Does bench model fidelity interfere in the acquisition of suture skills by novice medical students?

RAFAEL DENADAI1, MARIE OSHIIMA2, ROGÉRIO SAAO-HOSSNE3

1 Fellow of Surgery, Division of Coloproctology, Department of Surgery, Botucatu Medical School, Universidade Estadual Paulista (UNESP), Botucatu, SP, Brazil
2 Associate Professor, Department of Statistics, Faculdade de Tecnologia (FATEC), Marilia, SP, Brazil
3 Associate Professor, Division of Coloproctology, Department of Surgery, Botucatu Medical School, Universidade Estadual Paulista (UNESP), Botucatu, SP, Brazil

ARTIGO ORIGINAL

Objective: Although several inanimate bench models have been described for training of suture skills, so far, there is no ideal method for teaching and learning this skill during medical education. The aim was to evaluate whether bench model fidelity interferes in the acquisition of suture skills by novice medical students. Methods: 36 medical students with no surgical skills’ background (novices) were randomized to three groups (n = 12): theoretical suture training alone (control); low-fidelity suture training model (synthetic ethylene-vinyl acetate bench model); or high-fidelity suture training model (pig feet skin bench model). Pre- and post-tests were applied (performance of simple interrupted sutures and subdermal interrupted sutures on ox tongue). Three tools (Global Rating Scale with blinded assessment, effect size, and self-perceived confidence based on Likert scale) were used to measure all suture performances. Results: The post-training analysis showed that the students that practiced on bench models (hands-on training) presented better (all p < 0.0000) performance in the Global Rating Scale evaluation, compared with the control, regardless of the model fidelity. The magnitude of the effect (training) was considered large (> 0.80) in all measurements. Students felt more confident (all p < 0.0000) to perform both types of sutures after training. Conclusion: The acquisition of suture skills on the low-fidelity bench model was similar to that of the high-fidelity bench model, and the increase in the performance of participants that received bench model training was superior to those who received training based on theoretical teaching materials.

Keywords: Medical education; surgery; sutures; teaching; teaching materials.

©2012 Elsevier Editora Ltda. All rights reserved.

Resumo

Será que a fidelidade do modelo de bancada interfere na aquisição das habilidades de sutura por estudantes de medicina iniciantes na prática cirúrgica?

Objetivo: Embora vários modelos de bancada inanimados tenham sido descritos para o treinamento de habilidades de sutura, até o momento, não existe um método ideal para esse ensino e aprendizagem durante a formação médica. O objetivo foi avaliar se a fidelidade dos modelos de bancada interfere na aquisição de habilidades de sutura em estudantes de medicina iniciantes na prática cirúrgica. Métodos: 36 estudantes de medicina sem exposição prévia a habilidades cirúrgicas foram randomizados em três grupos (n = 12): treinamento de suturas baseado em materiais didáticos (controle); treinamento de suturas em modelo de baixa-fidelidade (modelo de bancada de etileno vinil acetato); ou treinamento de suturas em modelo de alta-fidelidade (modelo de bancada de pele de pata de porco). Foram aplicados pré e pós-testes (realização de pontos simples e pontos subdérmicos invertidos em língua de boi). Três ferramentas (Global Rating Scale com avaliação cega, tamanho do efeito e autoavaliação) foram utilizadas para mensurar todas as performances de sutura. Resultados: O análisis após o treinamento demonstrou que os estudantes que treinaram nos modelos tiveram um melhor (p < 0.0000) desempenho na avaliação pela Global Rating Scale, quando comparados com o controle, independente da fidelidade do modelo. A magnitude do efeito (treinamento) foi considerada grande (> 0.80) em todas as mensurações. Após o treinamento os alunos sentiram-se mais confiantes (p < 0.0000) para executarem os dois tipos de suturas. Conclusão: A aquisição de habilidades de suturas no modelo de baixa fidelidade foi semelhante à prática no modelo de alta fidelidade, sendo que a melhora no desempenho dos participantes que treinaram nesses dois modelos foi superior à aprendizagem baseada em materiais didáticos.

