Advancing the assessment of clinical competence in Latin America: a scoping review of OSCE implementation and challenges in resource-limited settings

Soledad Armijo-Rivera[1*] , Brenda Fuenzalida-Muñoz[2] , Scarlett Vicencio-Clarke[2] , Alexandra Elbers-Arce[3,4] , Sergio Bozzo-Navarrete[5] , Natasha Kunakov[6] , Cesar Miranda-Hurtado[4,7] , Hector Shibao-Miyasato[8,9] , Jacqueline Sanhueza[2] , Carla Cornejo[2] , Alix Soublette[2] , Ana María Sandoval[9] , Fresia Cicibel Casas-Bueno[7] and Ximena Delgado[9]

Abstract

Background Objective Structured Clinical Examination (OSCE) is important to assess clinical competencies in health professions. However, in Latin America, a region with limited resources, the implementation and quality of OSCEs remain underexplored despite their increasing use. This study analyses how the OSCE has been applied and how its quality has evolved in Latin America.

Methods A scoping review methodology was used, including a search across PubMed, Scopus, WOS, LILACS and Scielo, including studies on the implementation of OSCE in Latin America, written in English, French, Portuguese, or Spanish. Their quality was assessed using the AMEE guidelines 81 and 49 criteria and MMERSQI. Data were extracted regarding OSCE structure, evaluator training, validity, reliability, and the use of simulated patients.

Results   365 articles were obtained, of which 69 met the inclusion criteria. The first report on OSCE implementation in the region dates back to 2000. Three countries accounted for 84.06% of the reports (Chile, Mexico, Brazil). 68.12% was applied in undergraduate programs. In this group, the implementation was mainly in Medicine (69.57%), with lesser use in physiotherapy (7.95%) and nursing (2.9%). The number of stations and duration of each varied, with 18-station circuits being the most common. Evidence of validity and reliability of the OSCE was reported in 26.09%, feedback to students in 33,33%, and simulated patient training in 37.68% of the reports. A notable trend in the quinquennial analysis is the increased use of high-fidelity simulations and the shift towards remote OSCEs during the pandemic. The inclusion of inactive stations, inadequate training for simulated patients, and the absence of evidence supporting instrument validation are recurrently reported challenges in OSCE studies. The overall methodological quality has improved, as evidenced by OSCE Committee and Blueprint in nearly 50% of the studies and rising MMERSQI scores, especially in recent years.

*Correspondence:
Soledad Armijo-Rivera
s.armijo@udd.cl

Full list of author information is available at the end of the article

Conclusion While there has been progress in OSCE implementation, particularly in medical education, gaps remain in standardization, validation, training, and resource allocation. Further efforts are needed to ensure consistent quality, particularly in training simulated patients, addressing inactive stations, and ensuring instrument reliability. Addressing these gaps is crucial for enhancing the effectiveness of OSCEs in resource-limited settings and advancing health professional education across the region.

Keywords Objective structured clinical examination, OSCE, Quality education, Latin America, Assessment

Introduction/background

The Objective Structured Clinical Examination (OSCE) constitutes an evaluative form to assess clinical and communication skills for large groups of health profession students, and it has evolved since its first implementation in 1975 with medical students [1]. OSCEs are used to assess a range of competencies, including history taking, physical examination, communication skills (explanation/advice/consent), practical/technical skills, and clinical reasoning related to the patient [2]. There are practical guides for planning, implementation [3, 4], and analysis of OSCE [5] that provide an evidence-based approach to practice, including specific best practice guidelines for nursing OSCEs [6]. From the perspective of the Kane validity framework [7], OSCEs should primarily be used to test clinical and communication skills, which was the first intended use on the exam. Consensus on performance assessment highlights that the validity of an OSCE is influenced by various factors, such as the assessment of diverse skills, the use of rating scales rather than a checklist, and its correlation with other variables like clinical performance and the curriculum itself [8]. Content validity of the assessment must be ensured, considering various sources. The content of the OSCE must be explicitly explained to ensure it accurately assesses the intended objectives [9]. Blueprinting a test for learning objectives/outcomes is crucial, as it ensures alignment between what is tested and the learning objectives [10]. The OSCE blueprint should detail the balance of skills across domains and confirm that course content areas are adequately tested, particularly in clinical interactions [8, 11]. All these aspects should be considered to implement the OSCE with high standards.

A high proportion of OSCE reviews in the existing literature have been written in English, and non-English-language databases, like Scielo and LILACS, have been excluded, biasing information from Latin America [12]. This language-based exclusion is particularly relevant given the region's delayed adoption of OSCE as an alternative to traditional clinical assessment. In Latin America, the first report that mentioned OSCE as an alternative to traditional clinical assessment at the end of the medical career emerged in Brazil 22 years after the OSCE's introduction [13]. Similarly, the first report, including students' results of an implemented OSCE, was performed in Chile in 2000 [14], 25 years after the first OSCE implementation. These reports reflect the region´s lower adoption of innovative and objective clinical assessment approaches and motivate the need to understand the quality of the exams implemented, exploring the challenges and solutions tailored to advancing clinical competence assessment under resource constraints.

The need to understand how the OSCE has been implemented in Latin America over the years and what new challenges its application faces in Latin American countries led us to our research question: *How has the OSCE been implemented, and how has its quality evolved in Latin America 50 years after its creation?*

Materials and methods

The research teams were fifteen researchers from two countries (Chile and Peru). Five had overseen OSCEs in undergraduate programs, three had overseen OSCEs in revalidation of medical degrees in Chile, all had simulation training and four had a master's degree in medical education or general education. Three of the team members had been trained in international OSCE programs. The team includes physicians, nurses, nutritionists, and kinesiologists. All the team members had content expertise, and five had experience conducting scoping reviews. A librarian assisted the team in developing the definitive search strategy.

