



FACULDADE DE
MEDICINA
LISBOA

TRABALHO FINAL

MESTRADO INTEGRADO EM MEDICINA

Clínica Universitária de Cardiologia

A simulação realística na formação médica – Revisão sistemática

Hugo Tavares Campos Matias

Orientado por:

Prof.^ª Michele dos Santos Gomes da Rosa

Julho'2021

Resumo

A Simulação é considerada um excelente instrumento na formação médica, possibilitando o contacto com situações realísticas e potencializando a autonomia, trabalho em equipa, manejo de novas tecnologias e a segurança dos doentes. Enquanto ferramenta de ensino, a simulação é um método prático, baseado em problemas, em que o instrutor conduz a aprendizagem dos praticantes, desta forma, com um papel mais ativo.

O objetivo deste trabalho é estabelecer a importância do uso de simuladores na formação académica, essenciais na aprendizagem. Pretende-se definir as circunstâncias da sua utilização, considerando o processo de ensino e a etapa da formação em que o utilizador se encontra, as características de cada instrumento e como avaliar aptidões e conhecimentos objetivamente.

Foi feita uma pesquisa bibliográfica com posterior leitura e análise das fontes encontradas. Os estudos foram organizados por temáticas e os seus resultados colhidos e comparados. Procedeu-se também à avaliação dos mesmos de acordo com as escalas Delphi, PRISMA e GRADE.

Verificou-se que os especialistas tiveram melhores resultados (12%), mas os menos experientes tiveram maiores melhorias. Estes resultados são inconsistentes, com heterogeneidade de 30%, podendo este dado contribuir para que não sejam estatisticamente significantes. Os temas abordados incluem a avaliação de escalas e de instrumentos. A média do score nas escalas Delphi e PRISMA foi de 3,67 em 9 e de 13,97 em 27, respetivamente. Todavia, a maioria dos estudos parece ser forte a favor da recomendação do uso de simuladores segundo a escala GRADE.

Conclui-se que os simuladores e escalas de avaliação parecem ser indicados para formação médica, tanto de alunos (que parecem tomar maior proveito dos mesmos) como de especialistas, que também apresentam melhorias. Apesar de oferecerem prática nas várias áreas da atuação médica, são necessários estudos que investiguem a transferência destas aptidões para um contexto real, complementando o ensino tradicional.

Palavras-chave: Simuladores Educacionais; Usabilidade; Robótica; Simulação Realística; Aprendizagem

Abstract

Simulation is considered an excellent tool in medical education, enabling experiences with realistic situations, and enhancing autonomy, teamwork, handling new technologies and patient safety. As a teaching tool, it is a practical method, based on problems, where the instructor guides the participant, who have a more active role in their own learning experience.

The objective of this project is to establish the importance of simulators in medical education, which are essential in learning. It is aiming to define the circumstances of their usage, considering the teaching process and the stage of education in which participants are in, the tool's features and how to objectively evaluate skills and knowledge.

A research was conducted, followed with reading and analysis of the articles found. The studies were organized according to their themes and their results collected and compared. Every article was the evaluated according to the Delphi, PRISMA and GRADE scales.

Specialists had better results overall (12% higher), but the less experienced participants had bigger improvements. These results are inconsistent, though, with heterogeneity of 30%, which can translate as non-statistically significant results. The addressed themes include the evaluation of scales and tools. The mean score of the Delphi and PRISMA scales was 3,67 out of 9 and 13,97 out of 27, respectively. However, most studies seem to be strongly recommended, according to the GRADE scale.

Concluding, both simulators and evaluation scales seem to be indicated for use in medical education for students (who appear to better take advantage of them) and specialists, who also showed improvement while using these tools. Even though these offer practice in many professional environments, more studies are necessary to investigate skill transference to a real context to support the previous traditional methods.

Keywords: Educational Simulators; Usability; Robotics; Realistic Simulation; Apprenticeship

O Trabalho Final é da exclusiva responsabilidade do seu autor, não cabendo qualquer responsabilidade à FMUL pelos conteúdos nele apresentados.

Índice

RESUMO	1
ABSTRACT	2
ÍNDICE	4
ÍNDICE DE GRÁFICOS E TABELAS	5
AGRADECIMENTOS	6
INTRODUÇÃO	7
MÉTODOS.....	8
RESULTADOS	9
DISCUSSÃO	17
CONCLUSÃO	21
REFERÊNCIAS BIBLIOGRÁFICAS	22
ANEXOS	28

Índice de Gráficos e Tabelas

Gráfico 1.....	10
Gráfico 2.....	12
Gráfico 3.....	14
Gráfico 4.....	15
Gráfico 5.....	16
Tabela 1.....	11
Tabela 2.....	12
Tabela 3.....	14
Tabela 4.....	15
Tabela Suplementar 1.....	27
Tabela Suplementar 2.....	30
Tabela Suplementar 3.....	32

Agradecimentos

Agradeço a todos os que me apoiaram nesta caminhada. À minha mãe, sempre disponível, preocupada e pronta a ajudar, mesmo quando não pode. Ao meu pai, que se esforça sempre demais. À minha namorada, que me apoiou e ajudou durante os últimos momentos da redação deste trabalho. À Zelda, que dá tudo e muito mais. À Professora Michele Gomes da Rosa, pelo seu esforço incansável e disponibilidade imensa para esclarecer todas as questões e apoiar no que fosse necessário. Aos meus amigos, a quem levo para a vida. À restante família, pelo apoio. E, finalmente, aos que já não estão presentes, mas deixaram tudo para que pudesse chegar onde cheguei.

Muito, muito obrigado.

Introdução

Tradicionalmente, a cirurgia tem por base um ensino prático desde os seus primórdios (Custers & Cate, 2018). Todavia, com as crescentes preocupações relacionadas com a segurança dos doentes, a introdução de novas tecnologias, foco no custo-eficácia e redução da carga laboral e dos horários de trabalho (Maertens et al., 2016), o tempo de treino em contexto de bloco operatório tem vindo a ser progressivamente reduzido (Hoogenes et al., 2017). Desta forma, estes fatores acabam por culminar numa redução da qualidade do ensino cirúrgico. Assim, vários simuladores, protocolos de atuação e escalas de avaliação têm sido criados e estudados com o intuito de poder tornar o treino prático mais eficiente, permitindo alcançar os padrões de qualidade exigidos no tempo disponível (Abdelsattar et al., 2015).

É necessário avaliar a competência através da demonstração das suas aptidões. Contudo, estas avaliações têm focado competências baseadas no conhecimento e não em capacidades processuais (Morris et al., 2012). Tradicionalmente, as aptidões são adquiridas por observação e experimentação durante estágios, existindo um maior risco de lesão dos doentes e dos alunos (Guérard-Poirier et al., 2020). A maioria das complicações em bloco operatório ocorrem devido a erro humano, pelo que a experiência do cirurgião está inversamente relacionada à taxa de complicações (Brinkmann et al., 2017). O treino por simulação surgiu assim como um dos métodos de ensino mais eficazes que também conduz a uma redução dos erros, melhorando os cuidados e segurança dos doentes, podendo ser benéfico até para especialistas (Kahol et al., 2009).

Simulação é uma estratégia educacional que se foca em aptidões técnicas, procedimentos e operações, apresentado situações semelhantes à realidade, com o objetivo de melhorar a competência e confiança dos alunos, segurança dos doentes e melhorar a redução do erro (Shen et al., 2017). Os simuladores cirúrgicos são considerados seguros e eficazes para prática e avaliação nas fases iniciais do treino de um cirurgião, tendo instalações reproduzíveis que permitem uma avaliação longitudinal (Oquendo et al., 2018). Este novo contexto simulado passa a ser mais seguro e padronizado, ideal para uma aprendizagem inicial, já que no bloco operatório, o ambiente é mais aleatório e a segurança do doente pode não estar assegurada (Selvander & Åsman, 2013). Contudo, a avaliação objetiva das aptidões técnicas não

deve substituir testes cognitivos e subjetivos já usados (*Training and Assessment of Laparoscopic Skills*, 2004).

Neste sentido, este trabalho vai focar-se em vários métodos de avaliação ou de ensino alternativos através da análise de outros estudos, reunindo inferências numa revisão sistemática dos resultados encontrados. Pretende-se desenvolver as temáticas relacionadas com protocolos de ensino e avaliação, criação, validação e utilização de simuladores, capacidade de transição para um contexto realista e até seleção de avaliadores, análise das diferenças associadas à experiência e criação de modelos de auxílio ao ensino.

