Experience of a New Dedicated Meta-analysis Tool in Systematic Review and Meta-Analysis on Training Performances of Virtual Reality Bronchoscopy Simulation

Massoud Sokouti1, Ramin Sadeghi1,*, Saeid Pashazadeh2, Saeed Eslami Hasan Abadi3, Mohsen Sokouti4, Morteza Ghojazadeh5, and Babak Sokouti6,⋆

1 Nuclear Medicine Research Center, Mashhad University of Medical Sciences, Mashhad, Iran
2 Faculty of Computer and Electrical Engineering, University of Tabriz, Tabriz, Iran
3 Department of Medical Informatics, Faculty of Medicine, Mashhad University of Medical Sciences, Mashhad, Iran
4 Department of Cardiovascular Surgery, Tabriz University of Medical Sciences, Tabriz, Iran
5 Research Center of Evidence Based Medicine, Tabriz University of Medical Sciences, Tabriz, Iran
6 Biotechnology Research Center, Tabriz University of Medical Sciences, Tabriz, Iran

Corresponding authors: * raminsadeghi1355@yahoo.com; ⋆ b.sokouti@gmail.com

Abstract. The increasing employment of virtual reality bronchoscopy simulation makes it essential to validate the effectiveness of training. This systematic review and meta-analysis performs comparisons among trained groups. Well-known sources (i.e., PubMed, Scopus, and Google Scholar) were systematically searched using a Boolean query based on PRISMA guidelines. The inclusion criteria were peer-reviewed English articles considering virtual reality training simulation based on flexible bronchoscopy for novices and experts, and the studies without enough data or not relevant in content were excluded. Then, the meta-analysis procedure based on random-effects model, heterogeneity and publication bias tests, and subgroup analyses were performed. Eight studies comprising two randomized clinical trials, two prospective, and four observational studies were analyzed. The results for forest plots were illustrated and the significant p values of the target groups were indicative of the fact that the virtual reality bronchoscopy simulation was effective between the groups. The outcomes of meta-regression analyses for sample size and published year showed that the latter one can explain most of the heterogeneity between the studies. And, the funnel plots were indicative of no publication bias. Finally, the high rates of explained variance between studies achieved by subgroup analyses were 89% and 52%. The virtual reality bronchoscopy simulation for training the target groups will be beneficial for both novice trainees and experts. The outcomes of current study showed the effectiveness of virtual reality training technique along with the consistent results in terms of meta-analysis, publication bias, meta-regression, heterogeneity, and subgroup analysis. The stand-alone meta-mums tool is available by requesting through the corresponding authors’ email addresses.

Keywords. Virtual reality; Bronchoscopy; Training; Meta-analysis; Novices; Experts

MSC. 62-07; 90C90

Received: November 17, 2016 Accepted: August 27, 2017

Copyright © 2018 Massoud Sokouti, Ramin Sadeghi, Saeid Pashazadeh, Saeed Eslami Hasan Abadi, Mohsen Sokouti, Morteza Ghojazadeh and Babak Sokouti. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.
1. Introduction

Using the virtual reality environments for getting experts’ hands on bronchoscopy procedure takes time [11]. It needs anatomical experiences to be able to enter the airways from nose to trachea and segmental bronchioles atraumatically at any positions of supine, right and left sides. The novices must be able to identify upper and lower airways anatomy and understand the standard order of segmental airway examinations. For novices and trainees, the most common cause of failure in performing the bronchoscopy procedure is lack of enough training and experiences [3,9,17]. Virtual reality simulation-based training has shown effective trends in endotracheal bronchoscopy. The bronchoscopy training for medical students and fellowships is various in terms of teaching program [12]. A synthesis of data regarding the effectiveness of simulation-based bronchoscopy training would facilitate decisions for education and research requirements and can identify the instructors to use the bronchoscopy for novices, trainees, and even expert students [2,6,10,22]. In the literature, only one systematic review and meta-analysis study was found [13] in which Kennedy’s meta-analysis identified comparative studies of simulation-based bronchoscopy training by including much more variables.

