p=0.18).

There was no evidence of an association between maternity unit size and the rate of babies born with an Apgar score <7 5 mins (table 2).

DISCUSSION

There was no reduction in the proportion of term births with an Apgar score <7^{5mins} following PROMPT training (OR=1.01 (0.84 to 1.22)) in Scotland. The rate of Apgar score<7^{5mins} continued to increase (1.32% preintervention to 1.59% postintervention) and was higher in units with a high baseline rate of Apgar<7^{5mins} (OR=2.09 (1.28 to 3.41). The implementation of the intervention was very heterogeneous between units, with significant variation in the content of local courses, number of courses implemented and also the number of staff trained.

This study is the first to evaluate implementation of a multiprofessional training intervention for obstetric emergencies at national level. The clinical data were collected prospectively and nationally, independent of this study, limiting potential reporting bias and reducing the burden of data collection.

We employed a SW-RCT[37] design to evaluate an intervention with prior evidence of benefit. This design has been widely used in the context of limited delivery resources to evaluate educational or public health training programmes that are anticipated to be beneficial and, therefore, to prevent ethical objections.[38] In our study, it was deemed unethical to use a parallel (cluster) RCT design as there was insufficient equipoise for the intervention, and no alternative training could be proposed to the control group.[15 16 19] However, the SW-RCT design proved to be less robust than originally envisaged when used to study a training intervention.[39] While most units were keen to participate, adherence to the randomisation schedule was variable. Some units were able to commence training immediately after their T3 session and before their allocated step, whereas others started several months after the allocated starting point. We performed an ITT approach using the dates at which maternity units were randomised to commence training, irrespective of when they actually started local training, to define the 'control/non-exposed' and 'intervention/exposed' groups. We also performed a sensitivity analysis in the form of an 'as-implemented' analysis. We do not think that the ITT approach accurately captured the local implementation of the intervention. Therefore, we consider the 'as-implemented' approach to be the most accurate analysis method for the effect of intervention. While both ITT and 'as-implemented' analyses led to the same conclusion, the extent of the estimated risk was different between these two approaches, most likely because the definition of exposure to the intervention was fundamentally different between the two approaches and because of the pre-study Apgar<7^{5mins} status of the unit. A number of births have been alternatively considered as 'unexposed' in one approach but 'exposed' in the other approach (or vice versa) for the same time period. Moreover, units with low and average baseline rates of Apgar<7^{5mins} status—a high predictor of current Apgar<7^{5mins} rate—had the largest time departure from the planned implementation schedule, and therefore the largest intervention exposure misclassification in the ITT analysis compared with the 'as-implemented' approach. Other unknown factors might also explain the change in the effect size between the two approaches. This raises two important methodological considerations when modelling findings from SW-RCT. We agree with other methodologists that it is important to consider the impact of time trend on the estimated effect[33 35–37] and the necessity for time adjustment in the primary analysis. More importantly, our findings suggest that ITT analysis should not be considered in isolation when writing the statistical analysis plan of a SW-RCT. Sensitivity analysis such as our 'as-implemented' approach should also be employed, particularly for complex interventions for example, training.

We had initially envisaged the SW-RCT design as a pragmatic design to assess PROMPT at larger scale, within a reasonable timeframe and with restricted resources. However, the intervention implementation was challenging and the restrictions on logistics, time and funding meant that we were unable to provide ongoing implementation support to the participating units, other than by phone or email. Furthermore, maternity units were often unable to commence the intervention as planned, either starting earlier or later than the randomisation schedule. Finally, the content and frequency of local courses was heterogeneous. We recommend that future SW-RCT studies should consider a feasibility stage and a parallel process evaluation, as recommended for complex interventions,[40 41] to explore potential barriers and facilitators to randomisation and intervention implementation. The theoretical savings in time, resources and staffing that a SW-RCT appeared to offer, might be limited, especially when dealing with an intervention requiring behavioural changes or authentic implementation in busy health systems. Therefore, SW-RCT designs may not have that many advantages for this type of study.

We were also unable to provide funding to support maternity units to collate their baseline staffing composition and volume: the cumulative proportion of staff trained over the duration of the study could not be derived as initially planned.[26] Moreover, to reduce the burden of increased workload on already busy clinical staff, we requested only limited process data. Therefore, our investigation of the reasons underlying the delay, frequency, duration and content of local intervention implementation was limited, as was our investigation into the extent of the 'intervention penetration'; we could not test the intervention effect according to the proportion of staff trained at different time points. We recommend that funding should be sought to assess local intervention penetration in future studies of training; in particular, any SW-RCT evaluating training effectiveness should allocate resources to enable the measurement of the proportion of staff trained.