Métodos

Esta revisão foi feita com recurso à biblioteca eletrónica PubMed. Foram utilizados os termos de pesquisa “(“Assessment”[Journal] OR “assessment”[All Fields]) AND tools[All Fields] AND (“clinical competence”[MeSH Terms] OR (“clinical”[All Fields] AND “competence”[All Fields]) OR “clinical competence”[All Fields] OR (“clinical”[All Fields] AND “skills”[All Fields]) OR “clinical skills”[All Fields]) AND (“surgical procedures, operative”[MeSH Terms] OR (“surgical”[All Fields] AND “procedures”[All Fields] AND “operative”[All Fields]) OR “operative surgical procedures”[All Fields] OR “surgical”[All Fields] AND skills[All Fields] AND (“students, medical”[MeSH Terms] OR (“students”[All Fields] AND “medical”[All Fields]) OR “medical students”[All Fields] OR (“medical”[All Fields] AND “students”[All Fields]))” e foram selecionados os estudos no período entre 2004 e 15 de janeiro de 2021.

A pesquisa resultou em 38 artigos. Excluíram-se os artigos que não estivessem escritos em português ou inglês e que não façam menção aos resultados ou temática de interesse. De seguida, excluíram-se da análise estatística as revisões sistemáticas e narrativas, protocolos, guidelines, artigos de opinião ou aqueles que, mesmo apresentando resultados, estes não são possíveis de comparar com os restantes estudos.

Os artigos utilizados na amostra desta revisão foram agrupados consoante a temática geral abordada e também o *outcome* avaliado. Para além disso, foram também avaliados quanto aos critérios de qualidade utilizando as escalas Preferred Reporting Items for Systematic Reviews and Meta-Analyses ou PRISMA (Page et al., 2021), Delphi

(Verhagen et al., 1998) e Grading of Recommendations Assessment, Development and Evaluation ou GRADE (Goldet & Howick, 2013). A escala PRISMA é usada na avaliação de revisões sistemáticas ou meta-análises. Já a escala Delphi aplica-se a estudos clínicos. Dado existir vários tipos de estudos, os critérios foram adaptados para poderem responder a estas diferenças. A escala GRADE analisa o nível de recomendação da evidência com base no tipo de estudo, na presença de vieses e no impacto dos resultados.

Reuniram-se os valores de tempo de treino, valores de *performance* e da sua melhoria, separados de acordo com a experiência dos participantes, e parâmetros associados a segurança dos procedimentos. Estes resultados foram convertidos para valores percentuais para maior facilidade de comparação dos mesmos. Procedeu-se então ao cálculo da média da melhoria da *performance* entre os participantes. Esta análise foi feita com recurso ao programa RStudio para Mac.

Resultados

Após a pesquisa com os termos referidos, foram obtidos 38 artigos. Foram excluídos 1 artigo por estar escrito exclusivamente em francês e outros 3 por não abordarem os temas pretendidos: uma revisão sistemática em se excluíam estudos de “ferramentas para uso em contexto de simulação, para avaliar aptidões processuais ou cirúrgicas”, um artigo de opinião, sem variáveis, e um estudo em que se compara uma ferramenta cirúrgica para canhotos com a ferramenta original, para destros. Os artigos excluídos são mencionados na Tabela Suplementar 2. Desta forma, foram escolhidos 34 artigos. De forma a poder efetuar a análise estatística, foram excluídas as 2 revisões sistemáticas (Maertens et al., 2016; Morris et al., 2012), 1 protocolo (Guérard-Poirier et al., 2020) e 1 conjunto de *guidelines* (Beard et al., 2011), totalizando um conjunto de 28. Para além destes, mais 6 artigos não puderam ser utilizados na análise estatística: Meier et al. (2012) e Pellen et al. (2009), cujas escalas de avaliação não são convertíveis a percentagem, não sendo possível comparar os resultados com os restantes estudos, Vernez et al. (2017), em que o grupo experimental não se insere nos grupos que se pretendem avaliar, Hofstad et al. (2017), cujos resultados estão apresentados apenas sob a forma de gráficos, Bajka et al. (2010), cujos resultados estão descritos enquanto comparação entre os vários grupos em estudo (não apresentando valores numéricos

comparáveis) e Emken et al. (2004), que apresenta uma descrição de uma base de dados. Assim, foram 24 os artigos submetidos à análise estatística. O Gráfico 1 representa o *Flowchart* associado ao processo de inclusão e exclusão das publicações.

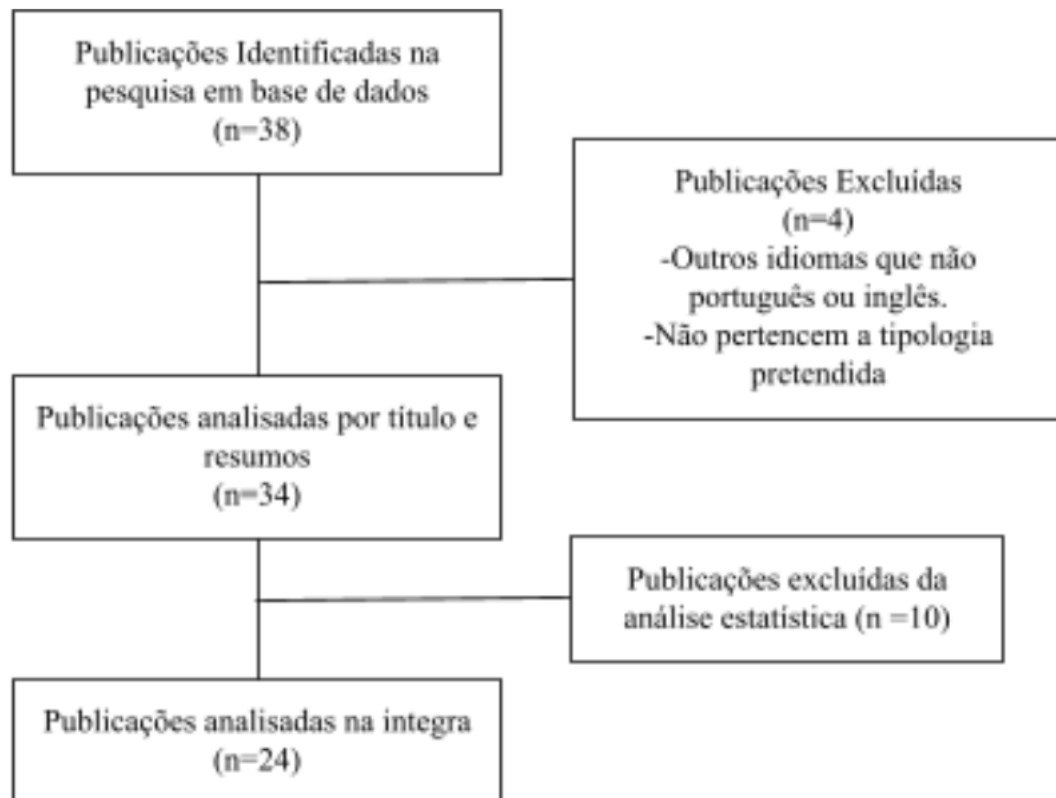


Gráfico 1: *Flowchart* representativo da exclusão e inclusão das publicações no estudo.

As temáticas abordadas nos vários estudos foram agrupadas em: análise de escalas/protocolos de avaliação e análise de instrumento, como representado na Tabela 1, de acordo com os objetivos de cada um. O grupo “análise de escalas/protocolos de avaliação” inclui os artigos que criam e avaliam protocolos ou escalas de avaliação e métodos de ensino, adaptam protocolos pré-existent para línguas diferentes, avaliam examinadores, avaliam a transferência de aptidões para um contexto real e também avaliam os focos e objetivos dos participantes segundo a sua experiência e o grupo “análise de instrumento” refere-se à criação, caracterização, avaliação, comparação e validação de simuladores e dos procedimentos e técnicas associadas e que estes pretendem mimetizar.

Estudo	Autor/Ano	População	Tema	Outcome Principal
Cataract surgeons outperform medical students in Eyesi virtual reality cataract surgery: evidence for construct validity	Selvander & Åsman, 2013	Oftalmologistas e estudantes de medicina	Caraterísticas de simuladores	Avaliação de instrumento
Objective Assessment of training surgical skills using simulated tissue interface with real-time feedback	Rafiq, 2008	Estudantes do 2ª ano de medicina	Avaliação de simulador	Avaliação de instrumento
Knowing the operative game plan: a novel tool for the assessment of surgical procedural knowledge	Balayla, 2012	Estudantes internos e cirurgiões gerais	Avaliação de método de ensino	Avaliação de Escala
Process-oriented instrument and taxonomy for teaching surgical procedures in medical training: the ultrasound-guided insertion of central venous catheter	Galvez, 2020	Cirurgiões	Criação de protocolo	Avaliação de Escala
Assessment of Basic Ankle Arthroscopy Skills in Orthopedic Trainees	Johnson, 2019	Estudantes de medicina e internos 1ªe 5ª	Criação de protocolo	Avaliação de Escala
Evaluation of problem- and simulator-based learning in lumbar puncture in adult neurology residency training	Sun & Qi, 2018	Internos do 1ºano	Adaptação de protocolo para outra língua	Avaliação de Escala
Assessing the validity evidence of an objective structured assessment tool of technical skills for neonatal lumbar punctures	Iyer, 2013	Residentes de pediatria	Avaliação de simulador	Avaliação de instrumento
Virtual reality and computer-enhanced training devices equally improve laparoscopic surgical skill in novices	Kanumuri, 2008	Estudantes 3º ano medicina	Avaliação de simulador	Avaliação de instrumento
Procedural surgical skill assessment in laparoscopic training environments	Uemura, 2016	Estudantes de medicina, internos e cirurgia geral	Avaliação de simulador	Avaliação de instrumento
The e-handover: applications for surgical training	Hand, 2016	Internos de cirurgia geral	Criação de instrumento de avaliação objetiva automática	Avaliação de instrumento
Assessment of lumbar puncture skill in experts and nonexperts using checklists and quantitative tracking of needle trajectories: implications for competency-based medical education	Clinkard, 2015	Estudantes de medicina, internos de anestesiologia e de emergência	Avaliação de ferramentas de avaliação	Avaliação de Escala
Validated Assessment Tools for Pediatric Airway Endoscopy Simulation	Jabbour, 2012	Estudantes, internos e cirurgiões gerais	Avaliação de vídeos	Avaliação de Escala