Unitermos: Educação médica; cirurgia; suturas; ensino; materiais de ensino.

©2012 Elsevier Editora Ltda. Todos os direitos reservados.
INTRODUCTION

Since suture technique is the foundation upon which most surgical skills will be built\(^1\), and general practitioners are routinely confronted with situations that demand performance of minor surgical procedures (e.g., cutaneous surgery)\(^{2,3}\), the acquisition of this technical skill during medical school (suture technique) is necessary. However, since this foundation is not acquired by a large percentage of students during their education\(^{4,5}\), it is necessary to establish a program of teaching and training of surgical skills during medical education\(^2,7\).

Aiming at this, the simulation-based training has been described\(^8,9\). However, the practice on fresh human cadavers and on live animals is associated with high costs, risks of infections, need for specialized facilities, and legal and ethical aspects\(^8,9\), and the use of virtual reality simulators is hampered by its high cost and lack of access\(^{10}\). As an alternative, inanimate simulators can be used\(^{11-21}\), including parts of postmortem animals (ox tongue, cattle digits, and pig, rat, chicken skins) and synthetic materials such as the ethylene-vinyl acetate (EVA) bench model described recently by this group\(^{20}\), polyurethane foam, and others.

These inanimate simulators vary widely regarding their level of fidelity or “realism” to living human patients\(^9,10\). High-fidelity bench models (e.g., pig, rat, and chicken skins) are limited by high costs, low availability, potential for transmission of infectious diseases, and ethical concerns. Lower-fidelity synthetic bench models (e.g., EVA plates and polyurethane foam) sacrifice “realism” in exchange for portability, lower costs, and potential for repetitive use\(^9,20\).

Despite the fact that the intuitive belief “the more realistic, the better”\(^2,22\) cannot be based on subjectivity alone, and, that there is no ideal model for training suture skills yet\(^2\), few studies have directly compared the effectiveness of synthetic bench models (low-fidelity simulator) versus postmortem animal bench models (high-fidelity simulator) on the acquisition of suture skills during medical education. Therefore, the purpose of this study was to assess objectively whether the fidelity of bench models alters the acquisition of suture skills by novice medical students. Two inanimate simulators made with plates of 4 mm of EVA (low-fidelity) and pig foot skin (high-fidelity) were compared through a randomized, controlled, and blinded study.

METHODS

PARTICIPANTS

The protocol consisted of 36 first- and second-year medical students with no surgical skills background (novices) from a single academic center that volunteered to participate in the study, enrolled after signing an informed consent, in accordance with the Helsinki Declaration of 1975, as amended in 1983. Local institutional research ethics board approval was obtained for this study.

STUDY DESIGN

This was a randomized controlled study with blinded examiners, comprising a pre-test, an one-hour practice phase, and a post-test. The pre- and post-tests were identical and consisted of the performance of five simple interrupted sutures and five subdermal interrupted sutures for closing two elliptical incisions (8 x 2 cm each) on ox tongue. Each student was tested individually and had a total of five minutes for each task. No feedback was provided during the pre- and post-tests.

PRE-TESTING

On the day of the experiment, all participants were taught how to use surgical instruments, as well as the techniques for both types of sutures (simple interrupted sutures and subdermal interrupted sutures), by means of an instructional video\(^{23}\) presentation, which was repeated and discussed six times (verbal teaching based on video). This stage took one hour. Next, all participants underwent a pre-test.

GROUPS AND TRAINING PROGRAM

Immediately after the pre-test, all students were randomized into one of the three study groups (n = 12). The three groups were placed in separate rooms, and were unable to communicate with one another. In group 1 (control), students received faculty-directed training based on theoretical materials (text books and instructional videos) about the handling of surgical instruments and the performance of sutures. In group 2, the students practiced handling surgical instruments and suturing using the EVA bench model (Figure 1) (lower-fidelity model) with the help of instructors (concurrent feedback) according to the training described by the authors\(^{20}\). Students in group 3 received similar training to group 2, but all learning was carried out on the pig foot skin bench model (Figure 2) (high-fidelity simulator) according to some features described by Purimi\(^{24}\). In order to standardize the learning, one faculty instructor was assigned for every four students\(^1\); all instructors were directed to teach suture skills using the same method. This stage lasted one hour\(^1\) for all three groups.