We conducted a literature search in two phases. The first phase was a preliminary search oriented to identify the reviews related to OSCE in the literature. This first phase was necessary to identify our research question and decide whether conducting a review focused on original data from Latin America was appropriate. The second phase followed a six-step approach to scoping reviews [15], including the program, students' level, the purpose of the OSCE (summative or formative), planning and implementation variables [3], and reported quality analysis [5]. As inclusion criteria, we considered articles reporting data on OSCE implementation in Latin America until March 2024. Exclusion criteria included articles lacking implementation data (such as reviews, letters to the editor, position papers, reflections, or recommendations) and assessments with only one station, even if referred to as an OSCE. Sources were selected based on predefined inclusion and exclusion criteria. Titles and abstracts were manually screened to identify relevant studies, followed by a full-text review to confirm

eligibility. Two independent reviewers conducted the screening process, resolving any discrepancies through discussion or consultation with a third reviewer. No software tools were used for screening. The analysis was performed with at least two independent researchers at each step. The report was prepared considering a six-step approach to scoping reviews [15] and the PRISMA-ScR checklist for scoping reviews [16].

Results

Part 1: preliminary bibliometrics search

Considering that "Objective Structured Clinical Examination" is a specific medical education term and that the acronymous "OSCE" is also used in other research fields, a senior researcher (SA) conducted the first preliminary research in WOS, Scopus, PubMed, and Scielo, identifying review articles related to these terms up until February 2024. We found 158 articles and performed a bibliometric analysis (SA, SV). A high proportion of the OSCE reviews in the existing literature at that time were written in English (143/157), with some reviews written in Japanese, German, French, Chinese, Greek, Croatian, and Danish (Appendix 1). None of these reviews were written in Spanish or Portuguese, nor considered Spanish terms used to refer to OSCE, and none included Scielo or LILACS databases, leading to a bias regarding information stemming from Latin America.

Subsequently, using the terms "Objective Structured Clinical Examination" and "ECOE" (the Spanish acronym for *Evaluación Clínica Objetiva Estructurada*), we searched the Scopus database and found 354 articles that included these terms, with authors affiliations from the 33 countries that the United Nations considers as part of South America, Central America, the Caribeann, and The Bahamas (Antigua and Barbuda, Argentina, Bahamas, Barbados, Belice, Bolivia, Brasil, Chile, Colombia, Costa Rica, Cuba, Dominica, Ecuador, El Salvador, Granada, Guatemala, Guyana, Haití, Honduras, Jamaica, Nicaragua, Panamá, Paraguay, Perú, República Dominicana, Saint Kitts and Nevis, Santa Lucía, San Vicente and the Granadinas, Suriname, Trinidad and Tobago, Uruguay, Venezuela), considering Mexico as the only North American Spanish-speaking country included in the Latin American region [17].

Part 2 scoping review

Step 1: identifying the research question

Considering the preliminary search of the literature performed in the first part, we recognized that a scoping review on the topic including Spanish terms and Latin America databases has not already been conducted; the number of articles collected from our region made it possible to define a broad and feasible question to analyze all

the literature in the topic during the defined period; and there is sufficient literature to warrant a scoping review.

The research question defined was: *How it has been implemented and how OSCE quality has evolved in Latin America 50 years after its creation?*
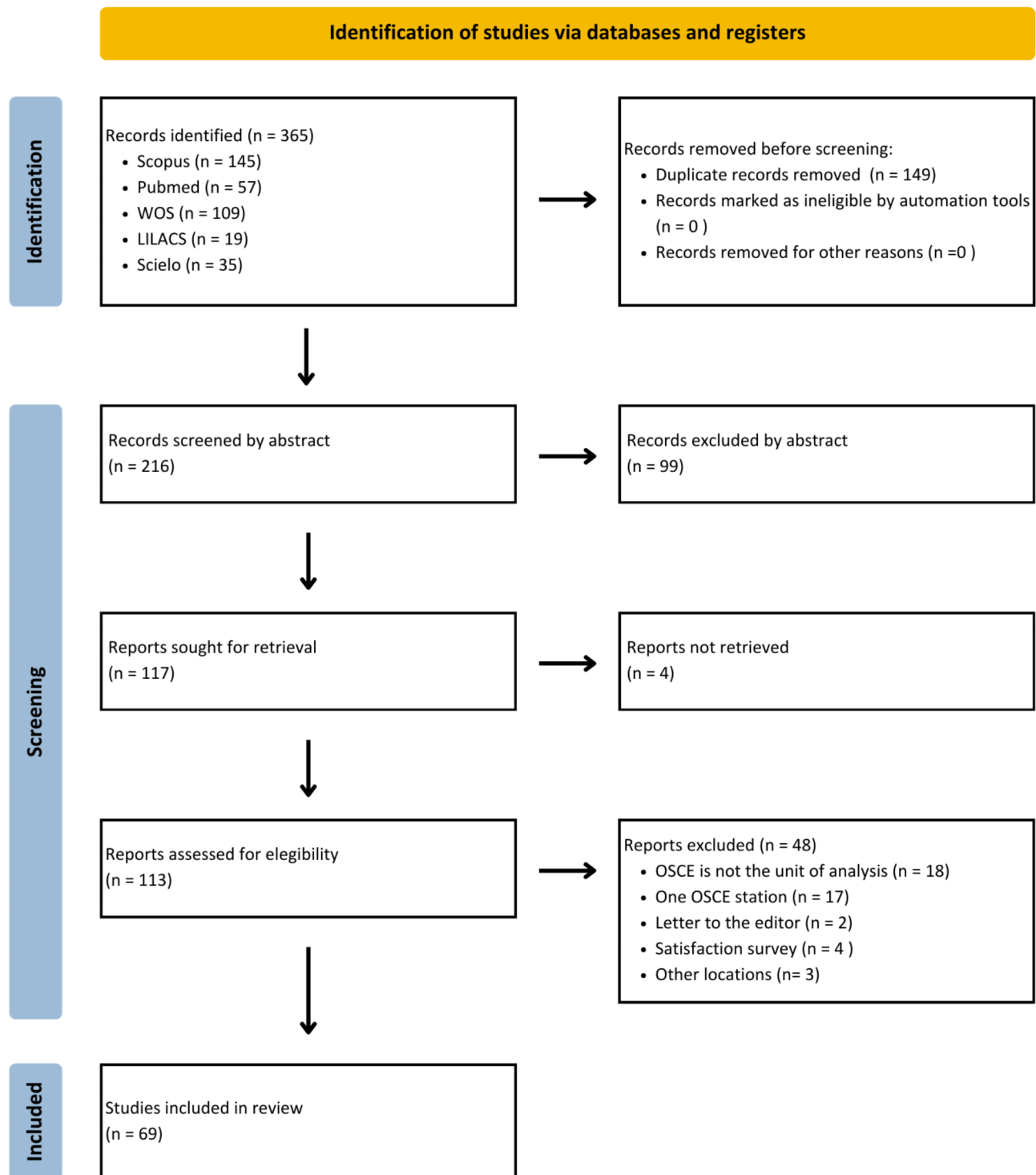
Step 2: identifying relevant studies