PROMPT training has been associated with improvements in outcomes in other contexts and associated with sustained local multiprofessional training where more than 90% of staff trained annually.[13 22 23] It is unlikely that any of the participating Scottish units were able to train 90%–95% of their staff. However, a higher level of implementation support was also provided by the central PROMPT team in some of these settings compared with the THISTLE units. Other national and international maternity units that have successfully implemented PROMPT took 1 year to train all of their staff and another year to demonstrate a difference in outcomes.[13 21] The PROMPT intervention involves behavioural changes and more research is required to understand the support needed by maternity units to implement local PROMPT authentically.

The substantial variation in the implementation of the PROMPT intervention by the Scottish units in this study may have affected the intervention authenticity and, therefore, our findings most likely reflect the effectiveness of the intervention as implemented in Scotland, rather than the efficacy that can be achieved if PROMPT is implemented exactly as recommended and has been observed among other early adopter sites. There may be several reasons for the variation in intervention authenticity; there were simultaneous and sometimes contradictory national training initiatives for example, a national training programme for Fetal Monitoring in Labour,[42] was introduced during the study period. In addition, although recommendations were made during the T3 programme to guide participating teams on the core components to include in their local multiprofessional PROMPT courses, units were allowed flexibility to facilitate implementation in their local context, to enhance ownership of the training. Furthermore, while we recommended start dates for local implementation for all participating units and that they should aim to train all of their maternity staff within 12 months, we could not mandate this or provide any further implementation support. Therefore, we were not able to ensure total authenticity of local implementation and in particular, the omission of the PROMPT fetal monitoring elements by some units is likely to have had a major impact on the primary outcome.

The use of Apgar scores to determine the quality of intrapartum care has been criticised.[43] However, an Apgar score $<7^{5mins}$ is associated with poor long-term outcomes,[31 32] can be improved by best care[13] and is independent of maternal demographics.[31] It would have been useful to have investigated other outcomes that have been improved with PROMPT for example, brachial plexus injuries[20] and future studies should consider more than a single measure of improvement.

Reducing avoidable harm in maternity care is an urgent priority and improved training for intrapartum care is at least part of the solution, but it must be both effective and sustainable. Unlike some other less-effective interventions[15] that appeared to show early promise, this research has not shown any evidence of harm related to the PROMPT intervention. Not all training for obstetric emergencies is effective,[44] however well intentioned, but the implementation of PROMPT at unit level has been associated with statistically significant, clinically important improvements in perinatal and maternal outcomes in different resource settings. This includes reductions in brachial plexus injuries,[14 20 23] encephalopathy or low 5 min Apgar score,[13 23] composite neonatal outcomes[45] as well as reductions in maternal deaths in Zimbabwe[21] and in the Philippines.[46]

Health and social care systems continue to be challenged by the problem that initially promising interventions often prove difficult to replicate and scale. The dynamic interplay between intervention and context means that it is often difficult to separate intervention from context.[47 48] Transplanting a programme in its entirety from one setting to another is rarely straightforward: the features of context (infrastructure, organisational systems, values, skills) may need to be reproduced too.[49–51] A recent ethnographic study conducted in the maternity unit where PROMPT was created have identified how this intervention and the unit contextual features shaped each other and how this process was influenced by the unit's structural conditions, such as staffing levels and physical environment.[52] The PROMPT team are now exploring a social science approach to scale up and spread, including social franchising.[53] A parallel process study, THISTLE Plus,[54] separately funded, has also been implemented in line with MRC recommendations[55] and will provide further insight into local barriers and facilitators that impacted the intervention implementation in Scotland.