POST TESTING

After the training phase, the instructors demonstrated to all students, in ten minutes, how to perform simple interrupted sutures and subdermal interrupted sutures. Subsequently, all medical students were randomly assigned to post-test.
BLINDING OF THE STUDY

All 144 pre- and post-tests maneuvers were recorded and saved. The digital videos were archived for later analysis and codified using randomly assigned numbers by one of the investigators. All of the recordings were independently and blindly evaluated by two experienced surgical specialists that had no prior knowledge of the groups.

PERFORMANCE MEASURES (QUANTITATIVE AND QUALITATIVE EVALUATIONS)

The quantitative evaluation was based on the number of students that completed one or more stitches and on the number of finalized stitches; only after the ends of the two surgical threads were cut was the stitch considered finalized. The previously validated global rating scale was used to evaluate objectively the suture performance (qualitative assessment) of each student in eight main areas, each of whom was rated on a behaviorally anchored five-point scale where one was the minimum score and five the maximum score, for a total maximum score of 40.

SELF-PERCEIVED CONFIDENCE BASED ON THE LIKERT SCALE

All 36 students completed pre- and post-training questionnaires to measure self-perceived confidence in performing both suture procedures (simple interrupted sutures and subdermal interrupted sutures); they rated their confidence on a five-point Likert scale. The lowest rating (“very unconfident”) was one and the highest rating (“very confident”) was five.

STATISTICAL ANALYSIS

In the descriptive analysis, data were summarized as means, medians, standard deviations, first and third quartiles, and minimum and maximum values. Bioestat® for Windows, version 5.0 was used for the statistical analyses. Student’s t-test was used for measurable variables; Fisher’s exact test was used for the analysis of categorical variables, due to the small sample set. Values were considered significant for a confidence interval of 95% (p < 0.05). Effect sizes were also calculated in order to identify the magnitude of the effect of the intervention regardless the sample size; effect sizes exceeding 0.80 were considered large.

RESULTS

QUANTITATIVE EVALUATION

During the pre-test, none of the 36 participants was able to perform any of the two proposed types of sutures and therefore, there were no differences in the comparative analyses between all three groups (all p > 0.05). In the post-test, a larger number of students (p = 0.0000) of groups 2 and 3 performed a larger number of the two types of sutures (p = 0.0000) when compared with group 1 (Table 1), with no differences in the comparisons between group 2 and 3 (p > 0.05). Comparing the two evaluated periods (pre-test versus post-test), a higher number of students (p = 0.0079) in group 1 completed a larger number of simple interrupted sutures (p = 0.0080) in the post-test, with no difference in the performance
of subdermal interrupted sutures (p > 0.05). In the same comparison (pre-test versus post-test) a higher number of students from groups 2 and 3 performed larger numbers of both types of sutures during post-test (all p = 0.0000).

**Qualitative assessment based on the global rating scale**

In blind evaluations of both pre- and post-tests, no interobserver difference was detected between the examiners on the qualitative evaluation of means in all the three groups (1, 2, and 3) for both types of sutures (all p > 0.05).

In all 72 performances, the means of the qualitative assessments performed in the pre-test period were less than 8.1; therefore, there are no differences between comparisons made group-to-group (all p > 0.05) (Table 1). Qualitative analysis of simple interrupted suture and subdermal interrupted suture performances in the post-test showed that groups 2 and 3 had better performance compared to group 1 (all p = 0.0000). There was no difference (all p > 0.05) in the comparison between groups 2 and 3 for both suture performances (Table 1). Comparing both qualitative measurements (pre-test versus post-test), there was a better performance of all three groups in the performance of simple interrupted sutures (p = 0.0135 for group 1; p = 0.0000 for groups 2 and 3) and subdermal interrupted sutures (p = 0.0019 for group 1; p = 0.0000 for groups 2 and 3) in the post-test (Table 1).

