

# Fetal biometry in ultrasound: a new approach to assess the long-term impact of simulation on learning patterns

G. Ambroise Grandjean<sup>1,2,3</sup>, C. Bertholdt<sup>1,2</sup>, S. Zuily<sup>4</sup>, M. Fauvel<sup>5</sup>, G. Hossu<sup>5</sup>, P. Berveiller<sup>6,7</sup>, O. Morel<sup>1,2</sup>

<sup>1</sup> Université de Lorraine, IADI - INSERM, F-54000, Nancy, France

<sup>2</sup> CHRU Nancy, Department of Obstetrics and Gynecology, F-54000 Nancy, France

<sup>3</sup> Université de Lorraine, Midwifery Department, Nancy F-54000, France

<sup>4</sup> Université de Lorraine, Hôpital virtuel de Lorraine, Nancy F-54000, France

<sup>5</sup> CHRU Nancy, Université de Lorraine, CIC-IT, F-54000 Nancy, France

<sup>6</sup> CHI Poissy Saint-Germain-en-Laye, Department of Obstetrics and Gynecology, F-78300, Poissy, France

<sup>7</sup> Université Versailles Saint-Quentin, EA 7404 - Gamètes Implantation Gestation, F-78180, Montigny le Bretonneux, France

## ABSTRACT

**Context:** Simulation-based education (SBE) has demonstrated its acceptability and effectiveness in improving ultrasound training. Because of the high cost of its implementation (investment in equipment and supervision), a pragmatic assessment of the transfer of skills learned in SBE to clinical practice and the identification of its optimal scheduling conditions have been requested to optimize its input.

**Objectives:** To quantify the long-term impact of simulation-based education (SBE) on the adequate performance of ultrasound fetal biometry measurements (I). The secondary objective was to identify the temporal patterns that enhanced SBE input in learning (II).

**Methods:** Trainees were arbitrarily assigned to a 6-month course in obstetric ultrasound with or without an SBE workshop. In the SBE group, the workshop was implemented 'before' or at an 'early' or a 'late-stage' of the course. Those who did not receive SBE were the control group. The ultrasound skills of all trainees were prospectively collected, evaluated by calculating the delta between OSAUS (Objective Structured Assessment of Ultrasound Skills) scores before and after the course (I). Concomitantly, the accuracy of trainees' measurements was assessed throughout the course by verifying their correlation with the corresponding measurements by their supervisors. The percentage of trainees able to perform five consecutive sets of correct measurements in the control group and in each SBE subgroup were compared (II).

**Results:** The study included 61 trainees (39 SBE and 22 controls). Comparisons between groups showed no significant difference in the quantitative assessment of skill enhancement (difference in the pre- and post-internship OSAUS score:  $1.09 \pm 0.87$  in the SBE group and  $0.72 \pm 0.98$  in the control group) (I). Conversely, the predefined acceptable skill level was reached by a significantly higher proportion of trainees in the 'early' SBE subgroup (74%, compared with 30% in the control group,  $P < 0.01$ ) (II).

**Conclusions:** The quantitative assessment does not support the existence of long-term benefits from SBE training, although the qualitative assessment confirmed SBE helped to raise the minimal level within a group when embedded in an 'early' stage of a practical course.

1  
2  
3 **Context:** Simulation-based education (SBE) has demonstrated its acceptability and effectiveness in improving  
4 ultrasound training. Because of the high cost of its implementation (investment in equipment and supervision), a  
5 pragmatic assessment of the transfer of skills learned in SBE to clinical practice and the identification of its  
6 optimal scheduling conditions have been requested to optimize its input.

7 **Objectives:** To quantify the long-term impact of simulation-based education (SBE) on the adequate performance  
8 of ultrasound fetal biometry measurements (I). The secondary objective was to identify the temporal patterns  
9 that enhanced SBE input in learning (II).

10 **Methods:** Trainees were arbitrarily assigned to a 6-month course in obstetric ultrasound with or without an SBE  
11 workshop. In the SBE group, the training was implemented 'before' or at an 'early' or a 'late-stage' of the course.  
12 Those who did not receive SBE were the control group. The ultrasound skills of all trainees were prospectively  
13 collected, evaluated by calculating the delta between OSAUS (Objective Structured Assessment of Ultrasound  
14 Skills) scores before and after the course (I). Concomitantly, the accuracy of trainees' measurements was  
15 assessed throughout the course by verifying their correlation with the corresponding measurements by their  
16 supervisors. The percentage of trainees able to perform five consecutive sets of correct measurements in the  
17 control group and in each SBE subgroup were compared (II).

18 **Results:** The study included 61 trainees (39 SBE and 22 controls) were recruited. Comparisons between groups  
19 showed no significant difference in the quantitative assessment of skill enhancement (difference in the pre- and  
20 post-internship OSAUS score:  $1.09 \pm 0.87$  in the SBE group and  $0.72 \pm 0.98$  in the control group) (I). Conversely,  
21 the predefined acceptable skill level was reached by a significantly higher proportion of trainees in the 'early' SBE  
22 subgroup (74%, compared with 30% in the control group,  $P=0.01$ )(II).

23 **Conclusions:** The quantitative assessment does not support the existence of long-term benefits from SBE  
24 training, although the qualitative assessment confirmed SBE helped to raise the minimal level within a group  
25 when embedded in an 'early' stage of a practical course.

