Original Article

Hysteroscopic Sterilization Using a Virtual Reality Simulator: Assessment of Learning Curve

Julienne A. Janse, MD*, Ruben S. A. Goedegebuure, MD, Sebastiaan Veersema, MD, Frank J. M. Broekmans, MD, PhD, and Henk W. R. Schreuder, MD, PhD

From the Department of Gynaecology and Obstetrics, Sint Antonius Ziekenhuis, Nieuwegein (Drs. Janse and Veersema), and Division of Woman and Baby, Department of Reproductive Medicine and Gynaecology, University Medical Center Utrecht, Utrecht (Drs. Goedegebuure, Broekmans, and Schreuder), the Netherlands.

ABSTRACT

Study Objective: To assess the learning curve using a virtual reality simulator for hysteroscopic sterilization with the Essure method.

Design: Prospective multicenter study (Canadian Task Force classification II-2).

Setting: University and teaching hospital in the Netherlands.

Participants: Thirty novices (medical students) and five experts (gynecologists who had performed >150 Essure sterilization procedures).

Interventions: All participants performed nine repetitions of bilateral Essure placement on the simulator. Novices returned after 2 weeks and performed a second series of five repetitions to assess retention of skills. Structured observations on performance using the Global Rating Scale and parameters derived from the simulator provided measurements for analysis.

Measurements and Main Results: The learning curve is represented by improvement per procedure. Two-way repeated-measures analysis of variance was used to analyze learning curves. Effect size (ES) was calculated to express the practical significance of the results (ES ≥ 0.50 indicates a large learning effect). For all parameters, significant improvements were found in novice performance within nine repetitions. Large learning effects were established for six of eight parameters (p < .001; ES, 0.50–0.96). Novices approached expert level within 9 to 14 repetitions.

Conclusion: The learning curve established in this study endorses future implementation of the simulator in curricula on hysteroscopic skill acquisition for clinicians who are interested in learning this sterilization technique. Journal of Minimally Invasive Gynecology (2013) ••• – ••• © 2013 AAGL. All rights reserved.

Keywords: Essure; Hysteroscopy; Sterilization; Training; Virtual reality

DISCUSS

You can discuss this article with its authors and with other AAGL members at http://www.AAGL.org/jmig-20-6-JMIG-D-13-00141R1

Drs. Janse and Goedegebuure contributed equally to this article.
Dr. Veersema has received honoraria from Conceptus for training sessions on the Essure device and was involved in development of the virtual images for the EssureSim.

The authors have no commercial, proprietary, or financial interest in the products or companies described in this article.


Corresponding author: J.A. Janse, MD, Department of Gynaecology and Obstetrics, Sint Antonius Ziekenhuis, Koekoekslaan 1, Nieuwegein 3430 EM, the Netherlands.
E-mail: Julienne.Janse@gmail.com

Submitted February 27, 2013. Accepted for publication April 22, 2013.
Available at www.sciencedirect.com and www.jmig.org

1553-4650/S - see front matter © 2013 AAGL. All rights reserved.
http://dx.doi.org/10.1016/j.jmig.2013.04.016
Virtual reality trainers have emerged and have been successfully implemented for a wide variety of surgical training purposes including laparoscopy, robot-assisted surgery, and cystoscopy [1–3]. In gynecology, virtual reality simulators have been used primarily for laparoscopic and robotic skills training [4–6]. Hysteroscopic skills training also can benefit from use of virtual reality simulators.

The effectiveness of training using simulators and other models largely depends on the validity of the models and simulators used [7,8]. Validity is defined by whether a simulator enables teaching or evaluation of what is intended to be taught or measured. To implement the model in training curricula, a learning curve provides helpful information and is essential in determining the training capability of the simulator [9]. The learning curve demonstrates the rate of learning and is defined as the relationship between the parameters measured through training repetitions [9].

