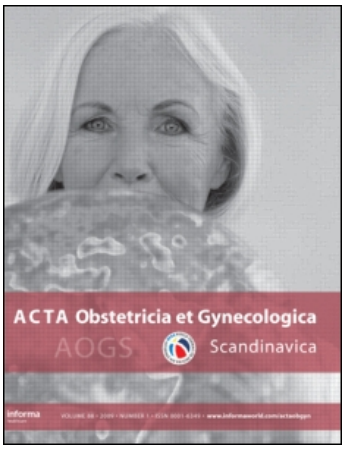

Acta Obstetricia et Gynecologica Scandinavica
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t716100748

The implementation and evaluation of a mandatory multi-professional obstetric skills training program

Jette Led Sørensen [a]; Ellen Løkkegaard [b]; Marianne Johansen [a]; Charlotte Ringsted [c]; Svend Kreiner [d]; Sean McAleer [e]
[a] Obstetric Department, The Juliane Marie Center for Children, Women and Reproduction, Copenhagen University Hospital, Rigshospitalet, Copenhagen, Denmark [b] Obstetric and Gynecological Department, Hillerød Hospital, Copenhagen, Denmark [c] Centre for Clinical Education, Copenhagen University and Capital Region, Rigshospitalet, Copenhagen, Denmark [d] Department of Biostatistics, Faculty of Health Sciences, University of Copenhagen, Copenhagen, Denmark [e] Centre for Medical Education, University of Dundee, Dundee, Scotland, UK

First Published on: 17 August 2009

To cite this Article Sørensen, Jette Led, Løkkegaard, Ellen, Johansen, Marianne, Ringsted, Charlotte, Kreiner, Svend and McAleer, Sean(2009)'The implementation and evaluation of a mandatory multi-professional obstetric skills training program',Acta Obstetricia et Gynecologica Scandinavica,99999:1,

To link to this Article: DOI: 10.1080/00016340903176834

URL: http://dx.doi.org/10.1080/00016340903176834

Acta Obstetricia et Gynecologica
2009, 1–11, iFirst article

informa
healthcare

ORIGINAL ARTICLE

The implementation and evaluation of a mandatory multi-professional obstetric skills training program

JETTE LED SØRENSEN[1], ELLEN LØKKEGAARD[2], MARIANNE JOHANSEN[1], CHARLOTTE RINGSTED[3], SVEND KREINER[4] & SEAN MCALEER[5]

[1]*Obstetric Department, The Juliane Marie Center for Children, Women and Reproduction, Copenhagen University Hospital, Rigshospitalet, Copenhagen, Denmark,* [2]*Obstetric and Gynecological Department, Hillerød Hospital, Copenhagen, Denmark,* [3]*Centre for Clinical Education, Copenhagen University and Capital Region, Rigshospitalet, Copenhagen, Denmark,* [4]*Department of Biostatistics, Faculty of Health Sciences, University of Copenhagen, Copenhagen, Denmark, and* [5]*Centre for Medical Education, University of Dundee, Dundee, Scotland, UK*

Abstract
Objective. To implement and evaluate a simulation-based training program. *Design.* Descriptive. Study period: June 2003–June 2006. *Setting.* Obstetric Department, Copenhagen University Hospital, Rigshospitalet, Copenhagen, Denmark. *Population.* Two training sessions were provided for all health professionals including doctors, midwives, auxiliary nurses, and 147 out 156 participants (94%) took part in the first training session and 192 out possible 201 (96%) took part in the second session. *Methods.* An intervention study of the impact of simulation-based training in management of postpartum bleeding, shoulder dystocia, basic neonatal resuscitation, and severe preeclampsia. *Main outcome measures.* Before, just after and 9–15 months following the training, data were collected on the confidence and stress levels relating to the carrying out of certain procedures. In addition, a written objective test on basic neonatal resuscitation was administered. Data on any changes in work-routines experienced by the participants were obtained by open-ended questions. Registry data from the Danish Medical Birth Registry and from the hospital administration were included in the analysis. *Results.* Ninety-two percent of all respondents had a positive attitude toward the training program. They considered management of shoulder dystocia, preeclampsia, and neonatal resuscitation less stressful and less unpleasant to perform after training. Confidence scores for all the trained skills improved significantly. A significant association was found between confidence in neonatal resuscitation and numbers of correct answers in the objective test. More than 90% found the training to have had a positive influence on their work. The need for organizational changes in the department became evident and necessary changes were implemented. Sick leave amongst midwives diminished significantly during the study period. *Conclusions.* A comprehensive evaluation of a mandatory simulation-based program, implemented in a obstetric department, demonstrated a positive impact at individual and organizational levels.

Key words: *Patient simulation, clinical competence, performance assessment, skills training, education, obstetrics, labor*

Introduction

Labor wards have a dual function in creating a relaxed atmosphere for normal childbirth and dealing with life-threatening emergencies. This makes them a challenging work place. It is a prerequisite that labor ward staff have knowledge, skills, and competence in clinical problem solving, communication, and cooperation with both the delivering women, their relatives, and all the different health professionals involved (1–4).