## 26 27 28 **INTRODUCTION**

29  
30 Ultrasound is used in many specialties because of its accessibility (1,2) and the ability to study items moving in  
31 real time at the patient's bedside (3). These advantages give it a key role in the monitoring and management of  
32 pregnancies, especially by fetal biometric measurements (4–9). It is nonetheless operator-dependent (10,11), a  
33 characteristic that causes errors and inaccuracy that can account for more than 10% of the variability of some  
34 parameter measurements (11). Improving intra- and interobserver reproducibility is therefore a major objective  
35 in obstetrics.

36 As we await the automated measurement techniques currently under development (12–14), it remains  
37 essential to improve the teaching of fetal biometry to limit the variability and optimize the quality and quality  
38 control of these measurements (15). High-fidelity simulation-based education (SBE) in ultrasound appears to be a  
39 promising solution (16). Its acceptability and its short-term utility for improving performance have now been  
40 established (17–21), although its benefits for the quality and cost of obstetric care have not been clearly  
41 confirmed (22). Currently, only Tolsgaard et al. have used an experimental approach to explore these aspects,  
42 offering pathways for assessing the medico economical impact of SBE (23). Their approach relies especially on  
43 the demonstration of a potential synergistic effect of SBE (16) that could make it possible to accelerate progress  
44 for the same quantity of teaching.

45 SBE remains expensive (both the equipment and logistics). It is therefore important to verify how the  
46 positive impacts of SBE on progression modify clinical practice in the long term and to assess the importance of  
47 the potential synergistic effect as a function of the timing of its use during the learning process.

48 Accordingly, the principal objective of our study was to compare the clinical progress over the long term  
49 between students who were or were not exposed to simulation training. The secondary objective was to study

50 the effect of the timing of the exposure (by comparing the different timing of the SBE workshops relative to the  
51 course) on the capacity to acquire and maintain the level of competence required.

52

## 53 **METHODS**

54

55 A systematic, prospective, and continuous collection of data about the practice of ultrasound for measuring fetal  
56 biometry during the initial phase of learning these skills took place at the University Hospital of Nancy in the  
57 maternity ward between November 1, 2017, and October 31, 2018.

58

### 59 **Participants**

60 The participants were student sonographers recruited during their initial training in the departments of  
61 obstetrics and gynecology, medical gynecology, and midwifery of the University of Lorraine. The inclusion criteria  
62 were enrollment in the specialized diploma program in obstetrics and gynecology or medical gynecology (first  
63 through eighth semester of their internship-residency, or enrollment in the second cycle of their midwifery  
64 studies). The exclusion criteria were: the failure to undergo an initial assessment or absence from the course for  
65 a duration exceeding one month during the study period.

66

### 67 **Description of the METHOD project**

68 This study is part of a university project aimed at optimizing the teaching and learning of ultrasound  
69 measurements of fetal biometry (METHOD project: MEasurement TeachHing in Obstetrics Design). French  
70 students currently learn ultrasound as apprentices in practical clinical training courses. For the METHOD  
71 pedagogical project, workshops with ultrasound simulators were progressively included into this curriculum.  
72 None of the participants included in the project had previously received this type of instruction. Three teachers  
73 (OM, CB and GA) coordinated this project and ran the clinical course. During the courses, students were  
74 supervised by experienced clinicians (obstetricians or midwives). The simulation workshop was developed and  
75 led by an instructor trained in teaching by simulation (GA).

76

77

78

#### 79 - Theoretical instruction

80 Before the clinical training course, all participants attended in person a two-hour theoretical presentation, the  
81 same for all, about the methodology and practice of ultrasound fetal biometry.

82

83

#### 84 - Pedagogical program of the simulation workshops

85 As recommended by Tolgaard et al., the workshops were conducted in groups of two participants. During a  
86 single 3-hour workshop with a simulator (Simbionix US Mentor®), the participants practiced a series of  
87 increasingly difficult exercises under the constant supervision of the teacher (GA). The program included in  
88 particular exercises of hand-eye coordination and of knobology (24), as well as the performance of fetal  
89 biometry.

90

#### 91 - Pre- and post-course quantitative evaluations

92 All participants included in the study were assessed by an OSAUS score before and after the course (Fig 1). The  
93 conditions of a stand-alone examination were reproduced with pregnant women (15–41 weeks of pregnancy)  
94 who volunteered, and each participant was assessed individually by one of the instructors (OM, CB or GA) as they  
95 took ultrasound fetal biometry measurements. The METHOD OSAUS score (customized for fetal biometry) was  
96 used to rate their skill level from 0 through 5 (25). The examination could not exceed 10 minutes, and no  
97 supervision or suggestions were offered during it. The pre-workshop assessments were scheduled early during or  
98 before the course and always before the simulator workshop (double-blinded assessment, both instructor and  
99 participant blinded about exposure to the simulation). The post-course assessments were scheduled after the

100 last day of the course. To limit the potential evaluation bias associated with knowledge of the student's exposed-  
101 unexposed status for the post-course OSAUS scores, these were assessed by four clinicians not involved in the  
102 study (single-blinding).

103  
104 - Continuous qualitative assessment

105 An online platform (the METHOD Logbook) was available to participants to enable the continuous collection of  
106 data about the instruction. It enabled participants to report in real time the results of the fetal biometry  
107 measurements they performed themselves during their training course (head circumference - HC, abdominal  
108 circumference - AC, femur length – FL, and estimated fetal weight – EFW). These examinations were  
109 systematically verified by an experienced clinician. For each examination performed, the students recorded both  
110 their set of measurements and that of their supervising clinician in the METHOD logbook. The METHOD Logbook  
111 platform determined the validity of each parameter measured, based on the difference between the student's  
112 measurements and those of the supervisor, that is, on the agreement of the student's and supervisor's  
113 measurements, with the latter considered as the reference measurements. Validation depended on an  
114 interoperator variability less than 0.8 of the Z-score for the EFW and less than 0.8, 0.7 and 0.6 of the Z-scores for  
115 FL, AC and HC, respectively. These cutoffs were defined by taking into account the variability inherent to each  
116 biometric parameter. These differences in variability are explained by their different levels of technical difficulty  
117 and the different distributions for each type of measurement (11).