The HystSim hysteroscopic surgery simulator system (VirtaMed AG, Zurich, Switzerland) is a validated virtual reality simulator for hysteroscopic interventions [10,11]. Recently a new training module (EssureSim; VirtaMed AG, Zurich, Switzerland) has become available on which the Essure sterilization method can be practiced (Conceptus, Inc., Mountain View, CA). As hysteroscopic sterilization using the Essure system is a technique that is becoming widely adopted, a valid training system is urgently needed. As with every new (endoscopic) procedure, new skills are necessary to gain sufficient proficiency and safety. Face and construct validity have recently been established for the EssureSim [12,13]. To investigate the effectiveness of repetitive training using the EssureSim, the objective of the present study was to determine the learning curve of the participants, whether novices can improve their skills, whether they reach a plateau phase, to what extent this plateau phase approximates the plateau phase of the expert participants, and whether there is retention of skills.

The study was exempt from institutional review board approval because there was no potential harm to humans or nonhumans. All participants gave oral consent before the start of the study.

Material and Methods

Participants

From May to June 2011, 30 novices (fourth- to sixth-year medical students at the University Medical Center Utrecht) and 5 experts voluntarily conducted a series of repetitions on the EssureSim device. The novices had never previously performed or assisted in a hysteroscopic procedure but as medical students had a basic understanding of hysteroscopy. All novices were invited to participate via oral and written means, and all agreed. Five gynecologists served as experts to set a reference for novice performance. For the present study, a gynecologist was considered an expert in hysteroscopic sterilization after performing >150 procedures using the Essure system. In the Netherlands, only a limited number of gynecologists were available who met this criterion (from University Medical Center Utrecht, Utrecht; Sint Antonius Ziekenhuis, Nieuwegein; Lucas Andreas Ziekenhuis, Amsterdam; and Ziekenhuis Rivierenland, Tiel). All gynecologists were personally invited, and all agreed to participate. None had any experience performing a simulated procedure using the EssureSim device.

Equipment

The EssureSim consists of an adapted diagnostic hysteroscope, simulation hardware, and software (Fig. 1). The software runs on standard laptop hardware (2.40 GHz Intel Core 2 DUO CPU P8600, 2 GB RAM, NVIDIA Quadro FX 2700 M graphic card). The system does not have haptic feedback. The software contains eight different cases with varying degree of difficulty (Introductory case, Video 1). A case consisting of a uterus with lateral insertion of the fallopian tubes was selected (case 6), and was considered difficult (Video 2).

Measurements

Before the exercises, a survey was administered to obtain baseline characteristics. The novices received a short standardized introduction to hysteroscopy, hysteroscopic sterilization using tubal microinserts, the simulator, and the study protocol. The experts received a standardized explanation of the simulator and study protocol. One investigator (R.S.A.G.) supervised all tests to limit intersupervisor bias. Both groups conducted a series of nine repetitions of the same case. After every repetition, an automatic feedback report was provided by the simulator, and points for improvement were discussed. The first, third, fifth, seventh, and ninth repetitions were used for data analysis. The other repetitions were meant for training by the supervisor and consisted of answering questions from the participants and giving tips and tricks about the procedure. For this reason, the even-numbered repetitions had a different duration and goal and were consequently excluded from further analysis. During the odd-numbered repetitions, no questions could be asked or any tips given because these performances were used solely for data analysis. To assess retention of skills, after 2 weeks, novices returned for a second series of five repetitions. The first, third, and fifth repetitions were used for data analysis. The other repetitions were meant for training by the supervisor and were excluded from further analysis.

Parameters were derived from the simulator software and included time, path length, patient comfort, trauma (cumulative number of cervical and uterine wall contacts), number of correctly placed devices, amount of distention fluid used, time during which the view was obscured, and time during
which the uterus collapsed. Participants were not assessed only on simulator-derived parameters because performing a procedure very fast or efficiently does not necessarily mean it is performed properly and/or with good results. For that reason, a 5-point Global Rating Scale (GRS) was used to assess competence from another (clinical) perspective. The GRS was adjusted for hysteroscopic sterilization (Fig. 2) and has not yet been validated [14]. Aspects that were rated included respect for tissue, handling of the hysteroscope, time and motion, flow and forward planning, and knowledge of the procedure. A description of all parameters used is given in Table 1. Blinding was not possible because of the clear differences between age and status of the groups and the necessity to score both the simulator screen and participant behavior.