The aim of the current systematic review and meta-analysis research is to investigate the effect of training for performing bronchoscopy procedure by using some simulations and the required times for the trainees to get skills.

2. Materials and Methods

According to PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analysis) guideline [18], articles were reviewed relevant to virtual reality bronchoscopy training with flexible bronchoscopy.

2.1 Search Method

Three open access databases including PubMed, Scopus, and Google Scholar, from 1st January 2000 to 16th May 2016, were systematically searched and reviewed in order to select and evaluate the related published original articles.

The PRISMA flowchart and the search strategy are shown in Figure 1. After thorough investigating and monitoring the extracted articles based on their titles, abstracts, and contents, 2831 articles were extracted and only eight articles were related to virtual reality performed with flexible bronchoscopy.

The eight studies involve virtual training of novices ($N = 90$) and experts or trainees ($n = 65$) by means of simulators for performing bronchoscopy techniques and comparing their states including pre-training and post-training parameters. The required time for performing bronchoscopy procedure before and after training period and obtaining skills using virtual reality based on simulations were measured by CASP checklists [7].
2.2 Data Abstraction

The Boolean expression for searching the databases in the medical subject and headings was used as below:

Query: (xbox OR kinect OR(virtual reality))(bronchoscopy OR bronchoscopic)

Relevant information was extracted from the databases and hence, the downloaded articles including virtual reality bronchoscopy based on simulation training. The required information was gathered separately for the parameters of learning, outcomes, and skills on performing bronchoscopy. Also, skills obtained based on time and process of success were recorded. The parameters such as mean and standard deviations for any types of methodologies and comparisons were included.
2.3 Inclusion and Exclusion Criteria

The inclusion criteria for the extracted articles include the English language, observational and prospective as well as RCT comparative studies.

Exclusion criteria were as follows: (i) unrelated articles, (ii) studies with inadequate data, (iii) non-English articles, case reports or letters to editor, (iv) systematic review, meta-analysis, and duplicated articles. The reported data for outcome measures, comparison groups, and sample sizes of trainees were recorded in a Microsoft Excel Sheet according to the checklist.

Two researchers independently and separately extracted the data from each study. The extracted data were sample size, mean ages, standard deviation, taken times for bronchoscopy training, and any types of comparisons. The required analysis was also performed for exploration of heterogeneity ($I^2$) and publication bias in all groups.

2.4 Characteristics of the Studies

The characteristics of eight included studies are summarized and shown in Table 1. The learners of the studies were mostly novices, trainees, experts (i.e., postgraduate residents, medical students, or other persons such as pulmonologists, anesthesiologists, and thoracic surgeons).