118 Each time students completed a set of measurements and entered it into the Logbook, they received  
119 feedback (for each parameter measurement and for EFW), presented as an LC–CUSUM (Learning Curve Using  
120 Cumulative Sum control charts) (26–28). The display setting for the METHOD Logbook showed a rising curve for  
121 each validated examination (strong agreement with the reference measurement) and a falling curve for each  
122 non-validated measurement (low agreement with the reference measurement) (Figure 2).

123  
124  
125 **Study plan and procedures**

126  
127 - Timing of exposure

128 The participants were randomly allocated to the control (no simulation workshop exposure during the  
129 observation period) and simulation (exposure to SBE) groups. The SBE workshops were planned randomly at  
130 different times on the course schedule, with the participants distributed within both the control and simulation  
131 (SBE exposure) groups. Within the simulation group, participants were distributed into three subgroups, the  
132 'before' group (exposure before the course), the 'early 'course group (interval between the beginning of the  
133 course and exposure < 2 months) and the 'late' course group (interval between exposure and the end of the  
134 course < 2 months); these groups defined different exposure timing patterns (Figure 1).

135  
136 - Endpoints

137 The primary endpoint was the percent delta (that is, percent change) in progression during the clinical training  
138 course, defined by the difference between the pre- and post-course OSAUS scores. We hypothesized that the  
139 mean change in progression differed between the simulation and control groups. The secondary endpoint was  
140 defined by the proportion of participants exposed to SBE able to acquire a threshold skill level according to their  
141 exposure timing pattern. This threshold level was defined by the validation of five consecutive EFW in the  
142 METHOD Logbook.

143  
144  
145 **Statistical analysis:**

146  
147 - Calculation of study size

148 In the usual learning conditions of initial training, the student's skill level is expected to progress. When this  
149 progression is assessed by a score, a delta/change between the scores obtained before and after the training  
150 expresses this progression (corresponding to the progression delta described above).

151 For the METHOD project, the expected progression was estimated from data from preliminary studies  
152 assessing the validity and predictive value of the METHOD OSAUS score (25). Based on these studies, the skill  
153 level in fetal biometry of a beginner in obstetric ultrasound corresponded to OSAUS scores ranging between  $1.5$   
154  $\pm 0.4$  and  $1.8 \pm 0.7$ . After practice experience for longer than 6 months or  $> 20$  supervised examinations, the  
155 operators reached an intermediate skill level corresponding to scores of  $3.3 \pm 0.6$  and  $3.3 \pm 0.8$  respectively  
156 (25,29).

157 Given these data, we estimated the expected progression of participants without SBE to be at least 1.25  
158 points on the OSAUS. Our hypothesis was that SBE would enable greater progress for the exposed participants  
159 (delta 1.75 points on the OSAUS). We calculated that we required a minimum of 20 participants in the exposed  
160 and unexposed groups to validate this hypothesis. In the SBE group, at least 8 participants were necessary in  
161 each subgroup to enable a comparative qualitative analysis between the different timing patterns. This required  
162 that we increase the size of the SBE group. [The modest headcount within subgroups did not allow to lead a](#)  
163 [relevant quantitative assessment for each timing pattern.](#)

164  
165 - Statistical tests  
166 The quantitative data are expressed as their means and standard deviations, with their range (minimum and  
167 maximum). The normality of the data was tested by the Shapiro Wilk test. Intergroup differences were studied  
168 by Student's t test (with equal variance). For the qualitative variables, Chi-square or Fisher tests were used  
169 according to the distribution of the data. A P value less than 0.05 was considered significant.

170 The statistical analysis was performed with R software, version 3.6.2 (2019-12-12).

171  
172 - Ethics and reporting of the results  
173 This study was reported to the CNIL (no. A2017-3). It is presented in accordance with the guidelines for Health  
174 Care Simulation Research and CONSORT (extension for pragmatic trials).

175  
176

## 177 RESULTS

178  
179 **Participants:**  
180 The study included 61 participants (36 residents and 25 midwifery students) (Figure 3). [Details of the cohort](#)  
181 [demographic characteristics are available in Table 1.](#) The mean observation period was 5 months and 15 days  $\pm$   
182 62 days.

183 After the allocation of the minimum number of individuals within each group and subgroup, there remained  
184 18 additional participants. These participants were distributed between the control group, and the early and late  
185 course subgroups (3, 14 and 1 additional participant, respectively). The distribution of the additional participants  
186 favored, for pedagogical reasons, early access to SBE.

187 At the beginning of the study, the pre-course OSAUS scores were similar in the exposed and control groups  
188 ( $2.61$  vs  $2.7$  respectively,  $P=0.83$ ) (Table 2).

189  
190  
191 Quantitative assessment  
192 The final post-course OSAUS score and the global delta for progression did not differ significantly between the  
193 SBE and control groups (Table 2). In comparison with the estimates from preliminary studies, the initial skill level  
194 of the participants was better than expected ( $2.63 \pm 1.01$  versus  $1.25$ ) and the global progression lower ( $0.95 \pm$   
195  $0.92$  versus  $1.75$ ).