The working conditions in the labor ward may have an influence on the staff's well-being and competencies. As a profession, midwives are described to be likely to suffer from burn out (5), and they have limited access to training and knowledge updating (6). The frequency of sick leave among midwives was

Correspondence: Jette Led Sørensen, Obstetric Department, The Juliane Marie Center for Children, Women and Reproduction, Copenhagen University Hospital, Rigshospitalet, Copenhagen, Denmark. E-mail: jette.led.soerensen@rh.regionh.dk

(Received 13 April 2009; accepted 9 July 2009)

ISSN 0001-6349 print/ISSN 1600-0412 online
DOI: 10.1080/00016340903176834

for several years antedating the present study among the highest compared to other staff group in the Copenhagen University Hospital, Rigshospitalet.

Previous research on obstetric training has been sparse. A review (7) concluded that 'few programs have been described and even fewer have been evaluated. Training methods need to be developed, described, and evaluated; further well-conducted research for this important intervention is urgently required'. It has been stressed that other specialties are ahead of obstetrics in the use of simulation (2,8). It is argued (2) that 'those of us involved in training must think creatively' and that there is a need for new training methods with more emphasis on team training. In a British study, a mandatory annual course for all labor ward staff was shown to be associated with clinical importance and sustained improvement in perinatal outcome (9). That observational study from 2006 was followed by a randomized study in obstetric simulation (10,11), where no additional benefit from training in a simulation center compared with training in local hospitals was found. Thus, it is important to address the questions about training in simulation centers versus training in local hospitals in obstetrics (12).

The present study describes the implementation and evaluation of a mandatory multi-professional simulation-based training program in a local hospital.

Material and methods

The evaluation in this study design was based on Kirkpatricks four levels of evaluation (13,14): (1) *Reaction-level* which measures the participants' satisfaction with the training program, (2) *Learning-level* comprising of the extent to which a training program has changed the participants' attitudes, affected their knowledge, and/or increased their skills, (3) *Behavior-level* which indicates the transfer of learning from an educational setting to real life, and (4) *Results-level* which looks at the impact on the organization and patient outcome as a consequence of the training program.

The study was undertaken at an obstetric department that had between 3,285 and 3,686 deliveries during the years 2003–2006, with approximately one-third being high-risk referrals. The study period was June 2003–June 2006. A total of 220 staff members (49 doctors, 105 midwives, 24 auxiliary nurses, and 42 nurses) were employed in the department during the training period.

Two training periods were undertaken consecutively. Details for the organization and the structure of the obstetric skills training program is shown in Figure 1. Each training session had 12 participants and was scheduled for two-and-a-half hours. It included a lecture followed by multi-professional training workshops with six participants in each workshop. Training material was developed by the local steering committee. A hierarchical task analysis was used, i.e. a task is broken down into a series of subtasks, as it is well suited to the development of training material and can be used to identify training needs, specify training objectives and elaborating training contents. It provides a logical rather than a psychological analysis of a task (15,16).

All the obstetric training scenarios were conducted in various rooms near the labor ward and not in a skills center. The training scenarios included simulated environments such as a delivery or a baby mannequin combined with relevant equipment, such as drips, medicines, catheters, and suction devices.

The training program was mandatory for all staff and was planned either during normal working hours or if outside normal working hours, the participants were paid extra per hour. All costs were covered by the departmental budget.

The main outcome measures were related to Kirkpatricks four level approach, and data for the analysis was collected before, immediately after and 9–15 months following the training. Most of the questionnaires used closed questions, with responses marked on a 5-point Likert scale as follows:

Kirkpatricks level 1 (Reaction): Measurement of the reaction of the participants toward the training program was obtained by questionnaires focusing on issues such as participants' opinion and perception of relevance of the training program.

Kirkpatricks level 2 (Learning): The value of participants learning was measured by participants self-assessed confidence (rated on a 5-point Likert scale from not confident = 1 to fully confident = 5) in performing specified procedures and self-assessed statements on whether clinical events were considered stressful and unpleasant (rated on a 5-point Likert scale from strongly disagree = 1 to strongly agree = 5). In addition, a written test of knowledge of skills (KOS) regarding basic neonatal resuscitation was undertaken. The KOS-test was composed of 7–8 short cases based on clinical scenarios (2–4 lines) followed by 3–6 questions with a yes/no/do not know response format. The test was administered pre- and twice post-training. The KOS test was chosen over an objective skills tests, as it could test many participants in a relatively short time and at low costs (17) and furthermore because it is reported to predict results in performance-based tests (17). The KOS-test was not developed as an instrument for individual testing, but was chosen to