Automatically rating trainee skill at a pediatric laparoscopic suturing task	Oquendo, 2017	Estudantes de medicina, internos e cirurgiões gerais	Avaliação de protocolo; Comparação de protocolos	Avaliação de Escala
Box- or virtual-reality trainer: which tool results in better transfer of laparoscopic basic skills? - A prospective randomized trial	Brinkmann, 2017	Estudantes de medicina	Avaliação de método de ensino; Avaliação de transferência de aptidões	Avaliação de Escala
Case-based learning and simulation: useful tool to enhance nurses's education? Nonrandomized controlled trial	Raurell-Torredà, 2015	Estudantes 2º de enfermagem	Comparação entre escalas	Avaliação de Escala
Do you see what I see? How we use video as an adjunct to General Surgery Resident Education	Abdelsattar, 2015	Internos 1ª e 5ª de medicina e cirurgia geral	Avaliação de método de ensino	Avaliação de instrumento
A randomized comparison of 2 robotic virtual reality simulators and evaluation of trainees' skills transfer to a simulated robotic urethrovesical anastomosis task	Hoogenes, 2017	Estudantes de medicina, internos e especialistas em otorrinolaringologia	Avaliação de protocolo	Avaliação de Escala
Assessing Wire Navigation Performance in the Operating Room	Taylor, 2016	Internos 3 ano de medicina	Avaliação de simuladores; Comparação de simuladores	Avaliação de instrumento
Reliability and validity study of the Spanish adaptation of the "Creighton Simulation Evaluation Instrument (C-SEI)"	Roldán-Merino, 2019	Estudantes de Enfermagem	Avaliação de ferramenta de avaliação e ensino	Avaliação de instrumento
A Novel Clinical-Simulated Suture Education for Basic Surgical Skill: Suture on the Biological Tissue Fixed on Standardized Patient Evaluated with Objective Structured Assessment of Technical Skill (OSATS) Tools	Shen, 2017	Estudantes de medicina	Avaliação de instrumento de avaliação/protocolo	Avaliação de Escala
Ureteroscopy and cystoscopy training: comparison between transparent and non-transparent simulators	Hu, 2015	Estudantes do 1º ano de medicina	Avaliação de protocolo	Avaliação de escala
Evaluating the use of cleft lip and palate 3D-printed models as a teaching aid	AlAli, 2018	Estudantes de medicina	Avaliação de instrumento de ensino e avaliação	Avaliação de instrumento
Cognitive simulators for medical education and training	Kahol, 2009	Internos do 6ºano medicina	Foco novatos vs especialistas	Avaliação de escala
Faculty feedback versus residents' self-assessment of operative performance: different but complementary	Tanoue, 2018	Internos e especialistas em medicina	Avaliação de modelos de auxílio de ensino	Avaliação de instrumento

Tabela 1: Descrição do tema e *outcome* principal dos vários estudos incluídos na análise estatística.

Cada um dos 24 artigos foi avaliado de acordo com as escalas PRISMA e Delphi, representados no Gráfico 2, e GRADE apresentada na Tabela 2.

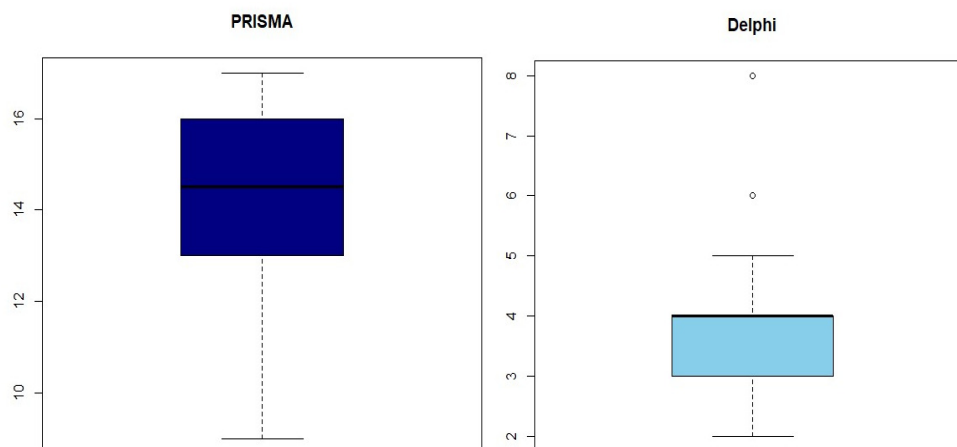


Gráfico 2: Distribuição dos scores atribuídos aos estudos com recurso às escalas PRISMA (azul escuro) e Delphi (azul claro).

Estudo	GRADE
Selvander & Åsman, 2013	Forte a favor de recomendação
Rafiq, 2008	Fraco a favor de recomendação
Balayla, 2012	Forte a favor de recomendação
Galvez, 2020	Fraco a favor de recomendação
Johnson, 2019	Forte a favor de recomendação
Sun & Qi, 2018	Forte a favor de recomendação
Iyer, 2013	Fraco a favor de recomendação
Kanumuri, 2008	Forte a favor de recomendação
Uemura, 2016	Fraco a favor de recomendação
Hand, 2016	Forte a favor de recomendação
Clinkard, 2015	Forte a favor de recomendação
Jabbour, 2012	Fraco a favor de recomendação
Oquendo, 2017	Forte a favor de recomendação
Brinkmann, 2017	Forte a favor de recomendação
Raurell-Torredà, 2015	Forte a favor de recomendação
Abdelsattar, 2015	Fraco a favor de recomendação
Hoogenes, 2017	Forte a favor de recomendação
Taylor, 2016	Forte a favor de recomendação
Roldán-Merino, 2019	Forte a favor de recomendação
Shen, 2017	Fraco a favor de recomendação
Hu, 2015	Forte a favor da recomendação
AlAli, 2018	Forte a favor de recomendação
Kahol, 2009	Forte a favor de recomendação
Tanoue, 2018	Forte a favor de recomendação

Tabela 2: Classificação atribuída aos estudos com recurso à escala GRADE.

PRISMA

A escala PRISMA é uma das escalas que avalia a qualidade dos estudos, variando entre 0 e 27 pontos. O score máximo foi atribuído a 4 dos artigos (Balayla et al., 2012; Galvez et al., 2020; Johnson et al., 2019; Roldán-Merino et al., 2019), no valor de 17 pontos. Pelo contrário, o score mínimo foi de 9, atribuído a 1 artigos (Rafiq et al., 2008). A média foi de aproximadamente 14,29.

Analisando os vários critérios, houve 6 que nunca foram verificados: “identificação do estudo enquanto revisão sistemática ou meta-análise no título”, “indicação de protocolo de revisão, onde aceder ao mesmo e fornecimento informação de registo”, “apresentação da estratégia de pesquisa”, “resultados da seleção de estudos”, “apresentação de dados sobre risco de viés dentro” e “entre os estudos”. Os critérios “descrição do raciocínio no contexto atual” e “apresentação sumária dos resultados por grupo, com efeitos estimados e intervalos de confiança associados” verificaram-se em todos os artigos.

Delphi

A segunda escala usada foi Delphi, cuja pontuação varia entre 0 e 9 pontos. O score máximo foi atribuído a 1 dos artigos (Hoogenes et al., 2017), no valor de 8 pontos. Pelo contrário, o score mínimo foi de 2, atribuído a 5 artigos (Jabbour et al., 2012; Kahol et al., 2009; Selvander & Åsman, 2013; Taylor et al., 2016; Uemura et al., 2016). A média foi de 3,67.

Analisando os vários critérios, houve 1 foi apenas verificado uma vez: “doente cego”, que informa sobre se o doente foi cego à decisão terapêutica. O critério que explicita o uso de estimativas ou medidas de variabilidade enquanto medidas do *outcome* primário, foi o mais atribuído, em 23 dos artigos.