<table>
<thead>
<tr>
<th>Study</th>
<th>Year</th>
<th>Training number</th>
<th>Study design</th>
<th>Instrument</th>
<th>Simulation</th>
<th>Modality</th>
<th>training time</th>
<th>Comparison group</th>
<th>Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>Colt</td>
<td>2001</td>
<td>5 (PG)</td>
<td>Prospective</td>
<td>Flex B</td>
<td>+</td>
<td>VR</td>
<td>4 (hour)</td>
<td>skilled physician</td>
<td>Speed (min)</td>
</tr>
<tr>
<td>Rowe</td>
<td>2002</td>
<td>7 (PG)</td>
<td>Observational</td>
<td>Flex B</td>
<td>+</td>
<td>VR</td>
<td>39 (min)</td>
<td>1st simulator first stimulator</td>
<td>Time &amp; skill</td>
</tr>
<tr>
<td>Moorthy</td>
<td>2003</td>
<td>9 (novice)</td>
<td>Prospective</td>
<td>Flex B</td>
<td>+</td>
<td>VR</td>
<td>-</td>
<td>experienced</td>
<td>Time &amp; skill</td>
</tr>
<tr>
<td>Blum</td>
<td>2004</td>
<td>10 (PG)</td>
<td>RCT</td>
<td>Flex B</td>
<td>+</td>
<td>VR</td>
<td>-</td>
<td>trained &amp; skilled</td>
<td>Time &amp; skill</td>
</tr>
<tr>
<td>Goldmann</td>
<td>2006</td>
<td>15 (PG)</td>
<td>Observational</td>
<td>Flex B</td>
<td>+</td>
<td>VR</td>
<td>4 days</td>
<td>expert &amp; trained novices</td>
<td>Time (FOI)</td>
</tr>
<tr>
<td>Konge</td>
<td>2011</td>
<td>42 (novice)</td>
<td>Observational</td>
<td>Flex B</td>
<td>+</td>
<td>VR</td>
<td>-</td>
<td>14 trainee &amp; 14 experts</td>
<td>Time &amp; skill</td>
</tr>
<tr>
<td>Krogh</td>
<td>2013</td>
<td>20 (MS)</td>
<td>RCT</td>
<td>Flex B</td>
<td>+</td>
<td>VR</td>
<td>71 (min)</td>
<td>self trained &amp; simulator trained</td>
<td>Time</td>
</tr>
<tr>
<td>Latif</td>
<td>2016</td>
<td>15 (PG)</td>
<td>Observational</td>
<td>Flex B</td>
<td>+</td>
<td>VR</td>
<td>90 (min)</td>
<td>post training proficiency</td>
<td>Time &amp; speed</td>
</tr>
</tbody>
</table>

Table 1. Characteristics of included studies

PG = postgraduate, MS = medical student, Flex B = flexible bronchoscopy, VR = virtual reality, FOI = fiberoptic incubation, RCT = randomized clinical trial

All the studies include virtual reality bronchoscopy training, obtaining skills based on simulation, and computer based programs. In one of the studies [10], the simulation training was done on a cadaver and in the remaining ones on live or anesthetized patients were scheduled. Only those studies were selected that only flexible bronchoscopy was included. None of the studies were addressed the advanced techniques such as trans-bronchial biopsy, Endo-bronchial
ultrasound, or rigid bronchoscopy for foreign body extraction, or pediatrics bronchoscopy. In two studies, the flexible bronchoscopy was performed with intubations. For correct performing of the bronchoscopy procedure, the monitor is responsible for displaying videos on the screen.

Among eight studies, two studies were RCT, two other studies were prospective, and four studies were observational. Regarding times taken for training, two studies had four hours and 39 minutes, respectively, and three studies reported over one day, and three other ones did not report their training times.

The outcome measurements include time, speed, and performances of skills accompanying by predicted performance of computer-generated tools.

Specifications of these studies are listed and summarized in Table 1.

2.5 Analysis

Eligible studies were extracted into a computerized spreadsheet for analysis. The meta-mums tool was used to assess the extracted data and the results were in agreement with CMA version 2.2.064. Pooled analysis was done on the studies to calculate effect sizes (Hedge’s g) that were used as effect measure estimates. Meta-analysis was performed with random effect model. Final included published articles were eight articles as mentioned above. The time and speed of bronchoscopy before and after training measures (Mean ± SD) were analyzed.

The heterogeneity and quantitative measures of heterogeneity including Q, p value, and I² were assessed. In the presence of heterogeneity, meta-regression was also performed. A p value of P < 0.05 for the Q test or I² > 50% indicated significant heterogeneity among the studies.

Subgroup analysis of groups A and B were performed based on published years of the studies (before and after 2005) and the heterogeneity could be decreased by subgroup analysis.

Graphical forest plots of effect sizes were generated for all three comparative groups. After generating the funnel plots and performing the required regression modeling such as calculating intercept of Egger’s regression tests and their p values, the publication bias of the studies was assessed. Based on various studies for assessing publication bias p value < 0.05 is regarded as being significant. The statistical analysis performed on all data was performed using both meta-mums tool and CMA version 2.2.0.064 while only the in-house implemented software results are proposed and illustrated in this study to show the meta-mums tool as an alternative software for CMA in future studies of researchers.