After three meetings with a librarian who assisted the team in refining the search strategy, we aimed to control for terms that introduced excessive "noise", and considering that the field has specific expressions not covered by Medical Subject Headings, we decided to use "Objective Structured Clinical Examination" or "Structured clinical examinations", or "Examen Clínico Objetivo Estructurado" OR "ECOE" OR "Exámenes Clínicos Estructurados" as keywords (Appendix 2). We searched the databases SCOPUS, WOS, Scielo, and PubMed without language filters, considering that the team could analyze all official Latin American languages. We did not register the review because PROSPERO does not accept scoping reviews, literature reviews or mapping reviews (Statement extracted from the website https://www.crd.york.ac.uk/prospero/#registernew).

Step 3: selecting studies to be included in the review

After collecting the citations from the search (145 from Scopus, 57 from PubMed, 109 from WOS, 19 from LILACS, and 35 from Scielo), we consolidated the data from the 365 reports data using Excel. One researcher (SA) manually removed the duplicates (149), resulting in a corpus of 216 abstracts for the initial review. These 216 titles and abstracts were reviewed and filtered by an expert (SA). Through this process, 99 articles related to other topics were excluded. Finally, a corpus of 117 articles were read in depth by two reviewers independently, and the differences was decided by a third reviewer (Fig. 1).

Step 4: charting the data

The team collaboratively developed the data extraction form (SA, SB, NK, HS, BF, SV), defining inclusion and exclusion criteria and the extraction procedure. The main extraction categories were: author, title, name of the institution, undergraduate or postgraduate, career, training level (year during the career), OSCE quality elements (instruments used to assess, standardized patient training process, evaluator training, time per station, total OSCE time, types of stations implemented, number of evaluated participants/evaluators/simulators/simulated patients, video recordings, feedback or debriefing, qualification criteria, type of postimplementation analysis, remediation for students). Concerning the type of stations implemented, we consider inactive stations to refer to those in

Identification of studies via databases and registers

Identification

Records identified (n = 365)
- Scopus (n = 145)
- Pubmed (n = 57)
- WOS (n = 109)
- LILACS (n = 19)
- Scielo (n = 35)

Records removed before screening:
- Duplicate records removed (n = 149)
- Records marked as ineligible by automation tools (n = 0)
- Records removed for other reasons (n =0)

Screening

Records screened by abstract (n = 216)

Records excluded by abstract (n = 99)

Reports sought for retrieval (n = 117)

Reports not retrieved (n = 4)

Reports assessed for elegibility (n = 113)

Reports excluded (n = 48)
- OSCE is not the unit of analysis (n = 18)
- One OSCE station (n = 17)
- Letter to the editor (n = 2)
- Satisfaction survey (n = 4)
- Other locations (n= 3)

Included

Studies included in review (n = 69)

*All the records were revised and excluded by a human.
Source: Page MJ, et al. BMJ 2021;372:n71. doi: 10.1136/bmj.n71.

Fig. 1 PRISMA flowchart OSCE Latam

which there is no direct interaction with a standardized patient (SP) or simulator.

After identifying eligible full-text articles, each was rigorously assessed for relevance and methodological integrity in accordance with predefined inclusion and exclusion criteria. Each study was evaluated for its relevance to the research objectives, methodological quality, and completeness of reported data. Following this process, 48 additional articles were excluded. The two main causes of exclusion were because OSCE was not the unit of analysis ($n = 18$) and because the term OSCE was used to one simulated station exam ($n = 17$). The final analysis included 69 articles (Fig. 1).

The Modified Medical Education Research Study Quality Instrument (MMERSQI) score was calculated based on the methodology reported by the authors [18]. Three researchers (SA, SB, NK) performed pilot testing of the extraction MMERSQI form on a small sample of five articles and sustained a meeting afterwards, concluding that the form was appropriate for use.

Step 5: collating, summarizing, and reporting the results

The research team conducted numerical and thematic analyses with the data extracted from all papers. To analyze and report on the evolution of their implementation quality, we organized the OSCE articles into five-year intervals, from 2000 to 2024.

Bibliometric characterization of the selected papers

Most of the selected articles come from the Scopus database, with 56 records representing 79.71% of the total. LILACS and Scielo contributed with 4 and 5 entries respectively, representing 13.05% of the total when combined. The predominant language of the entries is English (55.07%) followed closely by Spanish (43.48%) (Appendix 3).

The first article concerning OSCE in Latin America dates to 2000. Each quinquennial period shows an increasing volume of relevant publications, from 2000 to 2024 (Fig. 2).