GRADE

Finalmente, foi também utilizada a escala GRADE para uma avaliação qualitativa. Os estudos, representados na Tabela 2 aparentam ser, na sua maioria (N=17), fortes a favor da recomendação. Os restantes (N=7) foram avaliados como fracos a favor da recomendação.

Análise Estatística

A distribuição demográfica dos participantes encontra-se resumida na Tabela 3 e Gráfico 2.

Estatísticas descritivas					
	N	Mínimo	Máximo	Média	Desvio Padrão
Participantes	24	7	249	50,33	54,411
Homens	6	10	33	23,33	9,266
Mulheres	6	6	217	47,67	83,246
Idade	5	22,7	29,0	25,740	2,8527
Estudantes	17	3	249	39,65	61,317
Internos	16	2	80	25,44	22,118
Especialistas	9	3	59	12,44	17,643

Tabela 3: Resumo das características demográficas dos indivíduos inseridos nos estudos.

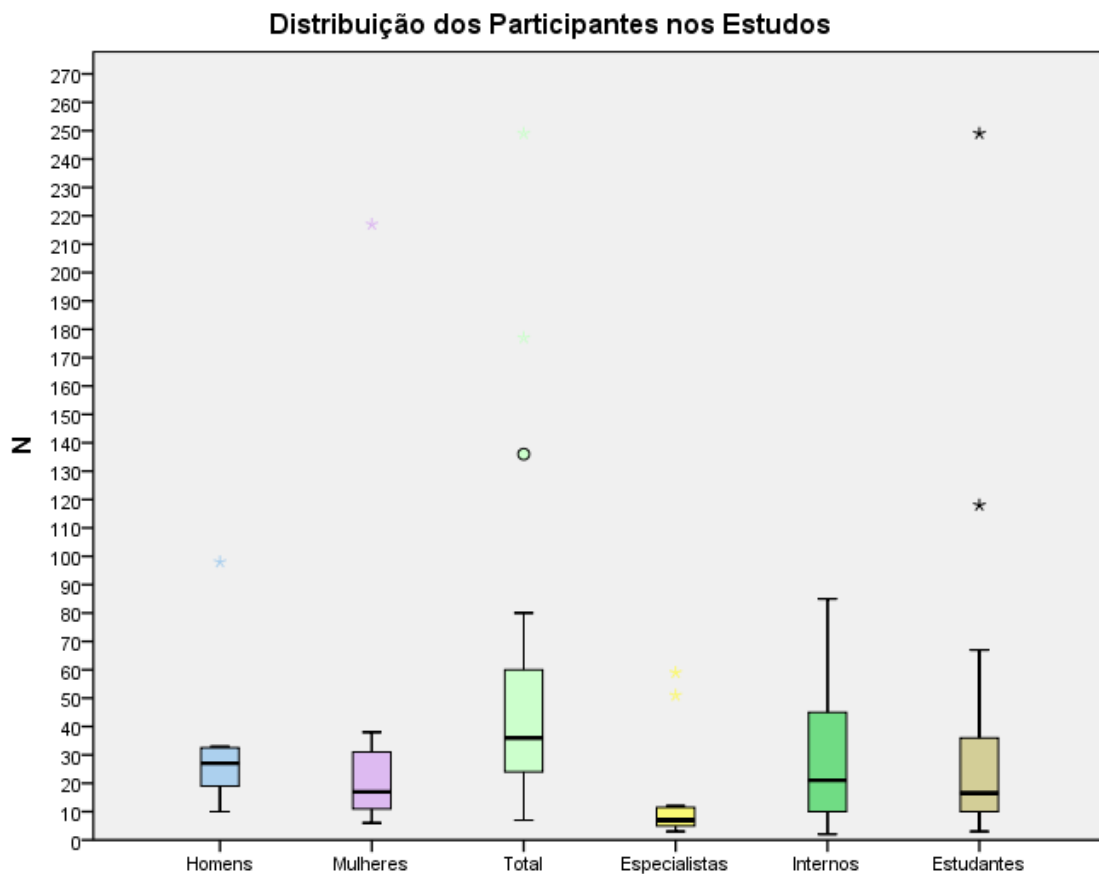


Gráfico 3: Distribuição demográfica das amostras na totalidade dos estudos.

Ao analisar a melhoria na *performance*, representada na Tabela 4 e no Gráfico 4, o grupo dos médicos internos mais experientes (tipicamente considerados entre o 3º ano de internato até ao final do mesmo) teve uma maior média da melhoria, em cerca

de 77,33%, seguido do grupo dos estudantes, com 45,04% melhores resultados. Os especialistas atingiram uma melhoria menor, de cerca de 33,35%.

Melhoria de <i>Performance</i>				
	Estudantes	Internos Novatos	Internos Experientes	Especialistas
Média Ponderada do Teste	45,04%	43,74%	77,33%	33,35%

Tabela 4: Melhoria de *performance*, em percentagem, estratificada de acordo com a experiência.

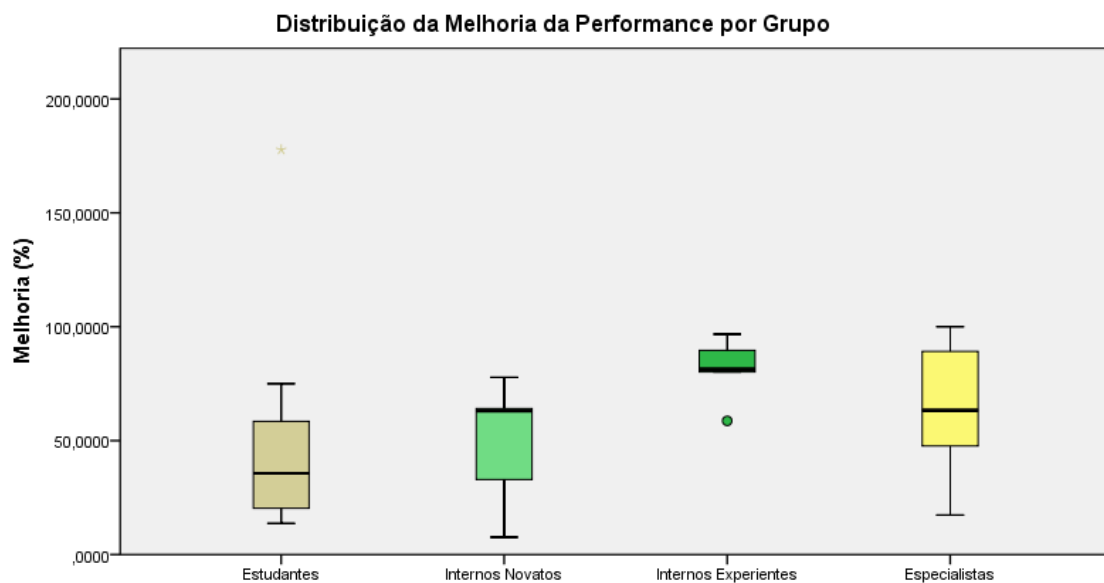


Gráfico 4: Distribuição da melhoria da *performance* por grupo.

No Gráfico 5, foram incluídos os estudos em que existiam resultados referentes a mais que um grupo de indivíduos no que diz respeito à sua experiência. Foram criados dois grupos: o grupo experimental que corresponde aos indivíduos menos experientes (estudantes e internos novatos) enquanto o grupo de controlo é representado pelos internos mais velhos e especialistas. O *Risk Ratio* é de 0,88, sendo que o seu intervalo de confiança se estende de 0,67 a 1,15, com *p-value* de 0,18. A heterogeneidade é de 30%.

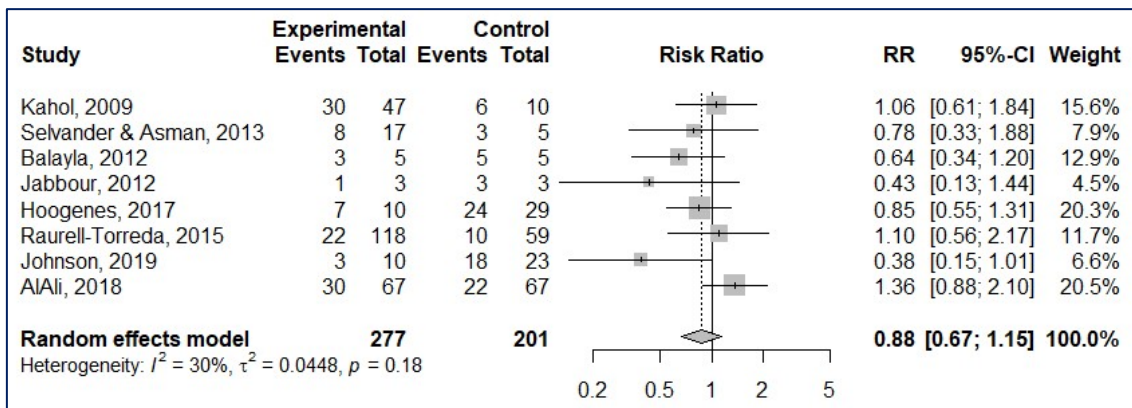


Gráfico 5: *Forest Plot* correspondente à análise dos resultados dos estudos que incluíam mais de um grupo de indivíduos (estudantes, internos e especialistas). O grupo *Experimental* corresponde aos menos experientes (estudantes e internos) enquanto o *Control* corresponde aos mais experientes (internos “sênior” e especialistas). *Events* corresponde ao número de acertos normalizado pelo número de participantes do estudo, representado na coluna *Total*. O *Risk Ratio* (RR) representa o efeito de cada estudo quanto ao resultado dos testes de performance. Se o valor for superior a 1, os menos experientes tiveram melhores resultados, estando o peso de cada estudo representado na coluna *Weight*. Finalmente, estão também representados o valor do teste de heterogeneidade (qui-quadrado), a inconsistência e o teste de hipóteses.