3. Results

The results of forest plot of effect size that compares pre-training and post-training of novices (group A) are as followings (Figure 2a):
Hedge’s $g = 1.038$  \( SE = 0.271 \)  \( V = 0.074 \)  \( LL = 0.506 \)  \( UL = 1.57 \)

\[ Z-value = 3.825 \quad p \text{ value} < 0.001 \]

The $p$ value is less than 0.05 and it significantly shows that virtual reality has more efficacies in training the novices.

The heterogeneity test shows high heterogeneity with the following achieved parameters $Q = 23.885$, $df = 8$, $p$ value $= 0.002$, $I^2 = 66.506$, $\Gamma^2 = 0.431$.

\[\text{Figure 2. Forest plots (2a, 2b, 2c)}\]
Systematic Review and Meta-Analysis on Training Performances . . .: M. Sokouti et al.

The results of the meta-analysis for publishing bias, including Egger’s regression test with intercept = 1.845, SE = 3.421, LL = −9.935, UL = 6.244, and \( p \) value = 0.606 does not show any significant publication bias.

The results of forest plots comparing experts and post training experts (group B) are as following (Figure 2b):

Hedge’s \( g \) = 0.829  \( \text{SE} = 0.283 \)  \( V = 0.08 \)  \( \text{LL} = 0.274 \)  \( \text{UL} = 1.383 \)

\( Z \)-value = 2.928  \( p \) value = 0.003

The \( p \) value is significant and it means virtual reality has shown more efficiency in training the experts.

The heterogeneity test also shows high heterogeneity with the parameters as below:

\( Q = 14.681, \ df = 5, \ p \) value = 0.012, \( I^2 = 65.943, \Gamma^2 = 0.313 \).
p value of heterogeneity test is significant, so meta-regression analysis results with sample size are Slope = −0.0169, SE = 0.03, LL = −0.0757, UL = 0.0418, Z-value = −0.565, p value = 0.572, \( \Gamma^2 = 0.31043 \).

p value is not significant and meta-regression with sample size has no relationship with Hedge’s g (Figure 3c). The results of meta-regression analysis with published years are Slope = 0.097, SE = 0.0308, LL = 0.0365, UL = 0.1574, Z-value = 3.1449, p value = 0.002, \( \Gamma^2 = 0.03286 \).

p value is significant and meta-regression with published years has shown direct relationship with Hedge’s g (Figure 3d). In group B, sample size and published year can explain 0 % and 89.5% of heterogeneity, respectively.

Figure 4. Funnel plots (4a, 4b, 4c)
And, the funnel plots are shown in Figure 4b. The result of the meta-analysis for publishing bias, including Egger’s regression test with intercept = 6.283, SE = 4.736, LL = −6.866. UL = 19.432, and p value = 0.255 does not propose any significant publication bias.

The results of forest plots comparing post training novice and experts (group C) are as following (Figure 2c):

Hedge’s g = −0.652  SE = 0.219  V = 0.048  LL = −1.082  UL = −0.223
Z-value = −2.975  p value = 0.003

p value is significant and it means trained novices performed bronchoscopy faster than experts.

The parameters of heterogeneity test are

Q = 6.869, df = 5, p value = 0.231, I² = 27.211, Γ² = 0.078. p value is not significant, so there is no evidence for heterogeneity.

The funnel plots are shown in Figure 4c. The result of the meta-analysis for publishing bias, including Egger’s regression test with intercept = 1.062, SE = 2.681, LL = −6.38, UL = 8.505, and p value = 0.712 and based on p value there is no evidence for existing publication bias.

The results of subgroup analysis of A, B groups before and after 2005 are as following:

In group A; Hedge’s g = 0.471, SE = 0.237, V = 0.056, LL = 0.007, UL = 0.936, p value = 0.047, and Hedge’s g = 1.713, SE = 0.247, V = 0.061, LL = 1.229, UL = 2.197, p value < 0.001. Overall Hedge’s g = 1.067, SE = 0.171, V = 0.029, LL = 0.732, UL = 1.402, p value < 0.001.