Descriptive characteristics of the programs using OSCE

The countries with the highest concentration of published articles in the region are Chile (34.78%), Mexico (28.99%) and Brazil (20.29%). The data from the World Bank Group [19] was used to graph the tendency of OSCE publications accumulated by country against the per capita income during 2023 (Fig. 3).
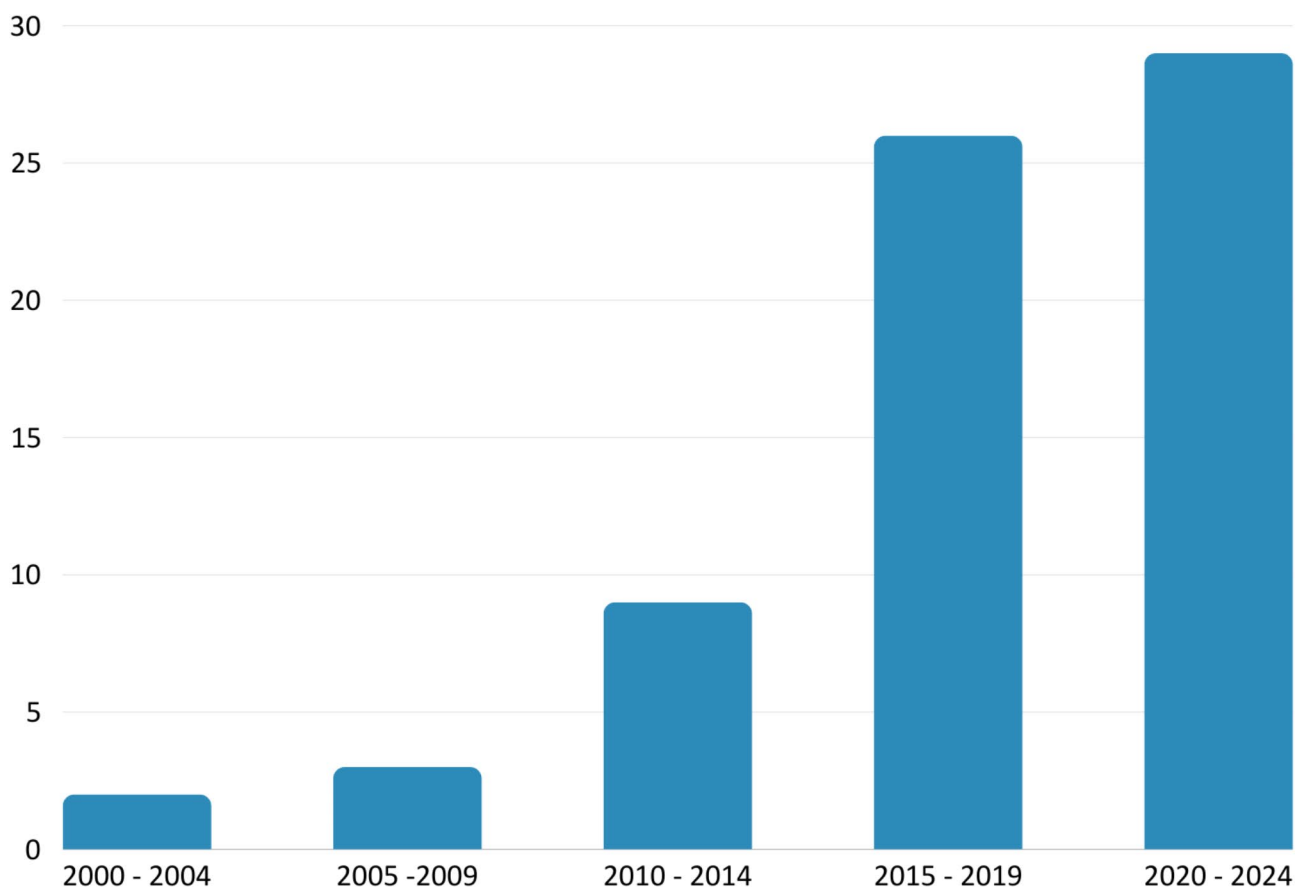
Fig. 2 OSCE articles published from Latam (trend 2000 to 2024)
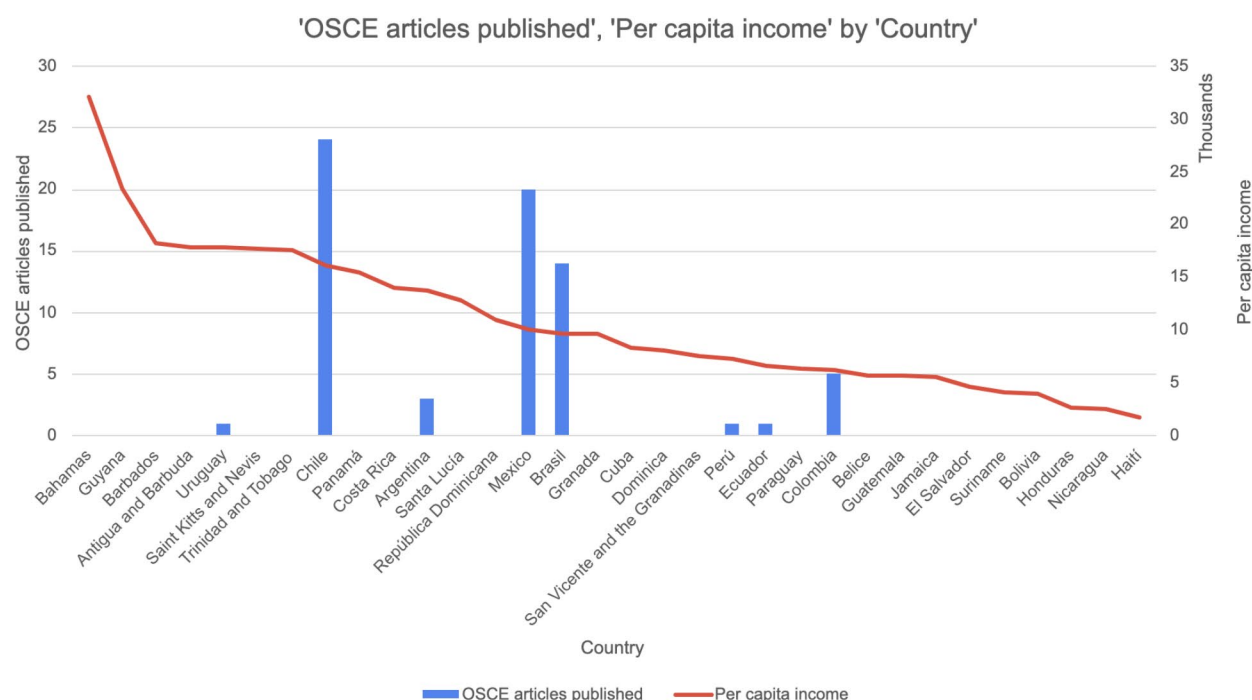
Fig. 3 OSCE articles and Per capita income 2023