196  
197

198  
199 Qualitative assessment  
200 The METHOD Logbook platform was used by 51 (82%) participants who generated LC-CUSUM, a percentage that  
201 did not differ between the SBE (79%) and control (87%) groups (Table 3). The proportion of participants able to  
202 reach the expected skill level (five consecutive validated examinations) was significantly higher in the 'early'  
203 group (74% vs 30% in the control group,  $P < 0.05$ ).

204 The mean number of examinations necessary to reach this skill level ranged from  $8.2 \pm 3.5$  to  $13.5 \pm 4.8$   
205 and did not differ significantly between the groups (Table 3).  
206  
207  
208  
209  
210  
211

## 212 **DISCUSSION**

### 213 **Main findings and interpretation of results**

214 At the end of a long observation period, the participants exposed to the simulation workshop did not progress  
215 significantly more than the control group. Nonetheless, the effect of the simulation might have been masked,  
216 especially because the participants' initial level was higher than expected, which left less margin to observe  
217 progression.  
218

219 Moreover, the proportion of participants able to reach the skill level expected was significantly higher in  
220 the subgroup of participants exposed to SBE near the beginning of the course. This result is crucial for  
221 understanding the impact of SBE on clinical practice. Standardizing the skill level guarantees a minimum  
222 threshold level of practice and is thus a major step toward improving the global quality of care. The number of  
223 practitioners able to reach this threshold skill level thus constitutes a more useful outcome measure than the  
224 mean or individual progression for assessing the potential clinical impact of a training activity.

225 These differences observed between the different timing patterns suggest the existence of different  
226 learning processes. This is probably explained by the optimized use of SBE by the participants who faced these  
227 clinical situations around the same time. This observation raises the question of the relevance of the "never the  
228 first time on the patient" adage, which was initially an argument for the development of simulation in the  
229 training of health professionals (HAS report 2012) (30). If this approach remains ethically justified for invasive  
230 procedures, it fails to incorporate the benefit of learning from mistakes, as described by L. Dyre et al (31).  
231 Learning from errors probably helps to optimize SBE, with students using the simulator to resolve problems  
232 encountered in clinical practice. This mechanism appears to optimize on the one hand training on the simulator,  
233 and on the other hand, increase the ability to transfer the skills acquired on the simulator to clinical practice.  
234 Among the skills used for ultrasound, hand-eye coordination and visual-cognitive skills (ability to mentally  
235 represent an ultrasound diagnosis) may comprise the major elements that explain this synergistic effect (16). A  
236 lower effect of learning from mistakes at the end of the course and a short interval between SBE and the final  
237 assessment may explain the less good performance observed in the late-course group.  
238

### 239 **Strengths and limitations**

240 The substantial number of participants combined with the duration of the observation in the METHOD study is  
241 one of its strengths. Only a single other study had a similar sample size (32), and for all of the observations  
242 concerning abdominal ultrasound (33), the mean observation periods were less than a month, compared with  
243 more than five months in ours.

244 This study also represents the first assessment of the effect of the timing of the exposure (by comparing the  
245 different timing of the SBE workshops relative to the course) on the capacity to acquire and maintain the level of  
246 competence required. The simultaneous use of quantitative outcome measures (score) and qualitative measures  
247 (LC-CUSUM), all relevant to clinical practice and not to performance on the simulator, is the second major

248 strength of this study. This methodological choice enabled us to avoid the known bias that results from the  
249 extrapolation to clinical settings of performance observed with a simulator (21).

250 The improvement of the performance of participants exposed to SBE as well as the clinical training  
251 course is due partly to the supplementary effect induced by the addition of training support for the SBE group.  
252 This inequality bias in principle is a limitation attributable to the fact that all studies up to now have simply  
253 assessed the interest of SBE, with no experiments proposing randomization between SBE alone and clinical  
254 practice alone (34). Because learning ultrasound is indissociable from access to clinical practice, these conditions  
255 cannot ethically be reproduced outside of an experimental context; this bias is thus inevitable. It should  
256 nonetheless be noted that the experimental and short-term comparison between SBE alone (i.e. without clinical  
257 practice) and practical training alone confirmed that SBE augments the students' progress (35).

258 The methods and the pedagogical process used in the METHOD study can be implemented easily in other  
259 medical schools and institutions. But the single-center nature of the data collection limits the possibility of  
260 generalizing the results, especially because of possible biases associated with local particularities in education  
261 methods and in the practice of obstetric ultrasound.

262

### 263 **Practical applications**

264 The experience acquired with the METHOD project also provides valuable information about the practical use of  
265 the OSAUS score. At the conclusion of the study, the investigators identified in particular pitfalls related to the  
266 difficulty in quantifying the participants' progress with a delta between two scores: the magnitude of this delta  
267 depends principally on the operator's initial level rather than on his or her weekly progress. This effect is  
268 explained by the nonlinear nature of the progress (more rapid and marked in the early learning phase). Given the  
269 intrinsic difficulty of obstetric ultrasound, progression later on appears slower and more difficult to quantify, and  
270 it is characterized by a lower rise in the score. This characteristic explains the difficulty in showing a difference in  
271 progression in the METHOD study due to the participants' initial skill levels, which were higher than anticipated.  
272 In the future, this experience will lead us to limit the use of changes in OSAUS scores to the observation of  
273 beginning operators and the comparison of groups with comparable initial levels and similar observation periods.