It is essential that we investigate how best to scale effective interventions, particularly for maternity care where the costs of litigation are unsustainable,[56] the human costs of preventable harm at birth notwithstanding. This study demonstrates that implementation can be difficult at local level and suggests that implementation support is required to ensure that the intervention is authentically adopted and effective. Training is not free,[57] however, effective training is very cost effective. Improving perinatal outcomes can substantially reduce litigation costs.[23 56]

The rate of Apgar$<7^{5mins}$ in Scotland went up throughout the study. This increase is likely to be multifactorial. The rate of prematurity—a key risk factor for low Apgar score—, as well as the rates of intrauterine death and of elective CS—two other exclusion factors for our studied sample—remained stable between the pre-study period and the study period (7% vs 8 %, 0.3% vs 0.3% and 14% vs 15%, respectively). This is suggesting that the studied sample has retained some comparability over time with regard to our exclusion criteria and/or factors related to low Apgar score status. Other factors such as maternal occupation or education status could have changed over time; the strength of their relationship with the Apgar score status is complex, depending, for example, on the mode of delivery, that is, absent or marginal in vaginal delivery but higher in CS delivery.[31] Changes in quality of care could explain the observed increase in low Apgar score and further work is now required to understand and reverse this trend. Research and clinical investigations in units which have maintained low rates of Apgar$<7^{5mins}$ would provide insights into this and could help to understand the 'active ingredients' of effective training, as well as strategies for the successful implementation of effective interventions at scale. Finally, it is important for clinical researchers to mobilise evidence up to policy makers. For example, system actors like state insurers are not patient safety organisations, but they do have a role in reducing, particularly aligning and focussing the system on improvement[56 58] and initiatives should be based on the evidence base for clinically effective improvement interventions.

Conclusion

PROMPT training, as implemented, had no effect on the rate of Apgar $<7^{5mins}$ in Scotland during the study period. Local implementation was found to be more difficult than anticipated and further research is required to understand why the positive effects observed in other single-unit studies have not been replicated in Scottish maternity units during the THISTLE Study.

Author affiliations
[1]Translational Health Sciences, University of Bristol, Bristol, UK
[2]Department of Women's Health and Children's Health, Southmead Hospital, North Bristol NHS Trust, Westbury on Trym, UK
[3]Institute for Women's Health, UCL, London, United Kingdom
[4]Centre for Healthcare Randomised Trials, Health Services Research Unit, University of Aberdeen, Aberdeen, UK
[5]Maternity Unit, Ninewells Hospital and Medical School, NHS Tayside, Dundee, UK
[6]Ian Donald Fetal Medicine Centre, Queen Elizabeth University Hospital, NHS Greater Glasgow and Clyde, Glasgow, UK
[7]Edinburgh Clinical Trials Unit, Centre for Population Health Sciences, Usher Institute, University of Edinburgh, Edinburgh, UK
[8]Institute of Applied Health Sciences, University of Aberdeen, Aberdeen, UK

ORCID iD
Erik Lenguerrand http://orcid.org/0000-0002-0371-731X

REFERENCES

1 Royal college of obstetricians and gynaecologists. Each baby counts 2018 progress report. London: RCOG, 2018

2 Magro M. *Five years of cerebral palsy claims. A thematic review of NHS Resolution data*. London: NHS Resolution, 2017.

3 Draper ES, Gallimore ID, Kurinczuk JJ, *et al*. MBRRACE-UK Perinatal mortality surveillance report, UK Perinatal deaths for births from january to december 2016. Leicester: The infant mortality and morbidity studies, department of health sciences, University of Leicester, 2018

4 O'Neill O. *Safe births: Everybody's business. An independent inquiry into the safety of maternity services in england*. London: King's Fund, 2008.

5 Lewis G. *The Confidential Enquiry into Maternal and child health (CEMACH). The Seventh Report on Confidential Enquiries into Maternal Deaths in the United Kingdom*. London: CEMACH, 2007.

6 Maternal and Child Health Research Consortium. Maternal and Child Health Research Consortium. Confidential Enquiry into Stillbirths and Deaths in Infancy. *London* 2001.

7 NHS Litigation Authority. *Ten years of maternity claims. An analysis of nhs litigation authority data*. London: NHS LA, 2012.

8 NHS Litigation Authority. *Cnst maternity clinical risk management standards*. London: NHS LA, 2006.

9 In:Royal College of OBSTETRICIANS and GYNAECOLOGISTS, Royal College of MIDWIVES, Royal College of ANAESTHETISTS,. Standards for Maternity Care. Report of a Working Party. London RCOG Press; 2008. et al.

10 Victorian Managed Insurance Authority (VMIA). Annual report. Melbourne VMIA; 2018.

11 NHS Litigation Authority. *Clinical negligence scheme for trusts – maternity clinical risk management standards*. London: NHS LA, 2010.