In group B; Hedge’s g = 0.173, SE = 0.388, V = 0.15, LL = −0.587, UL = 0.934, p value = 0.655, and Hedge’s g = 1.153, SE = 0.284, V = 0.081, LL = 0.596, UL = 1.709, p value < 0.001. Overall Hedge’s g = 0.811, SE = 0.229, V = 0.052, LL = 0.362, UL = 1.260, p value < 0.001.

In eight studies of group A and six studies of group B, subgroup analysis was performed due to significant heterogeneity (Figure 5, Table 2).

Table 2. Heterogeneity in subgroups of A and B before and after 2005

<table>
<thead>
<tr>
<th>Groups</th>
<th>Years</th>
<th>Q</th>
<th>df</th>
<th>p value</th>
<th>I²</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Before 2005</td>
<td>2.960</td>
<td>4</td>
<td>0.565</td>
<td>0</td>
</tr>
<tr>
<td>A</td>
<td>After 2005</td>
<td>5.444</td>
<td>3</td>
<td>0.142</td>
<td>44.892</td>
</tr>
<tr>
<td>A</td>
<td>within</td>
<td>8.404</td>
<td>7</td>
<td>0.208</td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>Between</td>
<td>15.481</td>
<td>1</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>overall</td>
<td>23.885</td>
<td>8</td>
<td>0.002</td>
<td>66.506</td>
</tr>
<tr>
<td>B</td>
<td>Before 2005</td>
<td>0.292</td>
<td>1</td>
<td>0.589</td>
<td>0</td>
</tr>
<tr>
<td>B</td>
<td>After 2005</td>
<td>7.286</td>
<td>3</td>
<td>0.063</td>
<td>58.828</td>
</tr>
<tr>
<td>B</td>
<td>within</td>
<td>7.579</td>
<td>4</td>
<td>0.108</td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>Between</td>
<td>7.103</td>
<td>1</td>
<td>0.001</td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>overall</td>
<td>14.681</td>
<td>5</td>
<td>0.012</td>
<td>65.943</td>
</tr>
</tbody>
</table>

In figures 5a and 5b, the p values of group A in years before and after 2005 are significant. Due to positive amounts of Hedge’s g, virtual reality training has positive effects on performing bronchoscopy training in both years (before and after 2005).
In group B, before 2005, $p$ value is not significant and hence, the virtual reality training has no influences (Figure 5a). In the years after 2005, $p$ value is significant and virtual reality training has positive effects on performing bronchoscopy training (Figure 5b).

In group A (Table 2), and subgroup analysis before and after 2005 there is no, moderate heterogeneity respectively, hence the subgroup analysis of published year ($Q_{Between} = 13.179$, $df = 1$, $p$ value < 0.001) can explain 89% of variance between studies.

In group B (Table 2), and subgroup analysis before and after 2005 there is no, moderate heterogeneity respectively, hence the subgroup analysis of published year ($Q_{Between} = 4.15$, $df = 1$, $p$ value = 0.042) can explain 52% of variance between studies.

In Table 2 both subgroups of groups A and B have insignificant $p$ values and there is no heterogeneity.

Also, there is no heterogeneity in group C, and therefore, subgroup analysis is not performed ($p$ value = 0.246).

Figure 5. Forest plots of subgroup A, B before 2005 (a) and after 2005 (b)
4. Discussions

The current study is a meta-analysis of information collected from systematic review of literature published articles on virtual reality simulation-based training bronchoscopy which also complies with PRISMA flowchart (Figure 1). The results showed that simulation based training on flexible bronchoscopy are effective in training novices, semi-skilled as well as skilled trainees.

One systematic review along with meta-analysis and two studies demonstrating a brief and full reviews were detected by the detailed description of modalities and effect of instructional design [8,13,21]. However, the two last reviews reported incomplete assessment of the studies based on simulation-based bronchoscopy training. So, they have not been included or compared in the current study. The included studies were selected via literature and their available data were used in the meta-analysis procedure. The results obtained for defining the time are similar to Kennedy’s study [13]. Indeed Kennedy’s study has included the advanced bronchoscopy training based on simulation bronchoscopy such as endobronchial ultrasound, transbronchial needle biopsy, foreign body removal, and rigid bronchoscopy with their specific data. The current meta-analysis assesses the outcome measure of times, speed and skills, and their comparative results before and after training of novices and experts.