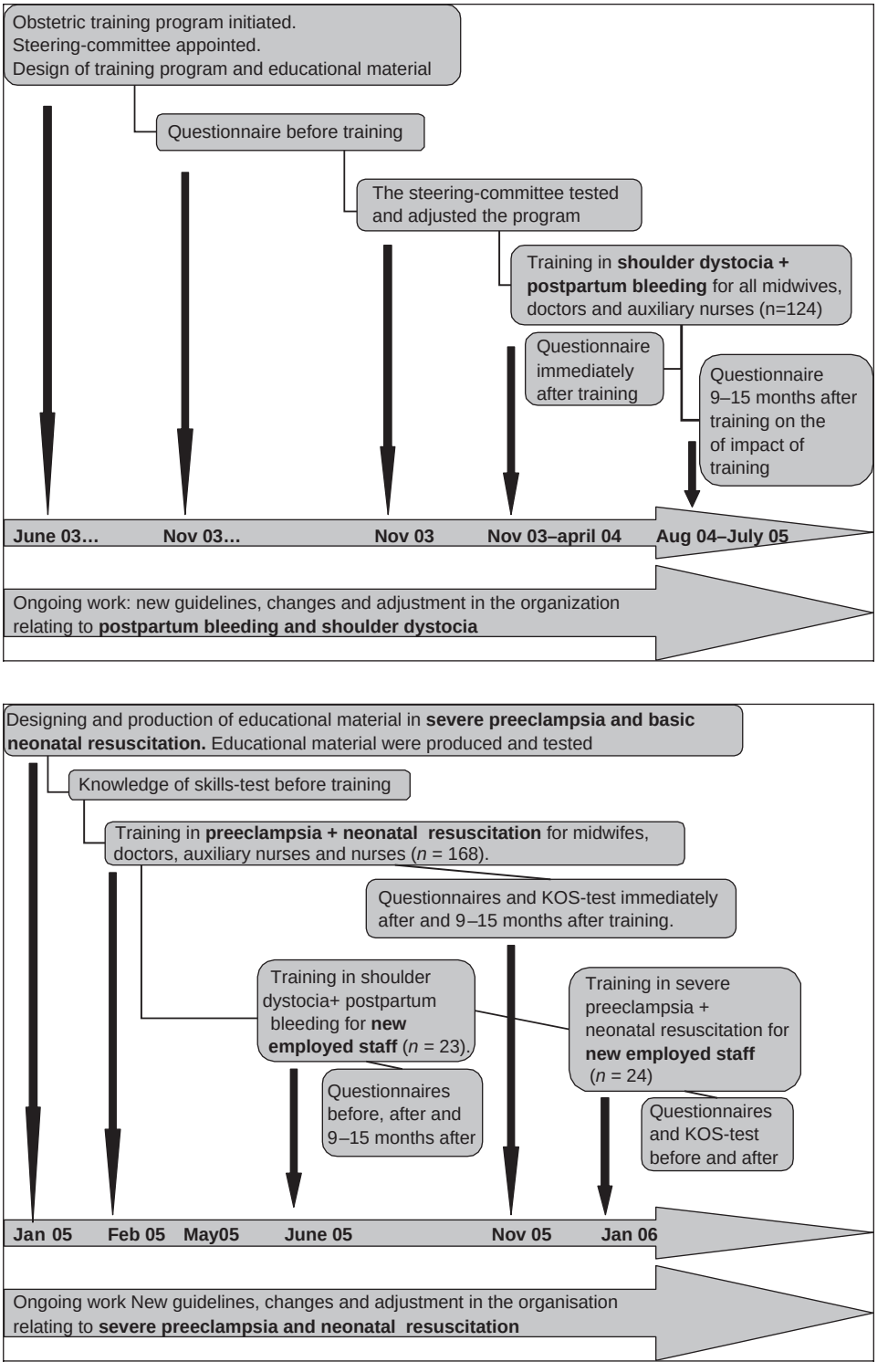

Figure 1. Flowchart of the obstetric skills training program over two training periods.

test variations in different staff groups over a period of time (18).

Kirkpatricks level 3 (Behavior): Data on behavior were collected by a questionnaire with semi-structured, open-ended questions to obtain information on how work routines were influenced by the training on the individual level.

Kirkpatricks level 4 (Results): Data were collected following each training session about the need for changes in working procedures at the organizational

level. Data from the Danish Medical Birth Registry were obtained to give information on changes in the diagnosis. Data were obtained from Hospital Administration about midwives' sick leave. Sick leave was measured as number of hours sick leave over hours of work per year.

An application for approval was filed to the Danish National Committee on Biomedical Research Ethics (number 16752). However, since the study did not involve patients, no approval was needed.

Each participant was given a project number, which was only known to the main author. The participants were ensured that during analysis and reporting, data would be treated as non-traceable information.

Statistical analysis

Quantitative data were analyzed using the Statistical Package for Social Sciences and SCD/DIGRAM (19). Log-linear chain graph models were used to analyze the correlations among responses to the questions. Most of the responses were measured on an ordinal 5-point rating scale, a Likert-like scale, by partial gamma coefficients, or partial rank correlations. The questions on whether or not attitudes and experiences in relation to the work had been influenced by training were tested by marginal and conditional homogeneity. The significance of the *p*-values was assessed, taking the association between the repeated measurements into account. The objective test by 'knowledge of skills tests' was analyzed by pairwise *t*-tests. Data from the Danish Medical Birth Registry were analyzed with odds ratios and 95% confidence intervals. Data from hospital statistics were analyzed with a chi-squared test for trend.

The qualitative data consisted of open-ended questions, which were analyzed by two investigators, i.e. the first author (JLS) and a person with no connection to the course and study. The two investigators together developed categories in order to quantify and condense the results for analysis.

Results

As regards management of shoulder dystocia and postpartum bleeding, 94% (147/156) of eligible staff participated and in training on the management of preeclampsia and neonatal resuscitation, 96% (192/201) participated (Tables I and II). Thirty training sessions were conducted.