12 NHS Resolution. *CNST Scheme 2017*. London: NHS Resolution, 2017.

13 Draycott T, Sibanda T, Owen L, *et al*. Does training in obstetric emergencies improve neonatal outcome? *BJOG: Int J O&G* 2006;113:177–82.

14 Draycott TJ, Crofts JF, Ash JP, *et al*. Improving neonatal outcome through practical shoulder dystocia training. *Obstet Gynecol* 2008;112:14–20.

15 MacKenzie IZ, Shah M, Lean K, *et al*. Management of shoulder dystocia: trends in incidence and maternal and neonatal morbidity. *Obstet Gynecol* 2007;110:1059–68.

16 Nielsen PE, Goldman MB, Mann S, *et al*. Effects of teamwork training on adverse outcomes and process of care in labor and delivery: a randomized controlled trial. *Obstet Gynecol* 2007;109:48–55.

17 Siassakos D, Hasafa Z, Sibanda T, *et al*. Retrospective cohort study of diagnosis-delivery interval with umbilical cord prolapse: the effect of team training. *BJOG* 2009;116:1089–96.

18 Sørensen JL, Løkkegaard E, Johansen M, *et al*. The implementation and evaluation of a mandatory multi-professional obstetric skills training program. *Acta Obstet Gynecol Scand* 2009;88:1107–17.

19 Fransen AF, van de Ven J, Schuit E, *et al*. Simulation-based team training for multi-professional obstetric care teams to improve patient outcome: a multicentre, cluster randomised controlled trial. *BJOG* 2017;124.

20 Crofts JF, Lenguerrand E, Bentham GL, *et al*. Prevention of brachial plexus injury-12 years of shoulder dystocia training: an interrupted time-series study. *BJOG* 2016;123:111–8.

21 Crofts JF, Mukuli T, Murove BT, *et al*. Onsite training of doctors, midwives and nurses in obstetric emergencies, Zimbabwe. *Bull World Health Organ* 2015;93:347–51.

22 Shoushtarian M, Barnett M, McMahon F, *et al*. Impact of introducing practical obstetric multi-professional training (PROMPT) into maternity units in Victoria, Australia. *BJOG* 2014;121:1710–8.

23 Weiner CP, Collins L, Bentley S, *et al*. Multi-professional training for obstetric emergencies in a US hospital over a 7-year interval: an observational study. *J Perinatol* 2016;36:19–24.

24 Scotland ISD. Births in Scottish Hospitals: ISD Scotland, 2016. Available: http://www.isdscotland.org/Health-Topics/Maternity-and-Births/Births/ [Accessed 02 Dec 2016].

25 Apgar V. The Newborn (APGAR) Scoring System: Reflections and Advice. *Pediatr Clin North Am* 1966;13:645–50.

26 Lenguerrand E, Winter C, Innes K, *et al*. THISTLE: trial of hands-on Interprofessional simulation training for local emergencies: a research protocol for a stepped-wedge clustered randomised controlled trial. *BMC Pregnancy Childbirth* 2017;17:294.

27 PROMPT Maternity Foundation. *Trainer's Manual*. London: Royal College of Obstetricians and Gynaecologists, 2012.

28 PROMPT Maternity Foundation. *Course Manual*. Cambridge: Cambridge University Press, 2012.

29 Carter BR, Hood K. Balance algorithm for cluster randomized trials. *BMC Med Res Methodol* 2008;8:65.

30 Raab GM, Butcher I. Balance in cluster randomized trials. *Stat Med* 2001;20:351–65.

31 Odd DE, Doyle P, Gunnell D, *et al*. Risk of low Apgar score and socioeconomic position: a study of Swedish male births. *Acta Paediatr* 2008;97:1275–80.

32 Odd DE, Lewis G, Whitelaw A, *et al*. Resuscitation at birth and cognition at 8 years of age: a cohort study. *Lancet* 2009;373:1615–22.

33 Hussey MA, Hughes JP. Design and analysis of stepped wedge cluster randomized trials. *Contemp Clin Trials* 2007;28:182–91.

34 Barker D, D'Este C, Campbell MJ, *et al*. Minimum number of clusters and comparison of analysis methods for cross sectional stepped wedge cluster randomised trials with binary outcomes: A simulation study. *Trials* 2017;18:119.