274 The continuous massive expansion of simulation in training health professionals calls for the  
275 development of scientific tools making it possible to reinforce the practical applicability of experimental results  
276 and the relevance of the studies. The publication of guidelines for Health Care Simulation Research (36)  
277 constitutes a first step, but the question of the cost-benefit relation for SBE must be explored in greater depth  
278 (37,38). In the light of the results of the METHOD study, this investigation must consider the timing of the SBE. At  
279 the same time, the systematization of the outcome measures used to assess the impact of simulation should  
280 make it possible to optimize comparisons between the studies as well as the work of meta-analysis. This  
281 procedure must be based on a Delphi procedure (39) to define a set of homogeneous outcomes enabling the  
282 assessment and comparison of procedures involving SBE. The possibility of concrete application of these results  
283 to our educational practices depends on these final two stages.

284

285

286

287

288

289

290

291

292

293

294

### **Conclusion:**

The assessment of the learning phase by the METHOD OSAUS score does not allow us to conclude that  
there is a difference in the mean progression between participants exposed to SBE and the control group. On the  
other hand, the association of SBE with clinical practice at the onset of the learning phase enabled a significantly  
higher number of participants to reach the required skill level, compared with the control group.

### **ACKNOWLEDGMENT**

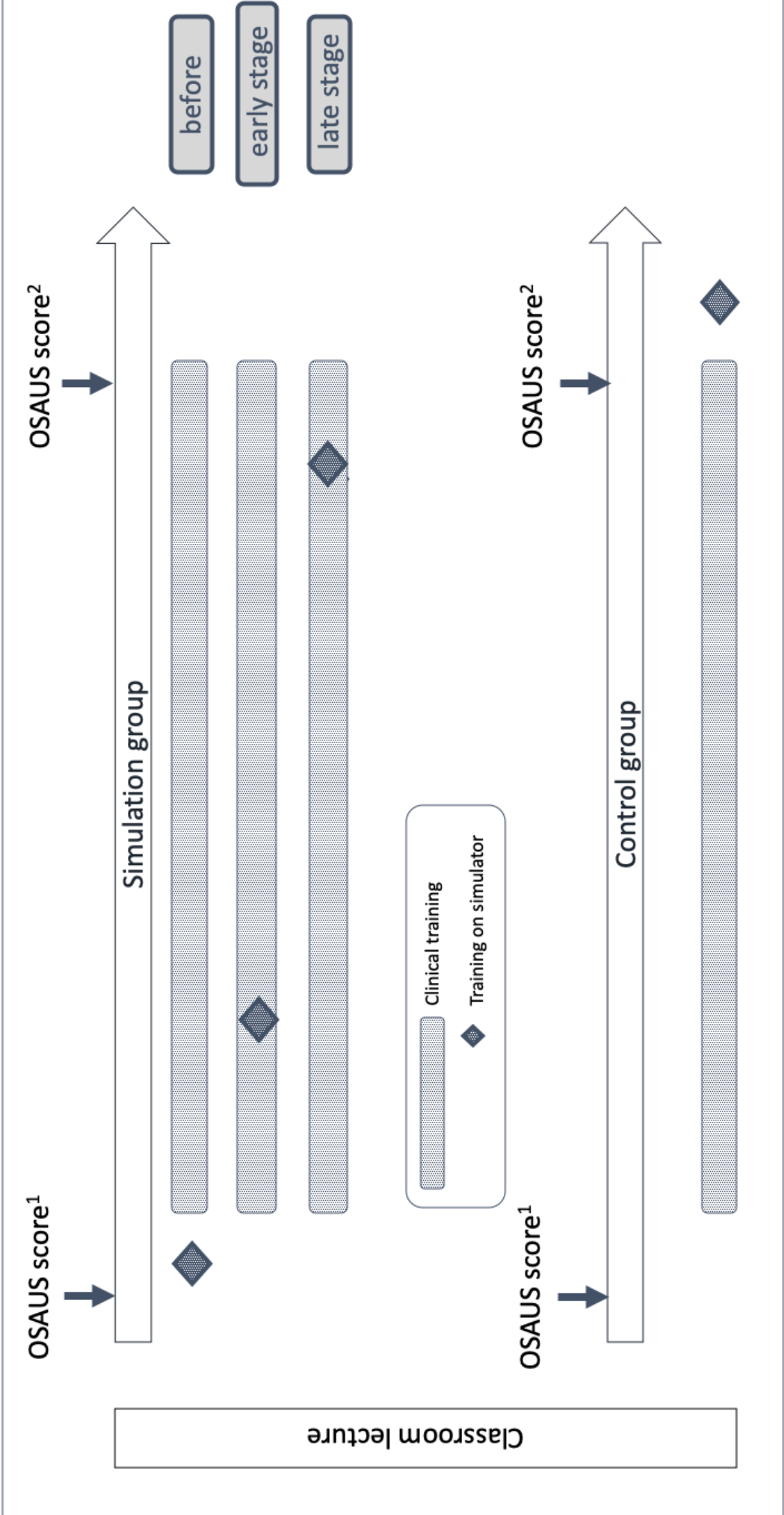
The authors would like to thank the investigators, coordinators, participants and patients for their involvement in  
the study. We also thank Jo Ann Cahn for ensuring the manuscript final translation and proofreading.