Response rates (Tables I and II) to the questionnaires before training varied from 88 to 98% between the different health professional groups. Immediately after training, the questionnaire response rate within the different health professional groups was 89–100% and 70–100% for the questionnaires delivered 9–15 months after training.

Response rates to the KOS-test in neonatal resuscitation before and immediately after training were 98% (190/192). For late testing, only 168 staff members were eligible (24 staff members were included too late for the late post-testing to occur). The late response rate was 76% (128/168).

The participants who were eligible for training, but did not manage to attend the training, did not differ from the participants involved in the training. Non-responders of the questionnaires did not differ concerning years of obstetric work experiences, self-assessed confidence, and numbers of correct answers in the first KOS-test.

Kirkpatrick level 1: reaction

Of the participants, 92–98% agreed or strongly agreed that each of the four training programs were good, and more than 80–92% agreed or strongly agreed that they considered the training relevant for their clinical work. The majority (85–95%) of all staff considered the multi-professional organization of the training program to be good or very good.

Kirkpatrick level 2: learning

Confidence scores for all the trained skills improved significantly when measured 9–15 months following training (Table I). Scores for confidence in management of postpartum bleeding did not improve immediately after training, but significantly when measured 9–15 months following training. An analysis of how confidence was retained revealed that for management of shoulder dystocia and severe preeclampsia, no significant reduction in the level of confidence was found when early post-testing was compared with late post-testing 9–15 months following training. However, for neonatal resuscitation, there was a significant reduction in confidence ($p < 0.001$) from early post-testing compared with late post-testing. A greater increase in the score for high self-assessment in confidence amongst the trainee doctors, auxiliary nurses, midwives, and nurses was found and less so amongst specialized doctors and specialized midwives (data not shown). Self-assessment of confidence (scores 4 and 5) in management of shoulder dystocia was significantly increased for trainee doctors with a rise from 15 to 50%, for specialized midwives from 89 to 100% and for midwives from 22 to 56%. Self-assessment of confidence (score 4 + 5) with clinical management of postpartum bleeding increased significantly for trainee doctors

Table I. Self-assessment of confidence before, just after and 9–15 months following training in the management of four obstetric emergency skills. The replies were rated on a 5-point Likert scale from not confident = 1 to fully confident = 5.

Trained obstetric emergency	Management of shoulder dystocia						Management of postpartum bleeding						Management of severe preeclampsia						Basic neonatal resuscitation					
Eligible for training	156[a]						156[a]						201[a]						201[a]					
Participated	147						147						192						192					
Questionnaires	Before training		After training		9–15 months after		Before training		After training		9–15 months after		Before training		After training		9–15 months after		Before training		After training		9–15 months after	
Total respondents	158[b]	100 (%)	139	100 (%)	127	101 (%)	165[b]	101 (%)	139	100 (%)	129	100 (%)	199[b]	101 (%)	187	99 (%)	141	100 (%)	201[b]	100 (%)	188	99 (%)	140	101 (%)
Not confident 1	20	13	1	1	2	2	6	4	1	1	0	0	19	10	2	1	4	3	46	23	0	0	4	3
Confidence 2	30	19	7	5	6	5	9	5	6	4	0	0	38	19	8	4	12	9	35	17	4	2	15	11
Confidence 3	50	32	35	25	34	27	46	28	30	22	24	19	63	32	66	35	47	33	66	33	40	21	53	38
Confidence 4	48	30	86	62	75	59	66	40	78	56	78	60	55	28	99	53	64	45	34	17	123	65	60	43
Fully confident 5	10	6	10	7	10	8	38	23	24	17	27	21	24	12	12	6	14	10	20	10	21	11	8	6
Not eligible[c]																	24						24	
No answer or don't know			8		20				8		18				5		27				5		27	
p-Value[d]			<0.001[d]						0.007[d]						<0.001[d]						0.001[d]			
Gamma γ[e]			0.35[e]						0.75[e]						0.19[e]						0.51[e]			

[a]Nurses: only training in management of preeclampsia and basic neonatal resuscitation.
[b]Not all staff that answered the first questionnaire was eligible for the training as they left the department.
[c]Some were included late in the training period and not eligible for 9–15 month questionnaire and test.
[d]Paired test for conditional and marginal homogeneity. Data before training are compared with data 9–15 month following training.
[e]Partial gamma coefficients are calculated in different strata of a multidimensional contingency table. A high gamma ($\gamma > 0.30$) indicates high correlation between variables, and a low ($\gamma < 0.15$) little or no correlation.

Table II. Questionnaire to whether management of obstetric emergencies were considered stressful and unpleasant before and 9–15 months following training. The replies were rated on a 5-point Likert scale from strongly disagree =1 to strongly agree =5.