Discussão

Após a leitura e seleção dos artigos, é possível destacar a sequência metodológica no que diz respeito à criação e implementação de um simulador ou protocolo de simulação (Emken et al, 2004). Após identificar as variáveis necessárias para avaliar o procedimento pretendido, tem que se criar um instrumento ou protocolo adequados para as avaliar. Estes terão que ser validados e assim poder proceder a uma seleção, certificação e avaliação das aptidões de cada profissional. Um dos métodos mais frequentemente utilizados para esta validação passa pela comparação dos resultados obtidos pelos grupos mais e menos experientes. Se os mais experientes obtiverem melhores resultados, pode considerar-se que a validade foi atingida (Beard et al, 2011). É fundamental que também se verifique uma aprendizagem das técnicas em que os simuladores se focam, sendo esse o verdadeiro “*outcome*” deste tipo de instrumento. Tal é verificado com uma melhoria dos resultados com o tempo de utilização dos instrumentos. Estas avaliações são feitas com recurso a escalas ou protocolos devidamente aprovados para os procedimentos em questão. Finalmente, é necessário verificar a transição das aptidões adquiridas para o contexto, por exemplo, do bloco operatório. Este parece ser, portanto, um processo moroso, razão pela qual existem poucos estudos nesta última fase.

Para construção de um protocolo de avaliação ou atuação, é preciso criar os critérios baseados nas fases ou momentos mais importantes do procedimento, normalmente identificados por especialistas nos mesmos, podendo, contudo, ser também específicos a uma única técnica. São comparados a outros métodos estabelecidos, de forma a perceber a sua utilidade perante o panorama mais recente e desta forma poder ser aplicado tanto a nível do ensino como da avaliação.

É de destacar que também é possível avaliar os avaliadores, tal como referido em Vernez et al. (2017). Este tem como objetivo perceber se a população não médica pode avaliar os procedimentos através de grelhas de avaliação com a mesma qualidade que os especialistas. Ou seja, esse estudo avalia a clareza dos critérios e a facilidade de utilização dos mesmos, independentemente da experiência do avaliador.

Dado que tanto os simuladores como as escalas de avaliação são, na sua maioria, específicos para um tipo de procedimento, este processo de criação está completo num reduzido número de casos e parece envolver amostras reduzidas. Devido a este menor número de indivíduos em cada estudo, as amostras são mal caracterizadas, sendo predominante a inexistência de uma inferência quanto ao género, idade ou grau de especialização dos mesmos.

Atendendo às avaliações de cada um dos estudos de acordo com as várias escalas, é possível verificar que não existem pontuações elevadas. A descrição das avaliações feitas através das várias escalas foi explicitada anteriormente, na secção dos resultados. Destaca-se a presença de alguns vieses, já que, devido à própria natureza destes estudos, a ocultação e randomização associadas são praticamente impossíveis. Pela ausência de estudos de segurança, poder-se-á afirmar que estão também presentes vieses de publicação, já que os que apresentavam métodos para estudo de segurança apresentaram resultados negativos. Por definição, dado que uma das características do ensino por simulação é a inexistência de doentes, a segurança dos mesmos não é posta em causa, podendo este facto justificar a ausência destes estudos. Contudo, estes são essenciais no que diz respeito à validação dos simuladores enquanto ferramentas que permitem a transição para um contexto real. Desta forma, apesar de os resultados serem promissores, parece ser prematuro concluir que todos os simuladores e todas as escalas criadas nestes estudos estão aptas a uma utilização generalizada por parte das escolas médicas. Aliás, parecem não existir critérios para avaliação de competências sob

a forma de ferramentas fiáveis e validadas, mesmo para procedimentos específicos, em que a validade é mínima (Morris et al., 2012). Refere-se também que os critérios menos selecionados com recurso à escala PRISMA estão tipicamente presentes em revisões sistemáticas ou meta-análises (nas quais são essenciais, sendo estas o alvo da escala referida). Tendo em conta que estas foram excluídas desta fase do estudo, seria de esperar que esses fossem os critérios menos atribuídos.

Conforme ilustrado na Tabela 4, a melhoria da *performance* tem uma tendência decrescente conforme a experiência do participante, tal como sugerido em Oquendo et al. (2018). O facto de estudantes terem uma taxa de melhoria maior através do uso de simulação, indica que estes novos instrumentos são capazes de levar a uma aprendizagem eficaz das aptidões básicas no que diz respeito a procedimentos cirúrgicos, conduzindo a uma maior confiança por parte dos estudantes. Apesar de mais reduzida, também os especialistas apresentaram uma melhoria. Tal poderá ser indicativo de que um método mais padronizado, com feedback imediato, com uma sequência pré-determinada de ações e talvez com acesso a técnicas mais avançadas poderá permitir uma aprendizagem nesta faixa com mais experiência. Outra hipótese que explica estes resultados pode ser o facto de estes instrumentos serem mais precisos que avaliadores ou instrutores, permitindo detetar imperfeições (como as associadas à velocidade de execução (Uemura et al., 2016) ou a maus hábitos adquiridos durante a prática clínica diária (Raurell-Torredà et al., 2015)) e, posteriormente, diminuir o risco de erro humano.

Como podemos observar no Gráfico 4, o valor do *Risk Ratio* é de 0,88, indicativo de que o grupo experimental teve resultados reduzidos em 12% quando comparados com o grupo de controlo, mais experiente conforme seria esperado. Porém, este resultado não parece ser estatisticamente significativo sendo o *p-value* de 0,18 e o intervalo de confiança encontra-se entre 0,67 e 1,15, ultrapassando a linha de indiferença. É de destacar, contudo, que a heterogeneidade da amostra é de 30%, representando uma inconsistência moderada dos resultados.

Ao comparar os resultados explicitados no Gráfico 4 com os da Tabela 4, demonstra-se um melhor resultado a nível da *performance* no grupo dos especialistas, mas são os menos experientes que parecem tirar maior proveito do ensino por simulação, com valores maiores de melhoria dos resultados.

Limitações

A presente revisão sistemática encontra-se principalmente limitada pelo conteúdo dos artigos avaliados. Primeiro, existe um número limitado de estudos que abordam o uso da simulação no processo da aprendizagem. Tal poderá conduzir a uma população mais heterogênea, como descrito anteriormente. Desta forma os grupos estudados terão qualidades distintas, especialmente no que diz respeito ao seu grau de formação, o que resulta numa maior dificuldade para realização de estudos mais criteriosos, com resultados coerentes. Tal facto poderá ter enviesado alguns resultados, já que muitos dos estudos apresentavam resultados referentes apenas a um dos graus de formação, sendo por isso necessário um melhor controlo das amostras.

Outra limitação passa pela interpretação das indicações fornecidas pelas escalas (Goldet & Howick, 2013; Page et al., 2021; Verhagen et al., 1998) usadas para avaliar os estudos. Para além de ser necessário decidir se cada item deve ser atribuído, foram adaptados alguns destes de forma a conseguir proceder à comparação direta dos vários artigos.

Conclusão

Com esta revisão, é possível concluir que a simulação realística e o seu uso na educação médica são áreas atualmente em exploração. Dado ser uma investigação recente, existem ainda poucos estudos realizados e a grande maioria deles pretende introduzir novos instrumentos e escalas ou protocolos. Dos que foram analisados, apresentam inferências restrita a um único grupo devido à heterogeneidade. Contudo, com os resultados promissores das primeiras fases, é aparente que seja esse o resultado esperado para cada um deles.