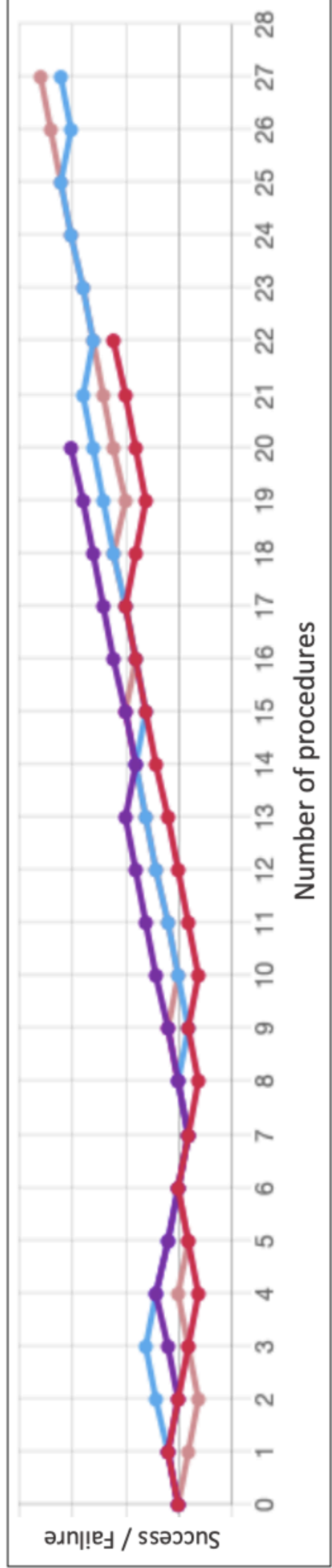
- 295  
296  
297  
298 1. Jenssen C, Gilja OH, Serra AL, Piscaglia F, Dietrich CF, Rudd L, Sidhu P. European Federation of Societies  
299 for Ultrasound in Medicine and Biology (EFSUMB) Policy Document Development Strategy - Clinical Practice  
300 Guidelines, Position Statements and Technological Reviews. *Ultrasound Int Open* 2019; **5**: E2-10.  
301 2. American Institute of Ultrasound in Medicine. AIUM Practice Guideline for the Performance of Obstetric  
302 Ultrasound Examinations. *Journal of Ultrasound in Medicine* 2013; **32**: 1083-101.  
303 3. Moore CL, Copel JA. Point-of-Care Ultrasonography. *New England Journal of Medicine* 2011; **364**:  
304 749-57.  
305 4. Massoud M, Duyme M, Fontanges M, Combourieu D. Courbe d'estimation de poids foetal 2014 par le  
306 Collège français d'échographie foetale (CFEF). *J Gynecol Obstet Biol Reprod* 2016; **45**: 80-5.  
307 5. Papageorgiou AT, Ohuma EO, Altman DG, Todros T, Ismail LC, Lambert A. International standards for  
308 fetal growth based on serial ultrasound measurements: the Fetal Growth Longitudinal Study of the intergrowth-  
309 21st Project. *Lancet* 2014; **384**: 869-79.  
310 6. Salomon LJ, Alfirevic Z, Berghella V, Bilardo C, Hernandez-Andrade E, Johnsen SL, et al. Practice  
311 guidelines for performance of the routine mid-trimester fetal ultrasound scan. *Ultrasound Obstet Gynecol* 2011;  
312 **37**: 116-26.  
313 7. Henriksen T. The macrosomic fetus: a challenge in current obstetrics. *Acta Obstet Gynecol Scand* 2008;  
314 **87**: 134-45.  
315 8. Boulvain M, Senat M-V, Perrotin F, Winer N, Beucher G, Subtil D, et al. Induction of labour versus  
316 expectant management for large-for-date fetuses: a randomised controlled trial. *Lancet* 2015; **385**: 2600-5.  
317 9. Chien LY, Whyte R, Aziz K, Thiessen P, Matthew D, Lee SK, et al. Improved outcome of preterm infants  
318 when delivered in tertiary care centers. *Obstet Gynecol* 2001; **98**: 247-52.  
319 10. Benacerraf BR, Minton KK, Benson CB, Bromley BS, Coley BD, Doubilet PM, et al. Proceedings: Beyond  
320 Ultrasound First Forum on Improving the Quality of Ultrasound Imaging in Obstetrics and Gynecology. *J*  
321 *Ultrasound Med* 2018; **37**: 7-18.  
322 11. Sarris I, Ioannou C, Chamberlain P, Ohuma E, Roseman F, Hoch L, et al. Intra- and interobserver  
323 variability in fetal ultrasound measurements. *Ultrasound Obstet Gynecol* 2012; **39**: 266-73.  
324 12. Hur H, Kim YH, Cho HY, Park YW, Won H-S, Lee M-Y, et al. Feasibility of three-dimensional  
325 reconstruction and automated measurement of fetal long bones using 5D Long Bone. *Obstet Gynecol Sci* 2015;  
326 **58**: 268-76.  
327 13. Ambroise Grandjean G, Hossu G, Banasiak C, Ciofolo-Veit C, Raynaud C, Rouet L, et al. Optimization of  
328 Fetal Biometry With 3D Ultrasound and Image Recognition (EPICEA): protocol for a prospective cross-sectional  
329 study. *BMJ Open* 2019; **9**: e031777.  
330 14. Ambroise Grandjean G, Hossu G, Bertholdt C, Noble P, Morel O, Grangé G. Artificial intelligence  
331 assistance for fetal head biometry: Assessment of automated measurement software. *Diagn Interv Imaging*  
332 2018; **99**: 709-16.  
333 15. Sarris I, Ioannou C, Dighe M, Mitidieri A, Oberto M, Qingqing W, et al. Standardization of fetal  
334 ultrasound biometry measurements: improving the quality and consistency of measurements. *Ultrasound Obstet*  
335 *Gynecol* 2011; **38**: 681-7.  
336 16. Tolsgaard MG. A multiple-perspective approach for the assessment and learning of ultrasound skills.  
337 *Perspect Med Educ* 2018; **7**: 211-3.  
338 17. Chalouhi GE, Bernardi V, Ville Y. Ultrasound simulators in obstetrics and gynecology: state of the art.  
339 *Ultrasound Obstet Gynecol* 2015; **46**: 255-63.  
340 18. Le Lous M, Tsatsaris V, Tesnière A, Grangé G. Improving students' ability to perform a standardized fetal  
341 biometry plane using ultrasound simulators. *J Gynecol Obstet Hum Reprod* 2017; **46**: 439-443