**Effect sizes**

The assessment of the intervention magnitude (training) was considered large (≥ 0.80) in all the measurements made (Table 1).

**Self-perceived confidence**

Regarding the perceptions of students about their confidence to perform both suture techniques, all 36 students were very unconfident (means = 1.0) before training, and therefore, there were no differences in the group-to-group comparison made among all three groups (all p > 0.05) (Table 1). After training, although groups 2 and 3 were similar (p > 0.05), students in groups 2 and 3 felt more confident (p = 0.0000) to perform both types of sutures when compared with group 1. Also after training, comparing the two types of sutures, there were no differences (all p > 0.05) in trust reported by students of all three groups. When comparing pre- and post-training, there was increased confidence (all p < 0.05) in all three groups for the performance of both suture techniques after training (Table 1).

**Discussion**

Since the majority of general practitioners performing minor surgical procedures do not have any formal surgical training, and the ability to close properly a wound is

---

**Table 1 – Quantitative, qualitative (global rating scale), and effect size (training) assessments and students’ perception on their confidence based on Likert scale to perform sutures**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Simple interrupted sutures</th>
<th>Subdermal interrupted sutures</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Group 1</td>
<td>Group 2</td>
</tr>
<tr>
<td>Quantitative evaluation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Students who have stitches (n = 12)&lt;sup&gt;a&lt;/sup&gt;</td>
<td>4 (33.33)</td>
<td>12 (100)</td>
</tr>
<tr>
<td>Total number of stitches (n = 60)&lt;sup&gt;b&lt;/sup&gt;</td>
<td>5 (8.33)</td>
<td>33 (55)</td>
</tr>
<tr>
<td>Number stitches for student (M ± SD)&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.42 ± 0.67</td>
<td>2.75 ± 0.75</td>
</tr>
<tr>
<td>Qualitative assessments (GRS)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-test (M ± SD)&lt;sup&gt;d&lt;/sup&gt;</td>
<td>8.08 ± 0.28</td>
<td>8.04 ± 0.20</td>
</tr>
<tr>
<td>Post-test (M ± SD)&lt;sup&gt;e&lt;/sup&gt;</td>
<td>11.35 ± 4.41</td>
<td>22.58 ± 2.47</td>
</tr>
<tr>
<td>Mean difference</td>
<td>3.27</td>
<td>14.54</td>
</tr>
<tr>
<td>p-value&lt;sup&gt;f&lt;/sup&gt;</td>
<td>0.01</td>
<td>&lt; 0.00</td>
</tr>
<tr>
<td>Effect size&lt;sup&gt;g&lt;/sup&gt;</td>
<td>11.68</td>
<td>72.70</td>
</tr>
<tr>
<td>Students’ perception based on LS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-training (M ± SD)&lt;sup&gt;h&lt;/sup&gt;</td>
<td>1.0 ± 0</td>
<td>1.0 ± 0</td>
</tr>
<tr>
<td>Post-training (M ± SD)&lt;sup&gt;i&lt;/sup&gt;</td>
<td>1.67 ± 0.65</td>
<td>3.17 ± 0.72</td>
</tr>
<tr>
<td>p-value&lt;sup&gt;j&lt;/sup&gt;</td>
<td>0.0023</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

Group 1, control; group 2, ethylene-vinyl acetate bench model; group 3, pig foot skin bench model; M, mean; SD, standard deviation; GRS, Global Rating Scale; LS, Likert scale; <sup>a</sup>p = 0.0000 for all comparisons between all three groups (group 1 < group 2 = group 3), except for group 2 versus group 3 (p = 1.000); <sup>b</sup>p > 0.05 for all comparisons between all three groups (group 1 = group 2 = group 3); <sup>c</sup>p = 0.0000 for all comparisons between all three groups (group 1 < group 2 = group 3), except for group 2 versus group 3 (p > 0.05); <sup>d</sup>p should be considered statistically significant (pre versus post) if p < 0.05; should be considered a large effect if > 0.80.
an essential and important skill in the setting of general practice, and because poor suturing technique and excessive suture tension are directly related to adverse wound healing and increased scarring\textsuperscript{27,28}, it is necessary to teach suturing to both new graduates and undergraduates.