Estas novas metodologias podem ser estendidas às restantes áreas da Medicina, já que poderão permitir uma aprendizagem autónoma e com feedback imediato de forma a poder consolidar e integrar os conhecimentos e aptidões adquiridos e aplicá-los a uma vivência clínica, complementando o ensino teórico mais tradicional sem comprometer a segurança dos doentes e sem necessitar de uma tão grande disponibilidade assistencial no ensino. De forma que tal seja possível, será necessário que terminem as várias fases dos simuladores e escalas que estão a ser criados e construir mais destes instrumentos de forma a expandir a sua utilização. É também útil a realização de estudos com intervalos temporais mais prolongados que investiguem as alterações na *performance* ao longo do tempo.

Referências Bibliográficas

- Abdelsattar, J. M., Pandian, T. K., Finnesgard, E. J., el Khatib, M. M., Rowse, P. G., Buckarma, E. H., Gas, B. L., Heller, S. F., & Farley, D. R. (2015). Do You See What i See? How We Use Video as an Adjunct to General Surgery Resident Education. *Journal of Surgical Education, 72*(6), 145–150. <https://doi.org/10.1016/j.jsurg.2015.07.012>
- AlAli, A. B., Griffin, M. F., Calonge, W. M., & Butler, P. E. (2018). Evaluating the Use of Cleft Lip and Palate 3D-Printed Models as a Teaching Aid. *Journal of Surgical Education, 75*(1), 200–208. <https://doi.org/10.1016/j.jsurg.2017.07.023>
- Bajka, M., Tuchschnid, S., Fink, D., Székely, G., & Harders, M. (2010). Establishing construct validity of a virtual-reality training simulator for hysteroscopy via a multimetric scoring system. *Surgical Endoscopy, 24*(1), 79–88. <https://doi.org/10.1007/s00464-009-0582-4>
- Balayla, J., Bergman, S., Ghitulescu, G., Feldman, L. S., & Fraser, S. A. (2012). Knowing the operative game plan: A novel tool for the assessment of surgical procedural knowledge. *Canadian Journal of Surgery, 55*(4 SUPPL. 2). <https://doi.org/10.1503/cjs.015411>
- Beard, J. D., Marriott, J., Purdie, H., & Crossley, J. (2011). Assessing the surgical skills of trainees in the operating theatre: A prospective observational study of the methodology. *Health Technology Assessment, 15*(1), 3–112. <https://doi.org/10.3310/hta15010>
- Brinkmann, C., Fritz, M., Pankratius, U., Bahde, R., Neumann, P., Schlueter, S., Senninger, N., & Rijcken, E. (2017). Box- or Virtual-Reality Trainer: Which Tool Results in Better Transfer of Laparoscopic Basic Skills?—A Prospective Randomized Trial. *Journal of Surgical Education, 74*(4), 724–735. <https://doi.org/10.1016/j.jsurg.2016.12.009>
- Clinkard, D., Moulton, E., Holden, M., Davison, C., Ungi, T., Fichtinger, G., & McGraw, R. (2015). Assessment of Lumbar Puncture Skill in Experts and Nonexperts Using Checklists and Quantitative Tracking of Needle Trajectories: Implications for Competency-Based Medical Education. *Teaching and Learning in Medicine, 27*(1), 51–56. <https://doi.org/10.1080/10401334.2014.979184>

- Custers, E., & Cate, O. T. (2018). The History of Medical Education in Europe and the United States, With Respect to Time and Proficiency. *Academic medicine: journal of the Association of American Medical Colleges*, 93(3S Competency-Based, Time-Variable Education in the Health Professions), S49–S54. <https://doi.org/10.1097/ACM.0000000000002079>
- Dydykin, S., & Kapitonova, M. (2015). The role of student surgical interest groups and surgical Olympiads in anatomical and surgical undergraduate training in Russia. *Anatomical Sciences Education*, 8(5), 471–477. <https://doi.org/10.1002/ase.1523>
- Emken, J. L., Mcdougall, E. M., & Clayman, R. V. (2004). Training and assessment of laparoscopic skills. *JSL : Journal of the Society of Laparoendoscopic Surgeons*, 8(2), 195–199.
- Galvez, V., Fuente, R. de la, Meneses, C., Leiva, L., Fagalde, G., Herskovic, V., Fuentes, R., Munoz-Gama, J., & Sepúlveda, M. (2020). Process-oriented instrument and taxonomy for teaching surgical procedures in medical training: The ultrasound-guided insertion of central venous catheter. *International Journal of Environmental Research and Public Health*, 17(11), 1–15. <https://doi.org/10.3390/ijerph17113849>
- Georgescu, M., Tanoubi, I., Drolet, P., Robitaille, A., Perron, R., & Patenaude, J. V. (2015). Simulation á haute fidélité: Moyen d'enseignement des principes de gestion de crises périopératoires aux résidents de chirurgie [Hi-Fi simulation: Teaching crisis resource management to surgery residents]. *La Tunisie medicale*, 93(2), 63–65.
- Goldet, G., & Howick, J. (2013). Understanding GRADE: An introduction. *Journal of Evidence-Based Medicine*, 6(1), 50–54. <https://doi.org/10.1111/jebm.12018>
- Guérard-Poirier, N., Beniey, M., Meloche-Dumas, L., Lebel-Guay, F., Misheva, B., Abbas, M., Dhane, M., Elraheb, M., Dubrowski, A., & Patocskai, E. (2020). An educational network for surgical education supported by gamification elements: Protocol for a randomized controlled trial. *JMIR Research Protocols*, 9(12:15). <https://doi.org/10.2196/21273>
- Hand, F., Stirling, A., Felle, P., Conlon, K., & Ridgway, P. (2016). The e-handover: applications for surgical training. *The clinical teacher*, 14(3), 211–215. <https://doi.org/10.1111/tct.12536>

- Hofstad, E. F., Våpenstad, C., Bø, L. E., Langø, T., Kuhry, E., & Mårvik, R. (2017). Psychomotor skills assessment by motion analysis in minimally invasive surgery on an animal organ. *Minimally Invasive Therapy and Allied Technologies*, 26(4), 240–248. <https://doi.org/10.1080/13645706.2017.1284131>
- Hoogenes, J., Wong, N., Al-Harbi, B., Kim, K. S., Vij, S., Bolognone, E., Quantz, M., Guo, Y., Shayegan, B., & Matsumoto, E. D. (2017). A Randomized Comparison of 2 Robotic Virtual Reality Simulators and Evaluation of Trainees' Skills Transfer to a Simulated Robotic Urethrovesical Anastomosis Task. *Urology*, 111, 110–115. <https://doi.org/10.1016/j.urology.2017.09.023>
- Hu, W. G., Feng, J. Y., Wang, J., Song, Y. J., Xu, X. T., Zhou, H., & Huang, C. B. (2015). Ureteroscopy and cystoscopy training: comparison between transparent and non-transparent simulators. *BMC medical education*, 15, 93. <https://doi.org/10.1186/s12909-015-0380-8>
- Iyer, M. S., Santen, S. A., Nypaver, M., Warriar, K., Bradin, S., Chapman, R., Mcallister, J., Vredevelde, J., & House, J. B. (2013). Assessing the validity evidence of an objective structured assessment tool of technical skills for neonatal lumbar punctures. *Academic Emergency Medicine*. 20(3), 321–324. <https://doi.org/10.1111/acem.12093>
- Jabbour, N., Reihisen, T., Payne, N. R., Finkelstein, M., Sweet, R. M., & Sidman, J. D. (2012). Validated assessment tools for pediatric airway endoscopy simulation. *Otolaryngology - Head and Neck Surgery (United States)*, 147(6), 1131–1135. <https://doi.org/10.1177/0194599812459703>
- Johnson, J. D., Cheng, C., Schmidtberg, B., Cote, M., & Geaney, L. E. (2019). Assessment of Basic Ankle Arthroscopy Skills in Orthopedic Trainees. *Foot and Ankle International*, 41(2), 193–199. <https://doi.org/10.1177/1071100719891418>
- Kahol, K., Vankipuram, M., & Smith, M. L. (2009). Cognitive simulators for medical education and training. *Journal of Biomedical Informatics*, 42(4), 593–604. <https://doi.org/10.1016/j.jbi.2009.02.008>