Few studies of the literature were focused on scoring of simulation based training bronchoscopy and did not have data about times of performing the procedures. So, these types of studies were not included or compared in the study. Kennedy and colleagues surveyed their study as a systematic review, but their given pooled effect size was lower than that of current study (standard mean difference (SMD) = 0.62, 95% CI, 0.12-1.13, n = 7) versus (standard mean difference (SMD) = 0.688, 95% CI, 0.244-1.132, n = 6) which shows large effect on effect sizes in our study. Given substantial heterogeneity ($I^2 = 55\%$) reported in Kennedy’s study, a low heterogeneity ($I^2 = 27.21\%$) is shown, so the current results can be regarded as significant with higher accuracy than Kennedy’s study. The other studies showed favorable effects on time of training procedures.

Funnel plots of comparing these two systematic reviews were symmetric and effect sizes revealed similar pooled results. Publication bias was not reported in Kennedy’s study whereas there are no evidences for existence of publication biases in the current meta-analysis.

The strength of current study is the inclusion of two prospective [4,19] and two randomized clinical trials studies [1,16]. Results analyzed on wide range of methods of meta-analysis and subgroup analysis also strengthen the outcomes. Although no obvious publication biases were found in all groups, however, groups A and B had high heterogeneity. Sample size meta-regression analysis of group A showed direct relationship with Hedge’s $g$ and could decrease the heterogeneity. Published year meta-regression analysis of group A demonstrated direct relationship with Hedge’s $g$ and also decreased the heterogeneity. Sample size meta-regression analysis of group B did not show direct relationship with Hedge’s $g$. Published year meta-
regression analysis of group B showed direct relationship with Hedge’s $g$ and it could also decrease the heterogeneity.

Subgroup analysis of groups A and B confirmed the results of published year meta-regression analysis. The results of analyses proposed that virtual reality bronchoscopy based on simulation trainings was effective in years after 2005 in both groups.

The weaknesses or limitations of the study are related to the relatively small sample size and absence of test scoring. Moreover, pooled effect sizes are not large enough and are moderate for time outcome.

The meta-analysis study of Kennedy [13] was about the comparison of two groups with and without intervention (virtual training based on simulation) while in this study the meta-analysis compared three groups A, B, and C to strengthen its reliabilities. The results for two meta-analyses showed significant effects favoring simulation for time outcome. Pooled effect sizes are considered in standard mean difference.

The effect size of Kenney’s study (standard mean difference with CI = 95%) was 0.62 (0.12-1.13) compared with that of current study 0.688 (0.244-1.132) which also equals to Hedge’s $g$ in group C. Effect size in the current study has higher effect than that of Kennedy’s study. The outcome of this study suggested that “the longer the structured training procedures, the more effective the performance of training of groups”.

5. Results and Conclusions

In summary, to the best of authors’ knowledge, this systematic review and meta-analysis is performed for the first time to assess the times and speed of learning process of trainees including the medical students, fellows, and experts.

Simulation based training can also accelerate the acquisition of bronchoscopy skills.

However, more RCT bronchoscopic training systematic reviews and meta-analysis are needed to confirm the effectiveness of simulation bronchoscopy in learning and validation of obtained findings on bronchoscopy skills including large groups of trainees.

Acknowledgment

This study as part of a PhD thesis (No. 931507) was approved and supported by Research Council of Mashhad University of Medical Sciences, Mashhad, Iran and carried out in Nuclear Medicine Research Center of Mashhad University of Medical Sciences, Mashhad, Iran.

Competing Interests

The authors declare that they have no competing interests.
Authors’ Contributions
All the authors contributed significantly in writing this article. The authors read and approved the final manuscript.

References