Trained obstetric emergency	Management of shoulder dystocia				Management of postpartum bleeding				Management of severe preeclampsia and eclampsia				Basic neonatal resuscitation			
Eligible in training	156[a]				156[a]				201[a]				201[a]			
Participated	147				147				192				192			
Questionnaires	Before training		9–15 months after training		Before training		9–15 months after training		Before training		9–15 months after training		Before training		9–15 months after training	
Total respondents (%)	155[b]	101 (%)	126	100 (%)	161[b]	100 (%)	128	100 (%)	194[b]	99 (%)	141	100 (%)	192	100 (%)	140	100 (%)
Strongly disagree	12	8	14	11	23	14	22	17	16	8	21	15	7	4	6	4
Disagree	49	32	50	40	87	54	71	56	71	37	73	52	41	21	49	35
Uncertainty	29	19	27	21	23	14	18	14	49	25	27	19	40	21	44	31
Agree	50	32	30	24	24	15	14	11	45	23	16	11	56	29	34	24
Strongly agree	15	10	5	4	4	2	3	2	13	7	4	3	48	25	7	5
Not eligible[c]											24				24	
No answer or don't know			21				21				23				23	
Don't know			0				0				4				5	
p-Value[d]			0.003[d]				0.24[d]				<0.001[d]				<0.001[d]	
Gamma γ[e]			0.52[e]				0.29[e]				0.53[e]				0.41[e]	

[a]Nurses: only training in management of preeclampsia and basic neonatal resuscitation.
[b]Not all staff that answered the first questionnaire was eligible for the training as they left the department.
[c]Some were included late in the training period and not eligible for 9–15 month questionnaire and test.
[d]Paired test for conditional and marginal homogeneity. Data before training are compared with data 9-15 month following training.
[e]Partial gamma coefficients are calculated in different strata of a multidimensional contingency table. A high gamma ($\gamma > 0.30$) indicates high correlation between variables, and a low ($\gamma < 0.15$) little or no correlation. Nurses: only training in management of preeclampsia and basic neonatal resuscitation.]

from 48 to 100% and for auxiliary nurses from 22 to 50%. In the management of severe preeclampsia, the increase was significant for trainee doctors from 33 to 43%. For management of basic neonatal resuscitation, the increase was significant for trainee doctors from 20 to 36%, for midwives from 24 to 53%, and for nurses from 13 to 21%.

Management of shoulder dystocia, severe preeclampsia, and neonatal resuscitation was considered stressful and unpleasent to perform before training, but the level of discomfort was significantly less following training for all health professional groups (Table II). In comparison, no difference was found for management of postpartum bleeding. Data on how the different health professional groups experienced the different skills in terms of being stressful and unpleasant showed for the auxiliary nurses, the trainee doctors, the midwives, and the nurses, a reduction in how stressful and unpleasant a skill was perceived. Less or no reductions were found for the specialized doctors and the specialized midwifes (data not shown).

For midwives, all trained skills except for management of postpartum bleeding were scored significantly less stressful and less unpleasent to perform 9–15 months following training.

Nurses were only trained in management of severe preclampsia and basic neonatal resuscitation. For both skills the nurses 9–15 months following training scored significantly lower in how stressful and unpleasant they considered performance of the skills.

KOS-test

The number of correct answers in the KOS-test in basic neonatal resuscitation increased significantly ($p < 0.001$) from 65% before training to 94% at early post-testing. There was a significant reduction ($p < 0.001$) in numbers of correct answers at the time of early post-testing compared with the late post-test results. However, at the time of late post-testing, it was still significantly higher ($p < 0.001$) than at the time of pre-testing.

The lowest average numbers of correct answers before training were found among auxiliary nurses (50%) and nurses (52%), whereas the average of correct answers before training were 70% for midwives, 75% for specialized midwives, 68% for trainee doctors, and 72% for specialized doctors. At the time of early post-testing, numbers of correct answers were 88–96%, with lowest numbers for auxiliary nurses and highest score among trainee doctors. Figure 2 shows a significant association between high confidence score before training and increasing numbers of correct answers in the KOS-test. The association between years of clinical work in obstetrics and the numbers of correct answers in the KOS-test can be seen in Figure 3. The partial gamma coefficient was 0.04, which indicated no correlation.

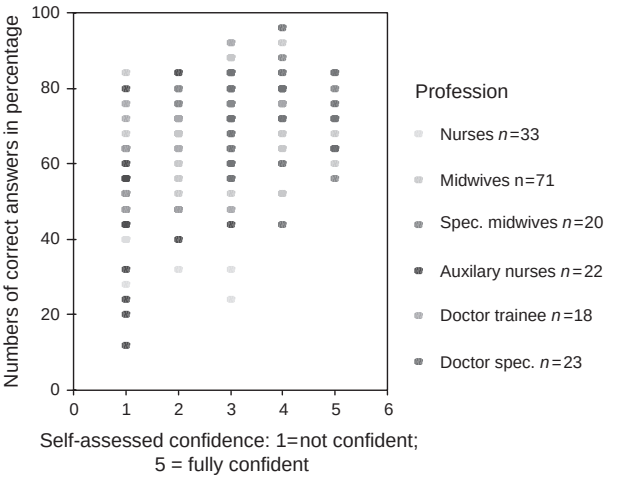

Figure 2. The association between self-assessed confidence and numbers of correct answers in KOS-test in basic neonatal resuscitation according to health professional group.
Statistical analysis: Partial gamma coefficients showed association between self-assessment score in confidence and number of correct answers in the KOS-test in neonatal resuscitation ($\gamma = 0.44$).
A significant association was seen between self-assessment and the numbers of correct answers in the KOS-test in neonatal resuscitation ($p < 0.001$, test of marginal and conditional homogeneity).