**Statistical Analysis**

Data were analyzed using commercially available software (SPSS version 15.0; SPSS, Inc., Chicago, IL). No power analysis was performed before the study. To analyze improvement within the novice group, a sample size of 30 was considered sufficient.

The independent t test and the $\chi^2$ test were used to compare general demographic data of the experts and novices. Two-way repeated-measures analysis of variance was used to analyze learning curves. The between-subject factor group was added to investigate novice and expert performance separately. Retention of skills was investigated by within-subject contrasts and was assessed by comparing the last repetition of both series; significant improvement was defined as a prolonged learning curve [15]. A p value of <.05 was considered statistically significant for all tests. Means and 95% confidence intervals were used to compare data for the learning curves because these are applicable to the analysis of variance.

The practical significance of the results was quantified using the effect size (ES), which is independent of sample size and a scale-free index [16]. Thus, the results can be further interpreted in terms of learning. Significant curves provide information on the evidence of any effect at all, and ES indicates whether a learning effect is meaningful or important. The ES was extracted from the analysis of variance output using SPSS software. ES of 0.10, 0.30, and 0.50 were considered to indicate small (negligible), medium (moderate), and large (crucial) learning effects, respectively. ES was considered relevant only if a significant (p < .05) result was obtained.

**Results**

**Demographic Data**

General demographic data for novices and experts are given in Table 2. As expected, there was a significant difference between sex (p < .05), age (p < .001), and number of tubal microinsert placements performed (p < .001). To assess retention of skills, novices returned after a median of 14 days (range, 7–21 days) for a second series of repetitions. All experts had previous experience using any other type of virtual reality simulator, and 80% of the expert group had
previously practiced for a short time on the hysteroscopy simulator before (without the Essure module).

**Learning Curve**

Results of novice performance in both series of repetitions is given in Table 3 (original measurements). A graphic presentation of the novice learning curve with the expert performance as a reference curve is shown in Figure 3.

Novices showed large improvement in the necessary time to complete a repetition ($p < .01$; ES, 0.77). A large difference was observed between time needed by novices and experts to complete a repetition, in favor of the experts. Novices progressed toward expert level in time, but no plateau phase was observed.

For path length, a similar large improvement in novice performance was observed ($p < .001$; ES, 0.73). A moderate difference between novices and experts was noted, favoring experts. No plateau phase was observed in the novice group.

Novices demonstrated large improvement in patient comfort ($p < .001$; ES, 0.50). During all repetitions, novices outperformed experts, although performance during the last repetition was approximately equal. In the graphic presentation, a plateau phase was recognized at the fifth repetition.

For trauma (cumulative number of contacts with the cervical wall and endometrial cavity), novices showed large improvement ($p < .001$, ES: 0.70), with a similar learning curve by the experts. A plateau phase was reached by the novices at the fifth repetition.

For mean GRS score, novices showed very large improvement ($p < .001$; ES, 0.96). Expert scores were quite stable from the beginning, which contributed to a large between-group difference in favor of the experts. No plateau phase was recognized for the novices.

There was moderate progression in the number of correctly placed microinserts by the novices ($p < .001$; ES,
Experts performed consistently, with zero placements failures. A plateau phase for the novices was recognized at the seventh repetition. For mean time that the uterus collapsed, moderate improvement in novice performance was noted (\( p < .05, \text{ES} = 0.31 \)). For mean time that the view was obscured, large improvement in novice performance was noted (\( p < .001, \text{ES} = 0.51 \)). Data for the expert group was inconsistent for these parameters.