Table 1 Distribution of OSCE articles among health careers in Latin America (2000–2024)

Country	Career						Total
	Medicine	Nurse	Physiotherapy	Dentistry	Speach Therapy	Unidentified	
Chile	17	1	2	3	1	0	24
Mexico	18	0	0	2	0	0	20
Brazil	10	1	1	1	0	0	14
Colombia	3	0	2	0	0	0	5
Argentina	3	0	0	0	0	0	3
Uruguay	0	0	0	0	0	1	1
Peru	1	0	0	0	0	0	1
Ecuador	1	0	0	0	0	0	1
Total	53	2	5	6	1	1	69

The data highlights that the medical program conducts the highest number of OSCEs (69.57%), followed by dentistry (8.7%) and physiotherapy (7.25%). On the other hand, nursing and speech therapy have the lowest percentage of publications, with 2.9% and 1.45%. Among the total OSCEs carried out by the medical career, 23.19% are in postgraduate programs. Additionally, it is important to note that among Latin American countries, Chile and Brazil have implemented the OSCE transversally in different health careers (Table 1).

Regarding the purpose of the OSCE, most were summative, followed by formative, and finally, diagnostic examinations. The most frequently assessed competencies are history taking, diagnosis, treatment, and communication, appearing in about 70% of the reports. Physical examination and procedural skills are included in less than 60% of the reports (Table 2).

The number of participants in the OSCE analyzed ranged from 9 to 5399. The number of stations varies between 2 and 25, being more frequently reported circuits with 18 stations (14.49%) followed by 6 stations (13.04%) and five stations (11.59%). The time spent at each station ranged from 4 to 25 min, with 6-minute stations being the most common (17,39%) followed by 8-minute stations (11,59%). The station time was not reported in 24,64% of the articles analyzed. Most reports referred to a single-circuit OSCE (46.38%). In 31,88% of the articles analyzed, the number of circuits was not reported (Table 2).

Among the 34 summative OSCE articles, 22 used instruments without describing validation evidence, while 12 used validated instruments. Reports on the reliability of the OSCE were presented in nearly 20% of the summative and formative OSCEs, showing a wide range

Table 2 Descriptives of OSCE implementation (2000–2024)

		n	%
Academic level of the program	Undergraduate	47	68.12
	Postgraduate	16	23.19
	Other	6	8.70
Intentionality of the assessment	Summative	34	49.28
	Formative	17	24.64
	Diagnostic	9	13.04
Competencies assessed	Diagnostic	50	72.46
	Treatment	49	71.01
	Communication	49	71.01
	History taking	48	69.57
	Physical Exam	41	59.42
	Procedures	40	57.97
Assessment tools used	Checklist	43	60.56
	Likert scales	19	26.76
	Rubric	16	22.54
	Specific Tools	11	15.49
Number of stations	2	1	1.45
	3	4	5.8
	4	5	7.25
	5	8	11.59
	6	9	13.04
	7	4	5.8
	8	5	7.25
	9	3	4.35
	10	4	5.8
	11	1	1.45
	12	3	4.35
	13	1	1.45
	15	5	7.25
	18	10	14.49
	20	1	1.45
	25	1	1.45
	Not reported	4	5.8
Time for each station (min)	4	1	1.45
	5	5	7.25
	6	12	17.39
	7	6	8.7
	8	8	11.59
	9	1	1.45
	10	4	5.8
	12	3	4.35
	15	5	7.25
	20	5	7.25
	25	2	2.9
	Not reported	17	24.64
Number of parallel circuits	1	32	46.38
	2	8	11.59
	3	2	2.9
	5	3	4.35
	6	1	1.45
	9	1	1.45
	Not reported	22	31.88

Table 3 Simulated patients training reports

		n	%
Responsible of the training	Clinician	11	15.94
	Simulation educator	9	13.04
	Multiple trainers	4	5.8
	Specialist in simulated patients	1	1.45
	Simulated patient	1	1.45
	Not reported	43	62.32
Psychosocial role training	No	14	20.29
	Yes	13	18.84
	Not reported	42	61.00
Biological-clinical case training	No	8	11.59
	Yes	28	40.58
	Not reported	33	48.00
SP roles reported	Portray the case	41	59.00
	Multiple roles	7	10.14
	Assess students	6	8.70
	Provide feedback	2	2.90
	Not reported	12	17.39

Table 4 Cronbach´s alpha reported in summative and formative OSCE (2000–2024)

	Summative		Formative	
Cronbach´s alpha	Yes	%	Yes	%
< 0.5	1	2.86	1	5.88
0.5–0.59	1	2.86	0	0
0.6–0.69	4	11.43	1	5.88
0.7–0.79	2	5.71	2	11.76
0.8–0.89	2	5.71	1	5.88
0.9-1	2	5.71	1	5.88
Not reported	23	65.71	11	64.71
Total	35	100	17	100

of values, including values < 0.5, and with no description of the sample's normality (Table 3). The standard-setting process most used was based on the judge's criteria in 40.58% of the reports, with 33,33% of the reports including a description of borderline methods.