Kirkpatrick level 3: behavior

The open-ended questions 9–15 months following the obstetric skills training in shoulder dystocia and postpartum bleeding showed that 89% of midwives (74/83), 94% (15/16) of auxiliary nurses, and 70% (23/33) of the doctors reported that training had a positive influence on their work.

Work-influence category results were as follows: 'feeling of confidence, safety and security' = 36–40%, 'coordination of management, better multi-professional effort, staff taking responsibility' = 26–27%, and 'changes in clinical management' = 23–24%.

Kirkpatrick level 4: results

Table III shows information obtained from the steering committee after every training session on the need for changes at the organizational level. These suggestions, along with considerations on implementation, were discussed in the steering committee and with the management team of the Obstetric Department and actions were taken.

Figure 3. The association between years of obstetrical work experience and correct answers in the KOS-test in neonatal resuscitation according to health professional group.
Statistical analysis: Partial gamma coefficients showed no association between self-assessment score and number of correct answers in the KOS-test in neonatal resuscitation ($\gamma = 0.04$ and no statistical significance $p = 0.286$). No further statistical analysis was performed.

Prevalence of obstetric emergencies

During the study period a significant increase ($p < 0.001$) of 42% in frequency of the ICD-code for postpartum bleeding was found. In 2002–2003, the ICD-code for postpartum bleeding was used in 3.4% (215/6,356) of the deliveries compared with 4.7% (309/6,506) of the deliveries in 2004–2005 (OR = 1.42; 95% confidence interval 1.19–1.70). Nationally, a significant increase in the use of this ICD code also was observed (OR = 1.07; 95% confidence interval 1.03–1.12).

The administration of uterotonics was significantly increased at 24% during the study period and was given to 4.5% (287/6,356) in 2002/2003 compared to 5.6% (365/6,506) in 2004/2005 (OR = 1.24; 95% confidence interval 1.07–1.42). Nationally this prevalence was unaltered with an OR = 0.9; 95% confidence interval 0.87–0.92.

Sick leave among midwives

Midwives sick leave dropped significantly ($p = 0.002$) from 8% (7,017/87,712) in 2000 to 2.9% (2,449/85,500) in 2006.

Discussion

An evaluation of the present mandatory simulation-based training program in a large obstetric department demonstrated a positive impact on the participant's satisfaction with the program, aspects of gaining in learning, and changes in work routines. At the same time an impact on the organizational level was found with changes in guidelines and equipment in the labor ward, reduced sick leave among midwives, and data on the prevalence of some obstetric emergencies, that seems to suggest that the identification and management of postpartum bleeding was given more focus following training.

Table III. Changes after implementation of the obstetric training program.

- Clinical guidelines in postpartum bleeding, shoulder dystocia and preeclampsia were composed or updated.
- An algorithm for basic neonatal resuscitation was designed in accordance with international guidelines. This was done in cooperation with the Department of Anesthesia and the Department of Neonatology.
- Highlights were extracted from guidelines and laminated action cards were designed and placed on open shelves in every labor ward suite.
- Filled in forms to order blood tests were made in a laminated form with information on which blood tests to order in emergency cases. These forms were placed on open shelves in every labor ward suite.
- A new observation form chart form was designed. It was developed in an attempt to achieve the loss of observations that seemed to happen on transfer between the labor ward and the postnatal ward.
- 'Bleeding boxes' and 'Preeclampsia boxes' with the most important medications were made. These boxes were placed in the labor ward, operating theatre and the postnatal ward.
- For neonatal resuscitation new self-expanding bags and masks were introduced. They were considered to be safer and more applicable for staff members who rarely use it.
- New clocks were placed over the delivery bed in every labor suite to ensure and make time taking easier.
- A condensed version of the teaching materials used during the training sessions were laminated and placed in folders in the labor ward easily accessible to all labor ward staff members.
- The cupboards in all the labor ward suites were organized according to the same standard.

Implementation of a training program in a clinical setting on a large scale is complex (20). The involvement of staff at all hierarchical levels, as well as support from the management team, were considered to be prerequisites for a successful implementation. The implementation of a mandatory training program was proven feasible as 94–96% of the eligible staff members participated. For a training program like this, it is essential to have a continuously updated database on staff members. The organization including the management team needs to be able to cope with designing and implementing such a multi-professional training program (9,21).

The data revealed that the participants had a positive attitude of the training program and considered it to be relevant. No major differences were obtained between the different health professional groups (Kirkpatrick level 1, reaction). Looking at the scores of confidence, the participating staff scored

higher at the time of early post-testing compared with pre-testing and confidence seemed to be retained over a period of 9–15 months, except for management of basic neonatal resuscitation. This may be explained by that the skill in managing neonatal resuscitation is less frequently needed compared with any other skills trained. Self-assessed confidence in each skill among midwives, trainee doctors, auxiliary nurses, and nurses improved more than it did among specialized doctors and specialized midwives, presumably due to less potential capacity for improvement existing in these two groups, as these scored higher before training (Kirkpatrick level 2, learning).