35 Hemming K, Taljaard M, Forbes A. Analysis of cluster randomised stepped wedge trials with repeated cross-sectional samples. *Trials* 2017;18:101.

36 Nickless A, Voysey M, Geddes J, *et al*. Mixed effects approach to the analysis of the stepped wedge cluster randomised trial-Investigating the confounding effect of time through simulation. *PLoS One* 2018;13:e0208876.

37 Hemming K, Haines TP, Chilton PJ, *et al*. The stepped wedge cluster randomised trial: rationale, design, analysis, and reporting. *BMJ* 2015;350:h391.

38 Brown CA, Lilford RJ. The stepped wedge trial design: a systematic review. *BMC Med Res Methodol* 2006;6:54.

39 Handley MA, Schillinger D, Shiboski S. Quasi-experimental designs in practice-based research settings: design and implementation considerations. *J Am Board Fam Med* 2011;24:589–96.

40 Craig P, Dieppe P, Macintyre S, *et al*. Developing and evaluating complex interventions: new guidance, 2006. Available: https://www.mrc.ac.uk/documents/pdf/complex-interventions-guidance/

41 Kristunas CA, Hemming K, Eborall HC, *et al*. The use of feasibility studies for stepped-wedge cluster randomised trials: protocol for a review of impact and scope. *BMJ Open* 2017;7:e017290.

42 About the Fetal Monitoring programme. Available: https://www.e-lfh.org.uk/programmes/electronic-fetal-monitoring [Accessed 11/05/2018].

43 Pyykönen A, Gissler M, Jakobsson M, *et al*. Determining obstetric patient safety indicators: the differences in neonatal outcome measures between different-sized delivery units. *BJOG* 2014;121:430–7.

44 Draycott T. Not all training for obstetric emergencies is equal, or effective. *BJOG* 2017;124.

45 Shoushtarian M, Barnett M, McMahon F, *et al*. *Impact of introducing Practical Obstetric Multi-Professional Training (PROMPT) into maternity units in Victoria, Australia*. BJOG: An International Journal of Obstetrics & Gynaecology, 2014: 121. 1710–8.

46 Republic of the Philippines Phillippine Information Agency. PROMPT implementation reduces maternal, neonatal mortality Quezon City: Republic of the Philippines, 2018. Available: http://pia.gov.ph/news/articles/1005787 [Accessed 08 May 2018].

47 Hawe P, Shiell A, Riley T. Theorising interventions as events in systems. *Am J Community Psychol* 2009;43:267–76.

48 Wells M, Williams B, Treweek S, *et al*. Intervention description is not enough: evidence from an in-depth multiple case study on the untold role and impact of context in randomised controlled trials of seven complex interventions. *Trials* 2012;13.

49 Bauman LJ, Stein REK, Ireys HT. Reinventing fidelity: The transfer of social technology among settings. *Am J Community Psychol* 1991;19:619–39.

50 Dixon-Woods M, McNicol S, Martin G. *Overcoming challenges: Lessons from the Health Foundation's improvement programme evaluations and relevant literature to improving quality*. London: Health Foundation, 2012.

51 Dixon-Woods M. The problem of context in quality improvement: Perspectives on context. London: Health Foundation, 2014: 87–101.

52 Liberati EG, Tarrant C, Willars J, *et al*. How to be a very safe maternity unit: An ethnographic study. *Soc Sci Med* 2019;223:64–72.

53 Greenhalgh T, Papoutsi C. Spreading and scaling up innovation and improvement. *BMJ* 2019;365:l2068.

54 Collins K, Moore A, Winter C, *et al*. *Understanding the implementation of obstetric emergency training: The THISTLE-Plus study*. RCOG World Congress. Cape Town, South Africa: BJOG, 2017: 43–4.

55 Moore GF, Audrey S, Barker M, *et al*. Process evaluation of complex interventions: Medical Research Council guidance. *BMJ* 2015;350:h1258.

56 Draycott T, Sagar R, Hogg S. The role of insurers in maternity safety. *Best Pract Res Clin Obstet Gynaecol* 2015;29.

57 Yau CWH, Pizzo E, Morris S, *et al*. The cost of local, multi-professional obstetric emergencies training. *Acta Obstet Gynecol Scand* 2016;95:1111–9.

58 Weinstein L. A multifacited approach to improve patient safety, prevent medical errors and resolve the professional liability crisis. *Am J Obstet Gynecol* 2006;194:1160–5.