The responsibility of the training process of the simulated patients was reported in 37.68% of the articles, mainly guided by clinicians or simulation educators. The training was mainly focused on biological and clinical cases (40.58%), with psychosocial role training described in 18.84% of the articles analyzed. Only two articles described the time invested in simulated patient training. The standardized patients participated in assessing students in 8.70% of the cases and provided feedback in 2.90% (Table 4). Evaluator training was described or mentioned in 41 articles (59.42%).

Regarding resources, 26 studies included inactive stations (37.68%), and simulators were used in 23 studies (33.33%). The use of video recordings was reported in 12 studies (17.39%). Safety protocols, including backups and contingency guidelines, were thoroughly

established in 24.64% of the articles reviewed. Feedback sessions, whether individual or group-based, conducted by various stakeholders and directed toward participants, were mentioned in 23 articles (33.33%). Additionally, mechanisms ensuring the transparency of records and outcomes were outlined, encompassing the levels of accessibility to participants and external parties, their intended uses, and safeguarding measures. These mechanisms were described in 4 articles, constituting 5.8% of the total documents analyzed.

The analysis of the articles by field of study indicates that in Medicine, up to 90% of the competencies evaluated were related to any part of clinical reasoning process (clinical history, diagnosis or treatment), followed by communication and physical examination (81.25% each). Checklists were used in 59.38% of the OSCE, global rating scores in 37.5% and rubrics in 31.25% of the OSCEs. In 12.5% of the OSCE in Medicine were used specific assessment tools. Safety protocols were reported in 29% of the articles. High-fidelity clinical simulation was used as a tool to evaluate knowledge (69%) and skills (72.7%). In the two nursing OSCE included the main competency evaluated was procedures, using low-fidelity simulation. In Kinesiology, the competencies evaluated were mainly procedural (100%), along with clinical history, treatment, and communication, but in this profession high-fidelity simulation was used to evaluate skills. In Dentistry, the main competencies evaluated were diagnosis at 81%, followed by clinical history and communication at 66.6% each. High-fidelity simulation was used in 83% of the articles to evaluate knowledge and skills. In Speech Therapy, the competencies evaluated were treatment, care processes, and communication. SP training was led by a simulation educator, emphasizing the psychosocial aspects of their role. Additionally, high-fidelity simulation with simulated patients was used to assess both knowledge and clinical skills of speech therapy students.

Evaluation of methodological quality and trends in Article quality performance (2000–2024)

The appraisal of the methodological quality of the articles is summarized in Table 5, which provides an overview of the mean, standard deviation, and range of MMERSQI score values across five-year periods from 2000 to 2024. Across the entire span from 2000 to 2024, the mean

Table 5 MMERSQI methodological quality appraisal (2000–2024)

Period	Mean MMERSQI Score	Std Dev	Range
2000–2004	41.00	13.44	31.5–50.5
2005–2009	32.50	21.26	12.5–55.0
2010–2014	48.50	15.33	15.5–65.0
2015–2019	51.00	8.53	31.5–62.0
2020–2024	51.50	11.30	23.0–71.0
2000–2024	51.50	11.70	12.5–71.0

value was 51.50, with a standard deviation of 11.70 and a range from 12.5 to 71.0, reflecting a moderate global variability. When analyzing the data in quinquennials, 2005–2009 shows the poorest performance, with a mean value of 32.50 and a standard deviation of 21.26, indicating greater variability, ranging from 12.5 to 55.0 in the scores. The data indicates an upward trend in mean values over time, with a reduction in variability over the last ten years, and MMERSQI scores exceeding 70 in the most recent quinquennial. This corresponds to two randomized trials published in 2023 [20, 21], both of which examine the effectiveness of virtual and hybrid assessment modalities in health sciences education.

We mapped the articles by quinquennial periods, using findings from thematic analysis, and summarized characteristics of OSCEs quality in each period. We identified the presence or absence of overarching categories: OSCE committee, blueprint, instruments validity or reliability, stations validation, simulated patient training, evaluators training, participants briefing, use of simulators, inactive stations, standard setting methods, post-implementation analysis, post-implementation adjustments, within the content of the identified publications. The themes are presented in Fig. 3, using a color gradient, with five groups, each representing a 20% increment.

First quinquennium 2000–2004

In the period 2000–2004, we identified two articles. The most important quality elements described during this period were the reporting of the blueprint, the simulated patient training process, and the standard-setting methods. The aspects of quality to improve were related to the absence of a description of the validity or reliability of the instruments used inside the OSCE and the high use of inactive stations.

Second quinquennium 2005–2009

Between 2005 and 2009, we identified three articles, with the first description of the reliability of the instruments used within the OSCE. Other quality elements were either described in lower proportion compared to the first period or not reported at all. All reports up to this quinquennial included inactive stations.

Third quinquennium 2010–2014

In the 2010–2014 period, the number of articles tripled compared to the previous timeframe. The most frequently discussed topics were the establishment of an OSCE Committee and post-implementation analysis. At least half of the articles addressed the blueprint, the validity of the instruments used, and the validation of stations, indicating a trend toward strengthening the quality of implementation reporting. However, reports on simulated patient training were limited, and there were