Validity of self-assessment is disputable and there is a difference between confidence and competence (22,23). Improvement in confidence is not necessarily translated into better competence and better outcome. In the present study a significant association was found between levels of confidence in the management of basic neonatal resuscitation and in the numbers of correct answers in the KOS-test. This result may indicate that the participants were capable of assessing their own competence in basic neonatal resuscitation skills. Whether this can be generalized to other areas remains to be investigated. In the present study a tendency towards participants' underestimation of own skills was seen. It has been hypothesized that people who underestimate their own competences are at risk of burning out. Some argue that guided self-assessment should be incorporated as an essential professional skill in continuous professional development (22).

No association was found between a high number of correct answers in the KOS-test and many years of work experience. Due to this lack of association, the skills that are required in basic neonatal resuscitation do not seem to be learnt and/or retained with years of work experience. This is in accordance with a previous review (24), where it was concluded that in half of previous studies there was an inverse relation between years of work experience and the quality of care that doctors provided.

Staff members reported the management of shoulder dystocia, severe preeclampsia, and basic neonatal resuscitation to be less stressful and less unpleasant following training. No difference was observed with regard to management of postpartum bleeding. One possible explanation for this may be that postpartum bleeding was scored less stressful before training, probably because it is a more frequently experienced clinical event than any of the other three skills trained.

Looking at the different health professional groups, significant improvements in general were observed for nurses and midwives. Less or no improvements were found for specialist doctors. However, they still approved the training program.

Relatively, little research has been carried out on stress levels among labor ward staff (25). In future studies, whether randomized or before-after studies, validated scales for burn-out scores or work-related stress could be applied. The KOS-test of neonatal resuscitation showed a significant increase in the numbers of correct answers after training. Although a drop back was observed 9–15 months following training, the score was still significantly higher than the pre-test score. Whether the KOS-test only reflects the performance in the simulated setting and not real clinical life remains to be investigated. An important question to address is when to reassess knowledge and skills, as immediate post-testing has been reported to be insufficient (26). The literature about retention of skills is sparse, but maybe retention is low already after 3–10 months post-training (27,28).

Behavior or transfer of learning (Kirkpatrick, level 3) is generally considered a difficult issue to address (13,14). In this study we concluded that the training did have a positive influence on work performance.

The need for organizational changes became obvious during the training period and the motivation to implement these changes was clearly present at all hierarchical levels of the organization (Kirkpatrick level 4). There was a significant increase in the prevalence of postpartum bleeding during the study period and no such increase was observed at a national level. It can be argued that the training resulted in a greater attention to cases with postpartum bleeding and that the staff, for this reason, more often used the diagnosis. The use of uterotonics also increased significantly during the training period.

The reduction in sick leave among midwives cannot be assigned causality in the present study. However the management team of the obstetric department considered the implementation of the training program to have played an important role on the observed reduction in sick leave.

A limitation of the present study was the lack of randomization and control groups. The present study design can be viewed as a 'compromise design', which is often seen in educational research (29). It can be argued that the changes observed following the training program may have happened even without the training program. The project was only carried out in one hospital setting, which gave local possibilities for measuring effects, but at the same time raised the questions about external validity of the study and on whether the project

would be applicable in other settings. In future studies on multi-professional simulation-based training in labor wards, a multicenter approach could be relevant.

The outcome measures fall short of direct observation of management in real emergencies, as the trained events occur rarely. The prevalence figures for maternal and fetal morbidity and mortality are so low that any significant changes may only be revealed in large multicenter studies. In obstetric research very high numbers of deliveries are required to measure direct outcomes (30).

Anderson et al. (21) suggested that methods to assess and optimize obstetric training are urgently required in order for women and their babies to benefit from these expensive and complex interventions. Evaluation should be integrated as an important part of all training activities, but evaluations are often inconsistent, or even lacking (13,14). In the present study we chose a broad range of evaluation tools and found these to complement each other. The study gave a good basis for changes at a local organizational level.

There are advantages in courses run on local basis, as they reduce costs and increase the access to training, compared to courses run on a national or international level, while the disadvantages could be problems with organization and quality of content, which suggests that locally organized training programs might be overseen at a national level to ensure quality (7). We found it was an advantage to develop and implement the present training program on a local basis, and that a skills center not was prerequisite for implementation. We agree as others (12) that skill centers and simulators are only of value within the context of a program or a total educational curriculum.

This mandatory obstetric training program thus had an impact on both the individual and on the organizational level. Implementation of the training program was feasible without a skills center**.** A vital resource behind the success described in this study was the development and implementation of a training program that involved representatives of all health professionals, staff with educational and obstetric competencies, plus a supportive management team.

Acknowledgements

Professor emeritus Knut Aspegren gave guidance on qualitative data analysis. Steen Rasmussen, National Board of Health, Denmark and Jens Langhoff-Ross, Copenhagen University Hospital, Rigshospitalet contributed with data from the Danish Medical Birth Registry. Statistical supervision was organized through the Advisory Service in Clinical Research Methodology and Medical Statistics, Department of Biostatistics, Faculty of Health Sciences, University of Copenhagen, Denmark.