- Kanumuri, P., Ganai, S., Wohaibi, E. M., Bush, R. W., Grow, D. R., & Seymour, N. E. (2008). Virtual reality and computer-enhanced training devices equally improve laparoscopic surgical skill in novices. *JSLs: Journal of the Society of Laparoendoscopic Surgeons*, *12*(3), 219–226.
- Kogan, J. R., Holmboe, E. S., & Hauer, K. E. (2009). Tools for direct observation and assessment of clinical skills of medical trainees: a systematic review. *JAMA*, *302*(12), 1316–1326. <https://doi.org/10.1001/jama.2009.1365>
- Maertens, H., Madani, A., Landry, T., Vermassen, F., van Herzeele, I., & Aggarwal, R. (2016). Systematic review of e-learning for surgical training. *The British journal of surgery*, *103*(11), 1428–1437. <https://doi.org/10.1002/bjs.10236>
- Meier, A. H., Boehler, M. L., McDowell, C. M., Schwind, C., Markwell, S., Roberts, N. K., & Sanfey, H. (2012). A surgical simulation curriculum for senior medical students based on TeamSTEPPS. *Archives of surgery (Chicago, Ill.: 1960)*, *147*(8), 761–766. <https://doi.org/10.1001/archsurg.2012.1340>
- Morris, M. C., Gallagher, T. K., & Ridgway, P. F. (2012). Tools used to assess medical students competence in procedural skills at the end of a primary medical degree: a systematic review. *Medical education online*, *17*. <https://doi.org/10.3402/meo.v17i0.18398>
- Oquendo, Y. A., Riddle, E. W., Hiller, D., Blinman, T. A., & Kuchenbecker, K. J. (2018). Automatically rating trainee skill at a pediatric laparoscopic suturing task. *Surgical Endoscopy*, *32*(4), 1840–1857. <https://doi.org/10.1007/s00464-017-5873-6>
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., McGuinness, L. A., ... Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ (Clinical research ed.)*, *372*, n71. <https://doi.org/10.1136/bmj.n71>
- Pellen, M. G. C., Horgan, L. F., Barton, J. R., & Attwood, S. E. (2009). Construct validity of the ProMIS laparoscopic simulator. *Surgical Endoscopy and Other Interventional Techniques*, *23*(1), 130–139. <https://doi.org/10.1007/s00464-008-0066-y>

- Rafiq, A., Tamariz, F., Boanca, C., Lavrentyev, V., & Merrell, R. C. (2008). Objective Assessment of Training Surgical Skills Using Simulated Tissue Interface with Real-Time Feedback. *Journal of Surgical Education*, 65(4), 270–274. <https://doi.org/10.1016/j.jsurg.2008.05.012>
- Raurell-Torredà, M., Olivet-Pujol, J., Romero-Collado, À., Malagon-Aguilera, M. C., Patiño-Masó, J., & Baltasar-Bagué, A. (2015). Case-Based Learning and Simulation: Useful Tools to Enhance Nurses' Education? Nonrandomized Controlled Trial. *Journal of Nursing Scholarship*, 47(1), 34–42. <https://doi.org/10.1111/jnu.12113>
- Roldán-Merino, J., Farrés-Tarafa, M., Estrada-Masllorens, J. M., Hurtado-Pardos, B., Miguel-Ruiz, D., Nebot-Bergua, C., Insa-Calderon, E., Grané-Mascarell, N., Bande-Julian, D., Falcó-Pergueroles, A. M., Lluch-Canut, M. T., & Casas, I. (2019). Reliability and validity study of the Spanish adaptation of the “Creighton Simulation Evaluation Instrument (C-SEI).” *Nurse Education in Practice*, 35, 14–20. <https://doi.org/10.1016/j.nepr.2018.12.007>
- Selvander, M., & Åsman, P. (2013). Cataract surgeons outperform medical students in Eyesi virtual reality cataract surgery: Evidence for construct validity. *Acta Ophthalmologica*, 91(5), 469–474. <https://doi.org/10.1111/j.1755-3768.2012.02440.x>
- Shen, Z., Yang, F., Gao, P., Zeng, L., Jiang, G., Wang, S., Ye, Y., & Zhu, F. (2017). A Novel Clinical-Simulated Suture Education for Basic Surgical Skill: Suture on the Biological Tissue Fixed on Standardized Patient Evaluated with Objective Structured Assessment of Technical Skill (OSATS) Tools. *Journal of Investigative Surgery*, 31(4), 333–339. <https://doi.org/10.1080/08941939.2017.1319994>
- Sun, C., & Qi, X. (2018). Evaluation of Problem- and Simulator-Based Learning in Lumbar Puncture in Adult Neurology Residency Training. *World Neurosurgery*, 109, 807–811. <https://doi.org/10.1016/j.wneu.2017.10.093>
- Tanoue, N., Korovin, L. N., Carton, M., Galvani, C. A., & Ghaderi, I. (2018). Faculty feedback versus residents' self-assessment of operative performance: Different but complementary. *American Journal of Surgery*, 215(2), 288–292. <https://doi.org/10.1016/j.amjsurg.2017.11.016>

- Taylor, L. K., Thomas, G. W., Karam, M. D., Kreiter, C. D., & Anderson, D. D. (2016). Assessing Wire Navigation Performance in the Operating Room. *Journal of Surgical Education*, 73(5), 780–787. <https://doi.org/10.1016/j.jsurg.2016.03.018>
- Torgerson, C. S., Brydges, R., Chen, J. M., & Dubrowski, A. (2007). Drilling simulated temporal bones with left-handed tools: a left-hander's right?. *The Annals of otology, rhinology, and laryngology*, 116(11), 819–826. <https://doi.org/10.1177/000348940711601105>
- Uemura, M., Jannin, P., Yamashita, M., Tomikawa, M., Akahoshi, T., Obata, S., Souzaki, R., Ieiri, S., & Hashizume, M. (2016). Procedural surgical skill assessment in laparoscopic training environments. *International Journal of Computer Assisted Radiology and Surgery*, 11(4), 543–552. <https://doi.org/10.1007/s11548-015-1274-2>
- Verhagen, A. P., de Vet, H. C., de Bie, R. A., Kessels, A. G., Boers, M., Bouter, L. M., & Knipschild, P. G. (1998). The Delphi list: a criteria list for quality assessment of randomized clinical trials for conducting systematic reviews developed by Delphi consensus. *Journal of clinical epidemiology*, 51(12), 1235–1241. [https://doi.org/10.1016/s0895-4356\(98\)00131-0](https://doi.org/10.1016/s0895-4356(98)00131-0)
- Vernez, S. L., Huynh, V., Osann, K., Okhunov, Z., Landman, J., & Clayman, R. V. (2017). C-SATS: Assessing surgical skills among urology residency applicants. *Journal of Endourology*, 31, S95–S100. <https://doi.org/10.1089/end.2016.0569>

Anexos

Título	DOI/PMID	Data	Autores	Língua	Tipo de estudo	Tema	Outcome principal
Cognitive simulators for medical education and training	10.1016/j.jbi.2009.02.008	2009	Kahol, K., et al	Inglês	Estudo Observacional	Avaliação e comparação de métodos de avaliação	Avaliação de escalas
Cataract surgeons outperform medical students in Eyesi virtual reality cataract surgery: evidence for construct validity	10.1111/j.1755-3768.2012.02440.x	2012	Selvander, M., Åsman, P.	Inglês	Estudo Retrospectivo	Avaliação de ferramentas de avaliação final	Avaliação de Escala
Objective Assessment of training surgical skills using simulated tissue interface with real-time feedback	10.1016/j.jsurg.2008.05.012	2008	Rafiq, A., et al	Inglês	Estudo Prospetivo	Avaliação de simulador	Avaliação de instrumento
Assessing the Validity Evidence of an Objective Structured Assessment Tool of Technical Skills for Neonatal Lumbar Punctures	10.1111/acem.12093	2013	Maya, S. et al	Inglês	Estudo Prospetivo	Criação de protocolo	Avaliação de Escala
Knowing the operative game plan: a novel tool for the assessment of surgical procedural knowledge	10.1503/cjs.015411	nov. 2011	Balayla, J., et al	Inglês	Estudo Observacional	Caraterização de procedimentos e técnicas	Avaliação de instrumento
Reliability and validity study of the Spanish adaptation of the "Creighton Simulation Evaluation Instrument (C-SEI)"	10.1016/j.nepr.2018.12.007	2019	Roldán-Merino, J., et al	Inglês	Estudo Transversal	Comparação de avaliadores	Avaliação de Escala
Virtual reality and computer-enhanced training devices equally improve laparoscopic surgical skill in novices	http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3015873/	2008	Kanumuri, P., et al	Inglês	Estudo Transversal	Adaptação de protocolo para outra língua/Validação de protocolo	Avaliação de Escala
Ureteroscopy and cystoscopy training: comparison between transparent and non-transparent simulators	10.1186/s12909-015-0380-8	2015	Hu, W. G., et al	Inglês	Estudo Observacional	Avaliação de simulador	Avaliação de instrumento