Time Frame	Articles	Category	Quality characteristic and Frequency											
			OSCE Committee	Blueprint	Instruments validity or reliability	Validation of stations	Simulated patient training	Evaluators training	Participants briefing	Use of simulators	Inactive stations	Standard setting methods	Post implementation analysis	Post implementation adequations
2000-2004	2	Yes	50% 1	100% 2	0% 0	50% 1	100% 2	50% 1	50% 1	100% 2	100% 2	100% 2	50% 1	50% 1
		No	50% 1	0% 0	100% 2	50% 1	0% 0	50% 1	50% 1	0% 0	0% 0	0% 0	50% 1	50% 1
2005-2009	3	Yes	33.33% 1	0% 0	33.33% 1	33.33% 1	33.33% 1	33.33% 1	33.33% 1	0% 0	33.33% 1	33.33% 1	33.33% 1	33.33% 1
		No	66.66% 2	100% 3	66.66% 2	66.66% 2	66.66% 2	66.66% 2	66.66% 2	100% 3	66.66% 2	66.66% 2	66.66% 2	66.66% 2
2010-2014	9	Yes	77.78% 7	55.56% 5	55.56% 5	55.56% 5	11.11% 1	44.44% 4	22.22% 2	11.11% 1	22.22% 2	44.44% 4	77.78% 7	0% 0
		No	22.22% 2	44.44% 4	44.44% 4	44.44% 4	88.89% 8	55.56% 5	77.78% 7	88.89% 8	77.78% 7	55.56% 5	22.22% 2	100% 9
2015-2019	26	Yes	46.15% 12	50% 13	26.92% 7	50% 13	53.85% 14	76.92% 20	46.15% 12	42.31% 11	34.62% 9	69.23% 18	38.46% 10	3.85% 1
		No	53.85% 14	50% 13	73.08% 19	50% 13	46.15% 12	23.08% 6	53.85% 14	57.69% 15	65.38% 17	30.77% 8	61.54% 16	96.15% 25
2020-2024	29	Yes	72.41% 21	48.28% 14	44.83% 13	55.17% 16	48.28% 14	51.72% 15	37.93% 11	31.03% 9	41.38% 12	62.07% 18	51.72% 15	6.90% 2
		No	27.59% 8	51.72% 15	55.17% 16	44.83% 13	51.72% 15	48.28% 14	62.07% 18	68.97% 20	58.62% 17	37.93% 11	48.28% 14	93.10% 27

Thematic insight Colour	0-19%	20-39%	40-59%	60-79%	80-100%

Fig. 4 Heat map of articles quality annalyzed by quinquennial

no mentions of post-implementation adjustments or adequations.

Fourth quinquennium 2015–2019
In the 2015–2019 period, we identified 26 articles, a threefold increase compared to the previous period. Positive quality aspects during this time included the detailed description of standard-setting methods and the training provided to evaluators. Although reports on simulated patient training remained limited, they showed an increase compared to the prior period. However, a low percentage of articles continued to address the validity or reliability of the instruments used or postimplementation adequations.

Fifth quinquennium 2020–2024
Between 2020 and 2024, we identified 29 relevant articles. Positive developments during this period included establishing an OSCE Committee and standard-setting methods. However, the percentage of participants receiving briefings remained low, and the blueprint report followed the same trend observed over the past decade. Notably, this five-year period saw the highest percentage of inactive stations recorded in the last 20 years. Data collection for this quinquennial concluded in March of last year. During this time, remote OSCEs were reported in both Medicine and Dentistry, with a notable increase in reports from surgical postgraduates following simulation training (Fig. 4).

Discussion
This review highlights the increasing adoption of OSCE in Latin America, with notable advancements in Chile, Mexico, and Brazil. While methodological rigor has improved, challenges remain, including variability in implementation, limited standardization, and inconsistent reporting of psychometric data. Additionally, the predominance of Scopus-indexed studies and the exclusion of non-English databases may have influenced the perceived trends in OSCE research across the region.

Our results reflect the broader disparities in medical education commonly seen in low- and middle-income countries (LMICs), where economic and infrastructural limitations hinder the timely adoption of modern educational methodologies [19]. Nonetheless, Latin America has demonstrated considerable progress in devising context-specific strategies for implementing OSCE, reflecting the region's resilience and creativity in fostering clinical competence despite limited resources, which is a core theme of this manuscript's objectives.

The analysis highlights key trends in bibliometric patterns (the dominance of Scopus-indexed studies), linguistic distribution (a predominance of publications in English and Spanish), and geographic concentration (with Chile, Brazil, and Mexico leading in research output). Additionally, OSCE implementation in Latin America predominantly focuses on medical education, with an emphasis on assessing clinical and communication skills. As previously discussed, the quality of studies across the region remains variable but has shown a progressive improvement over time, as reflected in an upward trend in methodological rigour, particularly in the most recent quinquennials.

The linguistic distribution of the publications emphasizes the regional focus of the research, with English and Spanish as the predominant languages. While English slightly leads, aligning with the global trend of English as the lingua franca of scientific communication,

the significant number of Spanish-language publications highlights the importance of incorporating non-English literature in comprehensive bibliometric analyses, especially in regions where multiple languages are spoken.

The bibliometric analysis reveals a strong preference for articles sourced from the Scopus database, which accounts for most selected records. This dominance suggests that Scopus is a key repository for research on the OSCE in Latin America, likely due to its extensive coverage and indexing of high-impact journals. Smaller, yet notable, contributions from LILACS and Scielo, each representing 7.04% of the total, highlight the importance of regional databases in capturing research outputs more specific to Latin American contexts, often published in local journals. It is important to acknowledge that the reliance on Scopus may limit the diversity of perspectives and findings, as it may not fully represent the breadth of regional research available in local publications. Thus, while the analysis provides valuable insights, it is constrained by the databases utilized and may not encompass the entirety of OSCE-related research in the region.

The concentration of publications in Chile, Mexico, and Brazil reflects these countries' leadership in advancing OSCE practice in the region. Even though these countries are within the highest development in the region, this fact is not the only explanation for their productivity in academic contributions related to OSCE, as other countries classified as high and medium-high income in 2023 exhibit different patterns in their contributions to the academic community concerning OSCE utilization. Furthermore, the number of OSCE-related articles does not correlate with the population sizes of these three countries; for instance, Chile has a smaller population compared to Brazil and other leading countries yet contributes a significant number of articles. This suggests that additional factors may be influencing the findings of our review.

Moreover, in resource-constrained environments, such as many countries of Latin America, the effective implementation of OSCEs could be hindered by the lack of financial resources and infrastructural capacity. One interesting case between middle-income countries in our region is Colombia, which shows higher productivity concerning OSCE than countries with similar income.

The implementation of OSCEs across various health disciplines in Chile and Brazil, extending beyond their primary application in medical education, further exemplifies their leadership in adopting comprehensive assessment strategies.