**Retention of Skills**

No significant decrease in performance between the last repetitions of both series was found for the parameters of time (\( p = .06 \)), patient comfort (\( p = .12 \)), trauma (\( p = .06 \)), correctly placed devices (\( p = .33 \)), time that the uterus collapsed (\( p = .45 \)), and time that the view was obscured (\( p = .91 \)), indicating retention of skills by the novices after 2 weeks. A significant increase in performance for path length (\( p < .01; \text{ES} = 0.58 \)) and mean GRS score (\( p < .01; \text{ES} = 0.52 \)) was observed, indicating a prolonged learning curve for both parameters in the novice group (Table 3). Novices approached expert levels after 9 to 14 repetitions.

**Discussion**

The present study assessed the effectiveness of repetitive training for performance using the EssureSim, a virtual reality simulator for hysteroscopic sterilization. For all parameters, significant improvements were found in novice performance within 9 repetitions. Retention of skills was demonstrated for all parameters, and although for some parameters an early plateau was reached, a prolonged learning curve was observed for path length and mean GRS score. Results of the study endorse future implementation of the simulator in curricula on hysteroscopic skill acquisition for clinicians who are interested in learning this sterilization technique. One or more training sessions substantially improve the speed of acquiring sterilization skills on the simulator, which could contribute to consistent, efficient, and safe daily practice. Training sessions include regular feedback from the simulator and, if needed, from a supervisor.

Strong points of this study are its realistic design, additional assessment of retention of skills, and that one investigator supervised all tests in both groups. Furthermore, the use of ES adds information as to whether significant learning curves can be translated into meaningful and important
learning effects. Moreover, use of a clinical parameter (GRS) in addition to the parameters derived from the simulator shows the direct improvement of skills. In addition, the GRS can be of value in future transfer of skills obtained using the simulator to a setting with patients. Although the GRS used has not yet been validated for hysteroscopic sterilization, similar rating scales have been used and validated for use in a general hysteroscopic training program, a HystSim (VirtaMed) myoma resection module and comparable cystoscopic procedures in urology [3,17,18].

The most difficult part of the procedure is placement of the microinserts at the appropriate angle in both tubal ostia. To challenge this aspect and avert measuring only basic hysteroscopic skills, of the eight available cases, a difficult case consisting of a uterus with lateral insertion of the fallopian tubes was chosen. Nonetheless, during the study several experts expressed doubt about the realism of placement angle on the right side. Three experts considered this angle possibly too lateral in comparison with difficult cases in reality. This might have led to a wider variance in results for a number of parameters and to lower scores on patient comfort for experts. That novices outperformed experts on patient comfort could be due to a tendency of novices to operate more carefully under new or difficult circumstances, whereas experts might tend to accept a higher level of patient discomfort to achieve efficient placements. An alternative to the present study design could be one in which hysteroscopic surgeons with no sterilization experience are compared with experts in the hysteroscopic sterilization technique.

Table 3

<table>
<thead>
<tr>
<th>Parameter</th>
<th>First series: learning curve</th>
<th>Second series: retention of skills</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>First repetition</td>
<td>Last repetition</td>
</tr>
<tr>
<td>Time (sec)</td>
<td>513.9 (95% CI, 440.0–587.7)</td>
<td>190.2 (CI, 171.4–208.9)</td>
</tr>
<tr>
<td>Path length (mm)</td>
<td>1710 (95% CI, 1356–2064)</td>
<td>577 (95% CI, 511–642)</td>
</tr>
<tr>
<td>Patient comfort</td>
<td>6.79 (95% CI, 6.36–7.23)</td>
<td>7.82 (95% CI, 7.56–8.08)</td>
</tr>
<tr>
<td>Trauma (No. of contacts)</td>
<td>85.1 (95% CI, 60.9–109.2)</td>
<td>11.8 (95% CI, 7.2–16.3)</td>
</tr>
<tr>
<td>Mean GRS (5-point scale)</td>
<td>1.38 (95% CI, 1.18–1.58)</td>
<td>4.69 (95% CI, 4.58–4.81)</td>
</tr>
<tr>
<td>Devices correctly placed (0–2)</td>
<td>1.53 (95% CI, 1.32–1.75)</td>
<td>1.97 (95% CI, 1.90–2.00)</td>
</tr>
<tr>
<td>Uterus collapsed (sec)</td>
<td>93.5 (95% CI, 37.8–149.2)</td>
<td>20.5 (95% CI, 19.3–39.1)</td>
</tr>
<tr>
<td>View obscured (sec)</td>
<td>6.2 (95% CI, 5.3–7.0)</td>
<td>3.0 (95% CI, 2.2–3.8)</td>
</tr>
</tbody>
</table>