With this purpose (suture training), given that some practical activities in patients may infringe on ethical and legal aspects\textsuperscript{29}, the learning of surgical skills outside the operating room and based on simulation is becoming widely used in medical education\textsuperscript{30-33}. In recent years, numerous inanimate bench simulators (high-fidelity and low-fidelity) that enable suture training have been described\textsuperscript{11-21}, but none are totally complete (or ideal)\textsuperscript{11}. Recently, the authors\textsuperscript{20} and another Mexican group\textsuperscript{31} described the synthetic EVA bench model as a tool for learning suture techniques by medical students and dermatology residents, respectively. However, in these two studies\textsuperscript{20,21}, the gain of suture skills was subjective and did not express the actual level of acquired skills\textsuperscript{35}, as in the training of surgical skills, competence acquisition should be assessed by an objective method\textsuperscript{36}. Thus, in the present study a previously validated certification and quantification tool (global rating scale)\textsuperscript{14,25} was used for the qualitative and objective assessment of suture performances. This evaluation tool is part of the Objective Structured Assessment of Technical Skills\textsuperscript{37}, which is currently considered the gold standard for the objective evaluation of teaching surgical skills\textsuperscript{38}.

In this study, the use of a control group (theoretical training alone) was based on studies\textsuperscript{30,32,33-35} that assessed the acquisition of technical skills, and the training tools (videos and textbooks) used by the theoretical training group are recognized as a form of learning technical skills\textsuperscript{35,37}. The completion of a pre-test after watching an instructional video and the option for training during one hour were based on other studies\textsuperscript{11,35} on suture teaching to medical students. For the blind evaluation of the recorded videos of all 144 performances, the evaluators were able to fast-forward the tapes; a procedure that, besides shortening evaluation time, has been shown to be as effective as showing the entire skill at its natural pace\textsuperscript{39}. Two types of suture were evaluated because the possibility of training with bi- and tri-dimensional sutures has been described as one of the advantages of using the EVA bench model\textsuperscript{20,21} in comparison with other low-fidelity and low-cost synthetic simulators.

To the best of the authors’ knowledge, this is one of the few studies that objectively evaluated the acquisition of suture skills by novice medical students by comparing a low-fidelity bench model with a high-fidelity one. An extensive literature review in English (MEDLINE, Scopus, Web of Science, EBSCO, and Embase databases) retrieved no relevant reports demonstrating the superiority of a high-fidelity model over a low-fidelity model (or the superiority of one low-fidelity model over another low-fidelity model) for teaching suture skills to novice medical students, the target of this study.

In this context, the choice of the pig foot skin bench model (high-fidelity simulator) was based on the idea of investigating a model adopted as a teaching tool in this area\textsuperscript{16}; the choice of the EVA bench model (low-fidelity simulator) was due to the previous experience of the authors\textsuperscript{20} with its use. Furthermore, to compare the two simulators, it is important to consider the differences in the applicability and enforceability of both as teaching tools: the pig foot skin bench model presents some characteristics that differentiate it from the EVA bench model (which is simple, portable, reproducible, versatile, low cost, and has easy accessibility and handling) that may hinder its use, including higher financial costs, material hardness (frozen) that can lead to loss of threaded needles, need for structure, space, and proper conditions for storage, and risk of infection\textsuperscript{16,19}.

As observed in other reports\textsuperscript{25,35}, the students were not able to perform the two proposed tasks before the training program. When comparing the two evaluated periods (pre-test versus post-test), a higher number of students that practiced suturing on EVA bench model (group 2) and on pig skin bench model (group 3) were able to make a larger number of stitches (quantitative) using the appropriate technique (qualitative) when compared with the theoretical training group (control group 1).