Assessing Wire Navigation Performance in the Operating Room	10.1016/j.jsurg.2016.03.018	2016	Taylor, L. K., et al	Inglês	Estudo Transversal	Avaliação de simulador	Avaliação de instrumento
Automatically rating trainee skill at a pediatric laparoscopic suturing task	10.1007/s00464-017-5873-6	2017	Oquendo, Y. A., et al	Inglês	Estudo Observacional	Avaliação de objetiva de técnica por meio de sensor	Avaliação de Instrumento
Validated Assessment Tools for Pediatric Airway Endoscopy Simulation	10.1177/0194599812459703	2012	Jabbour, N., et al	Inglês	Estudo Transversal	Criação de instrumento de avaliação objetiva automática	Avaliação de instrumento
Do you see what I see? How we use video as an adjunct to General Surgery Resident Education	10.1016/j.jsurg.2015.07.012	2015	Abdelsattar, J. M., et al	Inglês	Transversal	Avaliação de método de ensino	Avaliação de Escala
A randomized comparison of 2 robotic virtual reality simulators and evaluation of trainees' skills transfer to a simulated robotic urethrovesical anastomosis task	10.1016/j.urology.2017.09.023	2017	Hoogenes, J., et al	Inglês	Ensaio Randomizado	Avaliação de vídeos	Avaliação de Escala
Evaluation of problem- and simulator-based learning in lumbar puncture in adult neurology residency training	10.1016/j.wneu.2017.10.093	2018	Sun, C., Qi, X.	Inglês	Estudo Observacional	Avaliação de protocolo; Comparação de protocolos	Avaliação de Escala
Assessment of lumbar puncture skill in experts and nonexperts using checklists and quantitative tracking of needle trajectories: implications for competency-based medical education	10.1080/10401334.2014.979184	2015	Clinkard, D., et al	Inglês	Estudo Prospetivo	Avaliação de método de ensino; Avaliação de transferência de aptidões	Avaliação de Escala
Case-based learning and simulation: useful tool to enhance nurses's education? Nonrandomized controlled trial	10.1111/jnu.12113	2015	Raurell-Torredà, M., et al	Inglês	Estudo Observacional	Avaliação de simulador	Avaliação de instrumento
Procedural surgical skill assessment in laparoscopic training environments	10.1007/s11548-015-1274-2	2016	Uemura, M., et al	Inglês	Estudo Observacional	Avaliação de simulador	Avaliação de instrumento

Box- or virtual-reality trainer: which tool results in better transfer of laparoscopic basic skills? - A prospective randomized trial	10.1016/j.jsurg.2016.12.009	2017	Brinkmann, C., et al	Inglês	Ensaio controlado não-randomizado	Avaliação de protocolo	Avaliação de Escala
The e-handover: applications for surgical training	10.1111/tct.12536	2017	Hand, F., et al	Inglês	Estudo Observacional	Avaliação de simuladores; Comparação de simuladores	Avaliação de instrumento
A Novel Clinical-Simulated Suture Education for Basic Surgical Skill: Suture on the Biological Tissue Fixed on Standardized Patient Evaluated with Objective Structured Assessment of Technical Skill (OSATS) Tools	10.1080/08941939.2017.1319994	2017	Shen, Z., et al	Inglês	Estudo Observacional	Avaliação de ferramenta de avaliação e ensino	Avaliação de instrumento
Assessment of Basic Ankle Arthroscopy Skills in Orthopedic Trainees	10.1177/1071100719891418	2020	Johnson, J. D., et al	Inglês	Ensaio Observacional	Avaliação de instrumento de avaliação/protocolo	Avaliação de escala
Process-oriented instrument and taxonomy for teaching surgical procedures in medical training: the ultrasound-guided insertion of central venous catheter	10.3390/ijerph17113849	2020	Galvez, V., et al	Inglês	Estudo Observacional	Avaliação de protocolo	Avaliação de escala
Faculty feedback versus residents' self-assessment of operative performance: different but complementary	10.1016/j.amjsurg.2017.11.016	2017	Tanoue, N., et al	Inglês	Transversal	Avaliação de método de avaliação	Avaliação de escala
Evaluating the use of cleft lip and palate 3D-printed models as a teaching aid	10.1016/j.jsurg.2017.07.023	2018	AlAli, A. B., et al	Inglês	Estudo Observacional	Foco novatos vs especialistas	Avaliação de escala

Tabela Suplementar 1: Caracterização dos estudos que cumprem os critérios de inclusão para análise estatística.

Título	Data	DOI	Autores	Exclusão	Motivo de Exclusão
Tools for direct observation and assessment of clinical skills of medical trainees: a systematic review	2009	10.1001/jama.2009.1365	Kogan, J., et al	Do estudo	Revisão sistemática que exclui artigos que descrevem "ferramentas para uso em contexto de simulação, para avaliar aptidões processuais ou cirúrgicas" (Pág. 2, "Study Selection")
[Hi-Fi simulation: Teaching crisis resource management to surgery residents]	2015		Georgescu, M., et al		Artigo escrito apenas em francês
Drilling simulated temporal bones with left-handed tools: a left-hander's right?	2007	10.1177/000348940711601105	Torgerson, C., et al		Artigo que estuda a eficácia do uso de uma ferramenta para canhotos vs ferramenta para destros usada por destros vs ferramenta para destros usada por canhotos (com recurso a modelo)
The role of student surgical interest groups and surgical Olympiads in anatomical and surgical undergraduate training in Russia	2015	10.1002/ase.1523	Dydykin, S., Kapitonova, M.		Não há variáveis ou informações; Artigo de opinião
Tools used to assess medical students competence in procedural skills at the end of a primary medical degree: a systematic review	2012	10.3402/meo.v17i0.18398	Morris, M. C., et al	Da análise estatística	Não é estudo interventivo (revisão sistemática)
Assessing the Validity Evidence of an Objective Structured Assessment Tool of Technical Skills for Neonatal Lumbar Punctures	2013	10.3310/hta15010	Beard, J. D., et al		Não é estudo interventivo (conjunto de <i>guidelines</i>)
Training and assessment of laparoscopic skills	2004	https://www.ncbi.nlm.nih.gov/pubmed/15119670	Emken, J. L., et al		Descreve base de dados
C-SATS: Assessing Surgical Skills Among Urology Residency Applicants	2017	10.1089 / final.2016.0569	Vernez, S. L., et al		Grupo experimental não é comparável aos restantes
Systematic review of e-learning for surgical training	2016		Maertens, H., et al		Não é estudo interventivo (revisão sistemática)
A Surgical Simulation Curriculum for Senior Medical Student Based on TeamSTEPPS	2012	10.1002/bjs.10236	Meier, A. H., et al		Escala de avaliação não permite comparação com as restantes
Construct validity of the ProMIS laparoscopic simulator	2009	10.1001/archsurg.2012.1340	Pellen, M. G. C., et al		Escala de avaliação não permite comparação com as restantes

Establishing construct validity of a virtual-reality training simulator for hysteroscopy via a multimetric scoring system	2009	10.1007/s00464-008-0066-y	Bajka, M., et al		Resultados descritos enquanto comparação entre grupos de estudo
Psychomotor skills assessment by motion analysis in minimally invasive surgery on an animal organ	2017	10.1007/s00464-009-0582-4	Hofstad, E. F., et al		Resultados apresentados sob a forma de gráficos
An educational network for surgical education supported by gamification elements: protocol for a randomized controlled trial	2020	10.1080/13645706.2017.1284131	Guérard-Poirier, N., et al		Não é estudo interventivo (protocolo)

Tabela Suplementar 2: Caraterização dos artigos excluídos da amostra de estudo.

	Participantes	Homens	Mulheres	Idade	Estudantes	Internos	Especialistas	Delphi	PRISMA
Selvander & Asman, 2013	24	.	.	.	17	2	5	2	16
Rafiq, 2008	12	.	.	.	12	.	.	4	9
Balayla, 2012	35	.	.	.	5	25	5	3	17
Galvez, 2020	25	3	17
Johnson, 2019	33	27	6	29	10	23	.	4	17
Sun & Qi, 2018	60	60	.	4	14
Iyer, 2013	16	16	.	5	12
Kanumuri, 2008	16	10	6	26	16	.	.	6	15
Uemura, 2016	26	.	.	.	5	10	11	2	13
Hand, 2016	17	17	.	3	15
Clinkard, 2015	37	.	.	.	32	5	.	3	12
Jabbour, 2012	23	.	.	.	3	17	3	2	13
Oquendo, 2017	32	15	17	.	6	21	5	4	16
Brinkmann, 2017	36	.	.	.	36	.	.	3	12
Raurell-Torredà, 2015	177	.	.	23	118	.	59	3	13
Abdelsattar, 2015	62	.	.	.	10	45	7	5	13
Hoogenes, 2017	39	23	16	28	10	29	.	8	12
Taylor, 2016	7	7	.	2	14
Roldán-Merino, 2019	249	32	217	22,7	249	.	.	4	17
Shen, 2017	42	.	.	.	42	.	.	4	15
Hu, 2015	36	.	.	.	36	3	7	4	15
AlAli, 2018	67	.	.	.	67	.	.	4	14
Kahol, 2009	57	33	24	.	.	47	10	2	16
Tanoue, 2018	80	80	.	4	16
N válido	24	6	6	5	17	16	9	24	24
Mínimo	7	10	6	22,7	3	2	3	2	9
Máximo	249	33	217	29	249	80	59	8	17
Média	50	23	48	25,7	40	25	12	3,7	14,3
Desvio padrão	54	9	83	2,9	61	22	18	1,4	2,1

Tabela Suplementar 3: Caracterização demográfica dos estudos selecionados para análise estatística.