$CI = $confidence interval; ES = effect size (applicable only if result is significant); GRS = Global Rating Scale.

a Indicates significance of comparison between first and last repetitions of first series (analysis of variance).

b Implicates significance of comparison between last repetition of first and second series (analysis of variance).

c Indicates prolonged learning curve.

Considering the improvement in several parameters for both novices and experts, it can be argued whether the first repetition(s) should be used for data analysis in a study assessing a learning curve; the time needed to become familiar with a simulation model could influence results. The chosen study design reflects reality and integrates the possibility of feedback and training in the process [9]. The expert curve functions as a reference, and possible improvement during the first cases is likely to represent reality, in which experts also must become accustomed to the virtual environment [9,19].

Because of lack of a definition of an Essure expert in literature, gynecologists were included who had undeniable experience in hysteroscopic sterilization (had performed >150 Essure procedures), to ensure that the experts were indeed competent. This may have resulted in a small number of experts included in this study. Inasmuch as face and construct validity have been established, the expert group did not need to serve for statistical comparison; this simulator already proved its ability to differentiate between experience levels [12,13]. In the present study, the experts served as a reference to determine at what level they reached a plateau and to exhibit this plateau combined with the learning curves of the novices.

Four of the participating gynecologists already had limited experience on this virtual reality simulator in general but not with the sterilization module. It is possible that this might have resulted in a shorter learning curve for the expert group. Nevertheless, marked improvement in their performance in
The first repetitions was observed. A similar learning curve for experts has been described in other studies, indicating time is needed for anyone to become familiar with a simulation model [9,20].

The distention medium parameter is not reported because of a high number of missing values. This was possibly due to incorrect calibration of inflow and outflow valves on the hysteroscope and the chosen statistical analysis in which cases with missing values were excluded. For the parameters of mean time the uterus collapsed and mean time the view was obscured, missing values were observed in the expert group as well. Therefore, these parameters are not included in the graph of the curves of both groups shown in Figure 3.

In a recently published article establishing construct validity of the EssureSim, a correlation in a multivariate analysis was found between the number of cavity contacts and the
cumulative path length [12]. This finding might imply that improvement of the trauma parameter could be explained in part by shortening of the cumulative path length in general.

The results of the present study indicate a large training capacity of the virtual reality simulator for the hysteroscopic sterilization procedure because large improvements were made by novices training on the EssureSim. Furthermore, retention of skills was established for all parameters, and a prolonged learning curve was observed for path length and the clinical parameter GRS. Therefore, this virtual reality simulator can be recommended for future training purposes. With training using all eight cases provided by the simulator, a shortened learning curve possibly could be expected because of the variety of circumstances.

Assessment of the learning curve is an important step in determining the applicability of a simulator in a training curriculum. Another vital step is investigation of predictive validity, which indicates whether skills gained using a simulator can be transferred to the procedure in patients. In addition, the adjusted GRS for hysteroscopic sterilization could be validated for subsequent use. These are interesting focus areas for future studies.

Acknowledgments

We thank all those who voluntarily participated in the study, and M.J.C. (René) Eijkemans, PhD, for assistance with the statistical analysis.

Supplementary Data

Supplementary data related to this article can be found online at http://dx.doi.org/10.1016/j.jmig.2013.04.016.

References