- 342 19. Gueneuc A, De Garnier J, Dommergues M, Rivière M, Ville Y, Chalouhi GE. Impact de l'intégration de la  
343 simulation en échographie obstétricale dans le cursus des étudiants sages-femmes. *Gynécologie Obstétrique*  
344 *Fertilité & Sénologie* 2019; **47**: 776-782
- 345 20. Madsen ME, Konge L, Nørgaard LN, Tabor A, Ringsted C, Klemmensen AK, et al. Assessment of  
346 performance measures and learning curves for use of a virtual-reality ultrasound simulator in transvaginal  
347 ultrasound examination. *Ultrasound Obstet Gynecol* 2014; **44**: 693-9.
- 348 21. Tolsgaard MG, Madsen ME, Ringsted C, Oxlund BS, Oldenburg A, Sorensen JL, et al. The effect of dyad  
349 versus individual simulation-based ultrasound training on skills transfer. *Med Educ* 2015; **49**: 286-95.
- 350 22. Kirkpatrick D. Great Ideas Revisited. Techniques for Evaluating Training Programs. Revisiting Kirkpatrick's  
351 Four-Level Model. *Training and Development*. 1996; **50**: 54-9.
- 352 23. Tolsgaard MG, Tabor A, Madsen ME, Wulff CB, Dyre L, Ringsted C, et al. Linking quality of care and  
353 training costs: cost-effectiveness in health professions education. *Med Educ* 2015; **49**: 1263-71.
- 354 24. Enriquez JL, Wu TS. An introduction to ultrasound equipment and knobology. *Crit Care Clin*. janv 2014;  
355 **30**:25-45, v.
- 356 25. G. Ambroise Grandjean, P.Gabriel, G. Hossu, S. Zuily, O. Morel, P. Berveiller. Training in fetal ultrasound  
357 biometry : prospective assesment of Objective Structured Assessment of Ultrasound Skills (OSAUS) efficiency.  
358 *submitted to Journal of Gynecology Obstetrics and Human Reproduction* 2020.
- 359 26. Biau DJ, Williams SM, Schlup MM, Nizard RS, Porcher R. Quantitative and individualized assessment of  
360 the learning curve using LC-CUSUM. *Br J Surg* 2008; **95**: 925-9.
- 361 27. Biau DJ, Porcher R, Salomon LJ. CUSUM: a tool for ongoing assessment of performance. *Ultrasound*  
362 *Obstet Gynecol* 2008; **31**: 252-5.
- 363 28. Weerasinghe S, Mirghani H, Revel A, Abu-Zidan FM. Cumulative sum (CUSUM) analysis in the  
364 assessment of trainee competence in fetal biometry measurement. *Ultrasound Obstet Gynecol* 2006; **28**:  
365 199-203.
- 366 39. Tolsgaard MG, Ringsted C, Dreisler E, Klemmensen A, Loft A, Sorensen JL, et al. Reliable and valid  
367 assesment of ultrasound operator competence in obstetrics and gynecology. *Ultrasound Obstet Gynecol* 2014;  
368 **43**: 437-43.
- 369 30. Haute Autorité de Santé. Simulation en santé 2012. [Disponible en ligne: [https://www.has-  
370 sante.fr/jcms/c\\_930641/fr/simulation-en-sante](https://www.has-sante.fr/jcms/c_930641/fr/simulation-en-sante)]. Cité le 1 juin 2020.
- 371 31. Dyre L, Tabor A, Ringsted C, Tolsgaard MG. Imperfect practice makes perfect: error management  
372 training improves transfer of learning. *Medical Education* 2017; **51**: 196-206.
- 373 32. Andreasen LA, Tabor A, Nørgaard LN, Ringsted C, Sandager P, Rosthøj S, et al. Multicenter randomized  
374 trial exploring effects of simulation-based ultrasound training on obstetricians' diagnostic accuracy: value for  
375 experienced operators. *Ultrasound Obstet Gynecol* 2020; **55**: 523-9. 34.
- 376 33. Le Lous M, Tsatsaris V, Tesnière A, Grangé G. Improving students' ability to perform a standardized  
377 foetal biometry plane using ultrasound simulators. *J Gynecol Obstet Hum Reprod* 2017; **46**:439-443.
- 378 34. Chalouhi GE. Re: Sustained effect of simulation-based ultrasound training on clinical performance: a  
379 randomized trial. *Ultrasound Obstet Gynecol* 2015; **46**: 264-5.
- 380 35. Chao C, Chalouhi GE, Bouhanna P, Ville Y, Dommergues M. Randomized Clinical Trial of Virtual Reality  
381 Simulation Training for Transvaginal Gynecologic Ultrasound Skills. *J Ultrasound Med* 2015; **34**: 1663-7.
- 382 36. Cheng A, Kessler D, Mackinnon R, Chang TP, Nadkarni VM, Hunt EA, et al. Reporting Guidelines for  
383 Health Care Simulation Research: Extensions to the CONSORT and STROBE Statements. *Simul Health* 2016; **11**:  
384 238-48.
- 385 37. Tolsgaard MG, Tabor A, Madsen ME, Wulff CB, Dyre L, Ringsted C, et al. Linking quality of care and  
386 training costs: cost-effectiveness in health professions education. *Med Educ*. 2015; 49: 1263-71.
- 387 38. Osborne B, Parange N, Thoirs K. The effectiveness of the use of high fidelity simulators in obstetric  
388 ultrasound training: A systematic review. *Australas J Ultrasound Med* 2015; **18**: 107-11.
- 389 39. de Villiers MR, de Villiers PJT, Kent AP. The Delphi technique in health sciences education research. *Med*  
390 *Teach* 2005; **27**: 639-43.
- 391