Declaration of interest: The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

References

1. Drife J. Reducing risk in obstetrics. Qual Health Care. 1995; 4:108–14.
2. Johannsson H, Ayida G, Sadler C. Faking it? Simulation in the training of obstetricians and gynaecologists. Curr Opin Obstet Gynecol. 2005;17:557–61.
3. Veltman LL. Getting to havarti: moving toward patient safety in obstetrics. Obstet Gynecol. 2007;110:1146–50.
4. Guise JM. Anticipating and responding to obstetric emergencies. Best Pract Res Clin Obstet Gynaecol. 2007;21:625–38.
5. Kristensen TS, Hannerz H, Hogh A, Borg V. The Copenhagen Psychosocial Questionnaire – a tool for the assessment and improvement of the psychosocial work environment. Scand J Work. Environ & Health. 2005;31:438–49.
6. Ashcroft B, Elstein M, Boreham N, Holm S. Prospective semi structured observational study to identify risk attributable to staff deployment, training, and updating opportunities for midwives. BMJ. 2003;327:584–587.
7. Black RS, Brocklehurst P. A systematic review of training in acute obstetric emergencies. BJOG. 2003;110:837–41.
8. Macedonia CR, Gherman RB, Satin AJ. Simulation laboratories for training in obstetrics and gynecology. Obstet Gynecol. 2003;102:388–92.
9. Draycott T, Sibanda T, Owen L, Akande V, Winter C, Reading S, et al. Does training in obstetric emergencies improve neonatal outcome? BJOG. 2006;113:177–82.
10. Crofts JF, Ellis D, Draycott TJ, Winter C, Hunt LP, Akande VA. Change in knowledge of midwives and obstetricians following obstetric emergency training: a randomised controlled trial of local hospital, simulation centre and teamwork training. BJOG. 2007;114:1534–41.
11. Ellis D, Crofts JF, Hunt LP, Read M, Fox R, James M. Hospital, simulation center, and teamwork training for eclampsia management. A randomised controlled trial. Obstet Gynecol. 2008;111:723–31.
12. Issenberg SB, McGaghie WC, Petrusa ER, Lee GD, Scalese RJ. Features and uses of high-fidelity medical simulations that lead to effective learning: a BEME systematic review. Med Teacher. 2005;27:10–28.
13. Kirkpatrick DL. Evaluating training programs. San Francisco, CA: Berrett-Koehler Publishers; 1998.
14. Morrison J. ABC of learning and teaching in medicine: evaluation. BMJ. 2003;326:385–7.
15. McLeod PJ, Steinert Y, Trudel J, Gottesman R. Seven principles for teaching procedural and technical skills. Acad Med: J Assoc Amer Med Colleges. 2001;76:1080.
16. Stewart A. Instructional design. A practical guide for medical teachers. In: Dent JA, Harden RM, editors. Edinburgh: Churchill Livingstone; 2001. pp. 195–204.
17. Remmen R, Scherpbier A, Denekens J, Derese A, Hermann I, Hoogenboom R, et al. Correlation of a written test of skills and a performance based test: a study in two traditional medical schools. Med Teacher. 2001;23:29–32.

18. Kramer AW, Jansen JJ, Zuithoff P, Dusman H, Tan LH, Grol RP, et al. Predictive validity of a written knowledge test of skills for an OSCE in postgraduate training for general practice. Med Educ. 2002;36:812–9.
19. Kreiner S. Introduction to DIGRAM. Research report 2003;03:10. Department of Biostatistics, Faculty of Health Sciences, University of Copenhagen, Denmark.
20. Campbell M, Fitzpatrick R, Haines A, Kinmonth AL, Sandercock P, Spiegelhalter D, et al. Framework for design and evaluation of complex intervention to improve health. BMJ. 2000;321:694–6.
21. Anderson ER, Black R, Brocklehurst P. Acute obstetric emergency drill in England and Wales: a survey of practice. BJOG. 2005;112:372–5.
22. Duffy FD, Holmboe ES. Self-assessment in lifelong learning and improving performance in practice. Physician know thyself. JAMA. 2006;296:1137–9.
23. Ward M, Gruppen L, Regehr G. Measuring self-assessment: current state of the art. Adv Health Sci Educ: Theor Prac. 2002;7:63–80.
24. Choudry NK, Fletcher RH, Soumerai SB. Systematic review: the relationship between clinical experience and quality of health care. Ann Intern Med. 2005;142:260–73.
25. Engelbrecht S. Motivation and burnout in human service work. The case of midwifery in Denmark. Thesis/Dissertation, National Institute of Occupational Health, Copenhagen and Roskilde University, Faculty of Psychology, Philosophy and Science Studies, 2006.
26. Arthur W, Bennett W, Stanush PL, Mc Nelly TL. Factors that influence skill decay and retention: a quantitative review and analysis. Human Performance. 1998;11:57–101.
27. Hamilton R. Nurses' knowledge and skill retention following cardiopulmonary resuscitation training: a review of the literature. J Adv Nurs. 2005;51:288–97.
28. Trevisanuto D, Ferrarese P, Cavicchioli P, Fasson A, Zanardo V, Zacchello F. Knowledge gained by pediatric residents after neonatal resuscitation program courses. Paediatr Anaesth. 2005;15:944–7.
29. Cohen L, Manion l, Morrison K. Research methods in education. Routhledgefalmer: London and New York; 2003.
30. Mongelli M, Chung TK, Chang AM. Obstetric intervention and benefit in conditions of very low prevalence. BJOG. 1997; 104:771–4.