In the present study, no differences were observed in the comparison between groups 2 and 3 (hands-on training), demonstrating that the fidelity of the model did not affect the acquisition of suture skills. This was also described in previous studies\textsuperscript{22,33,34}, demonstrating that training on high-fidelity bench models was similar to training on low-fidelity bench models, and both types of models (high-fidelity and low-fidelity) were significantly better than the theoretical training alone. Other reports\textsuperscript{11,13,22,27,28,31,35,39-41} used the performance of sutures on bench models for the investigation of other relevant aspects in surgical education (e.g., optimal instructor: student ratio, and evaluating the need for a faculty mentor to teach basic surgical skills); therefore, objective comparisons between different bench models (high-fidelity versus low-fidelity) were not made.

In this study, the effect sizes (of suture training) were large, indicating that the significant improvement was most likely related to the intervention and not to sample size\textsuperscript{26}. Evaluations of all 144 performances conducted by two independent and blinded evaluators decreased the possibility that this increase in performance was a result of rater bias or expectations from the non-blinded raters\textsuperscript{11}.

This study also showed that, after one hour of training, there was an increase in confidence levels to perform the skills taught (suture techniques), similar to what has been observed in other reports\textsuperscript{12,41}.
In summary, the present results confirm that acquisition of technical skills on bench models or in medical skills laboratories (hands-on training) has greater effectiveness than faculty-directed learning from theoretical materials.

Based on similar results (low-fidelity versus high-fidelity) found in this study, it is believed that in some situations the two bench models (EVA and pig foot skin) can be complementary and, consequently, enhance the already established teaching and learning methods. In training programs that use parts of postmortem animals (e.g., ox tongue, and pig and chicken skins) in classrooms (surgical technique labs), EVA plates (that are accessible, low-cost, and easy to handle) can be adopted as an alternative and complementary tool for training sessions at home. Therefore, students can train under the supervision of instructors in the classroom (immediate feedback) and can also practice at home again and again, bringing the EVA plates to clarify doubts with the instructor (summary feedback). In both moments, important aspects to promote good wound healing and cosmesis (e.g., instrument handling, correct positioning of the needle holder, angle of needle insertion in the “skin”, needle exit at an equidistant point from the insertion point for both bites, closeness of “wound edges” with the appropriate level of tension, evasion and apposition of the “wound edges”, and meticulous “tissue” handling) can be assessed and taught to the student again, promoting a gain of skills over time.

There were some limitations in this study that must be acknowledged. First, the retention of acquired skills was not ascertained. In this context, it has been demonstrated that the retention of surgical skills is stronger when acquired with periods of rest, instead of teaching in a single time; this form of teaching should be adopted in order to retain and improve the learned skills. Second, the transfer of skills to the clinical setting was not investigated. In previous literature on surgical simulator-based training, there is evidence that the surgical skills developed on inanimate bench models can result in performance improvements on corpses, animal models, and in the operating room. A third limitation was that both pre- and post-tests were performed on an inanimate bench model. Since the use of inanimate bench models as a teaching tool alternative to the use of live animals, fresh human corpses, and living human patients is the background of the present study, it is believed that the adoption of any live simulator to assess the acquisition of skills before and after the training phase would be contradictory. A fourth limitation was that the aesthetic outcomes were not directly assessed. In this context, in the global rating scale (adopted as an assessment tool in the present study), there is an area used to measure the quality of the final product, which can partially measure the suture's aesthetic aspects of the “wound” on bench models. Another limitation is the fact that the present study assessed only a basic surgical skill (suturing) and does not meet all the needs of medical students in training, which should include the acquisition of other tasks and surgical skills (e.g., performing excision, dissection, and ligating structures); therefore, care should be taken about the generalization of results to other technical skills.

Further studies are required to measure the retention of skills over time, transferability to a real surgery setting, and the acquisition of suture techniques by trainees in other levels of training (e.g., final year medical students and residents).

**CONCLUSION**

The acquisition of skills of both two- and three-dimensional sutures by novice medical students on the low-fidelity simulator (synthetic EVA bench model) was similar to the practice on the high-fidelity simulator (pig foot skin bench model) after one hour of training, and the qualitative and quantitative improvement of suture performances by participants that trained on both simulators (hands-on training) was greater than those who used theoretical materials.

**REFERENCES**