392	
393	Tables & Figures :
394	
395	Table 1 : Demographic characteristics of the trainees cohort.
396	
397	Table 2: Allocation and results of pre- and post-course assessment in the different groups
398	
399	Table 3: Logbook utilization rates and LC-CUSUM (learning curve cumulative sum) analysis
400	
401	Figure 1: Presentation of the different timing of the METHOD cohort participants' simulation exposure and of
402	their assessments.
403	
404	Figure 2: Example of four LC-CUSUM (Learning Curve Using Cumulative Sum control charts) obtained at the end
405	of the course for the measurement of the estimated fetal weight (each curve represents the individual
406	progression of one participant).
407	
408	Figure 3: CONSORT3 flow diagram of enrollments, allocations, follow-up and analysis for METHOD study.
409	





Trainees assessed for eligibility (n=63)

Excluded (n=1)

- Did not attend pre-course lecture and pre-OSAUS<sup>1</sup> assessment

Trainees eligible for random allocation (n=62)

Allocated to "Simulation" group (n=39)

Allocated to "Control" group (n=23)

Excluded (n=1)  
• Leave of absence

Lost to follow-up :

- quantitative assessment<sup>2</sup> (n=0)
- qualitative assessment<sup>3</sup> (n=8)

Lost to follow-up :

- quantitative assessment<sup>2</sup> (n=1)
- qualitative assessment<sup>3</sup> (n=4)

Analyzed :

- quantitative assessment (n=39)
  - qualitative assessment (n=31)
- Excluded from analysis (n=0)

Analyzed :

- quantitative assessment (n=22)
  - qualitative assessment (n=19)
- Excluded from analysis (n=1)

Table 1 : Demographic characteristics of the trainee cohort

Training curriculum	Mean age <sup>1</sup>	Mean year of medical education <sup>1</sup>	Gender <sup>2</sup>	
			Female	Male
Midwife trainees (n=25)	26.2 ± 0,6	4 ± 0	22	1
Residents (n=36)	29.9 ± 2.1	8.8 ± 1.9	29	7

<sup>1</sup> Results are presented as mean ± standard deviation

<sup>2</sup> Results are presented as absolute numbers

Table 2: Allocation and results of pre- and post-course assessment in the different groups

	Total (n=61)	Control group (n=22)	SBE* group (n=39)	P-value
Pre-OSAUS1	2.63 ± 1.01 [0.4; 4.8]	2.7 ± 1.1 [0.4; 4.2]	2.61 ± 0.97 [0.4; 4.8]	0.83
Post-OSAUS2	3.58 ± 0.72 [1.8; 4.8]	3.38 ± 0.69 [1.8; 4.8]	3.7 ± 0.72 [1.8; 4.8]	0.1
Delta3	0.95 ± 0.92 [-1; 3.2]	0.71 ± 0.98 [-1; 2.6]	1.09 ± 0.72 [-1; 3.2]	0.41

Results are reported as mean ±SD (range).

<sup>1</sup>OSAUS, pre-course assessment via OSAUS METHOD score (Objective Assessment of Ultrasound Skills – Measurement Teaching in Obstetrics Design)

<sup>2</sup>Post-OSAUS, post-course assessment via OSAUS METHOD score (Objective Assessment of Ultrasound Skills – Measurement Teaching in Obstetrics Design)

<sup>3</sup> Difference between post-course and pre-course OSAUS scores

\*Simulation based education (in ultrasound)

Table 3: Logbook utilization rates and LC-CUSUM (learning curve cumulative sum) analysis

	Control group (n=21)	SBE* group (n=39)	'Before' (n=8)	'Early' stage (n=22)	'Late' stage (n=9)
Logbook users <sup>1</sup>	20 (95)	31 (79)	4 (50)	19 (86)	8 (89)
Effective skill level reached <sup>2</sup>	6 (30)	17 (55)	2 (50)	14 (74)	1 (13)
Numbers of exams required to reach the skill <sup>3</sup>	10.7 ± 4.5 [5; 16]	9.5 ± 4.3 [5; 18]	11 ± 2.8 [9; 13]	8.6 ± 4 [5; 18]	18 /

\*Simulation based education

<sup>1</sup> Results are presented as absolute numbers (percentages)

<sup>2</sup> Effective skill level is considered as reached when 5 consecutive correct exams were performed (correct exam corresponding to a difference less than 0.8 Z-score between trainee and supervisor for estimated fetal weight). Results are presented in absolute numbers (percentage)

<sup>3</sup> Results are presented as means ±SD [range].