The data reveals a strong focus on OSCEs within medical education, with a striking 77.5% of publications addressing this field. The integration of OSCEs into postgraduate medical programs further emphasizes the method's role in assessing clinical competencies at advanced training levels. However, the relatively low representation of other health professions, such as nursing [23, 24], physiotherapy [25, 26, 27, 28, 29], and speech therapy [30], highlights a potential area for growth and development. The limited use of OSCEs in these fields may reflect either a lack of resources or a slower adoption rate, suggesting that targeted initiatives could encourage broader implementation.

All the OSCE analyzed are aligned with Kane validity framework [7], because they are primarily used to test clinical and communication skills, as was the original design intention of Harden for the examination, a principle that continues to hold true according to his later reflection [1]. The purpose statement for the OSCE should underpin all stages of the OSCE process, from design and delivery to data analysis and outcomes [4]. This includes making the purpose explicit to all stakeholders, such as faculty, candidates, examiners, employers, regulatory bodies, and the public [2]. The published studies included in the review fulfil the common purposes for which the methodology is intended.

When constructing the OSCE, sufficient sampling (stations, duration, examiners) must be planned and reflected in the blueprint, as inadequate sampling can undermine the reliability of pass/fail decisions. OSCEs with fewer than 12 stations or less than 150 min of testing time are generally considered unlikely to produce reliable results [22]. Methodologically, the OSCEs reported in the analyzed articles demonstrate a diverse range of structures, with significant variation in the number of stations and the duration of each station. The frequent use of 18-station circuits, followed by 6- and 5-station circuits, suggests a preference for medium-sized assessments, likely balancing the need for comprehensive evaluation with logistical feasibility. However, the fact that 23.94% of the articles did not report station times points highlights a gap in methodological transparency that should be addressed in future studies to ensure comparability and reproducibility.

Additionally, to implement a quality OSCE, an appropriate marking/scoring scheme should be selected to appraise the participants' competencies. While years ago, rating scales were considered more suitable for advanced clinical learners [31], more recent research has shown that a careful combination of rating scales and checklists may be appropriate, especially for technical skills testing [32]. In our study, there is a tendency to include checklists and global rating scores, and some of the studies included in our analysis describe the validity evidence for specific scales appropriate for assessing competencies in a transversal manner, including a scale to assess patient-centred communication by simulated patients [33], a modified Objective Structured Assessment of Technical Skills (OSATS) to assess procedural skills in surgeons

[34] and specific instruments to assess transfusion [35]. Psychometric analyses of the instruments used in the OSCE are crucial for quality assurance, as they assess reliability and station-level issues [5]. To justify academic decisions based on test scores, all sources of validity evidence should be gathered [7, 8, 9].

Standard-setting methods are detailed in one-third of the reports in our final sample, showing improvement in this regard over the past decade. The Borderline Regression Method is currently favoured for standard-setting in OSCEs, even in small cohorts within medicine [36] and dentistry [37], provided there is a suitable combination of global grades and checklist scores, alongside trained evaluators.

Examiner training is considered crucial for OSCE implementation, but its effectiveness is mixed [38]. Rather than striving for uniform examiners, reliability may be enhanced through multiple observations by diverse evaluators. Embracing this variability can be beneficial, as it may arise from design issues in stations and marking schemes, contributing to validity. Recent training efforts increasingly emphasize examiner conduct, behaviours, and the mitigation of conscious and unconscious biases [5]. Descriptions of evaluator training were included in nearly 60% of the studies, though the specific characteristics or effects of such training were not extensively highlighted.

The analysis of the methodological quality of OSCE publications over the 24-year period indicates a progressive improvement in reporting practices and quality measures, particularly in recent years. The upward trend in MMERSQI (Medical Education Research Study Quality Instrument) scores reflects increasing rigour in study design and reporting, with the last quinquennial (2020–2024) showing the highest quality scores. This improvement is particularly evident in the adoption of randomized trials and the increasing attention to reliability and validity assessments, although these aspects still present areas for further enhancement [20–21].

One of the most significant findings is the variability in the quality of OSCE implementation across the region. While some countries and institutions have developed robust systems with thorough training for simulated patients and evaluators, others still lack consistency in these critical areas. The low reporting of participant briefing and the persistent use of inactive stations in many OSCEs suggest ongoing challenges in optimizing the assessment process. Additionally, the limited use of validated instruments and the lack of detailed reliability reports in many studies highlight the need for more standardized approaches to ensure the accuracy and fairness of OSCEs. Ensuring both validity and confidentiality is essential for the OSCE to function effectively as a fair assessment tool and to promote high-quality clinical competence learning and development.

Moreover, the low implementation of safety protocols, including backups and contingency guidelines, in a quarter of the articles reflects the low awareness of the need to mitigate risks associated with high-stakes assessments. In the same sense, the low incidence of feedback mechanisms and transparency measures in the OSCE process highlights an area that requires further attention to enhance the overall educational value and fairness of these examinations. It is important to keep in mind that feedback plays a critical role in motivating learning, and providing meaningful feedback should be a routine practice [39].

The evolution of the quality of OSCEs shows a progressive increase in the number of articles in each quinquennium analyzed, highlighting a growing interest in and the importance of OSCE as an assessment method. There has been a gradual improvement in reporting on reliability, blueprinting, and validity, but these areas still need further emphasis in many studies.

Simulated patient training is a crucial step in achieving valid and reliable OSCEs, yet this aspect has not been consistently well-reported, suggesting it may be under-prioritized despite its importance for realistic assessments. The continued presence and eventual increase in inactive stations is a concern, as these stations reduce the overall engagement and value of the examination experience.

The shift toward remote OSCEs, particularly during the pandemic, represents a significant adaptation in the field, though it also introduces new challenges regarding standardization and validity of these examinations.

This review provides critical insights into the implementation of OSCE in Latin America, addressing a significant gap in medical education research. As the first comprehensive analysis of this topic in the region, it highlights the increasing adoption of OSCE while identifying persistent challenges such as variability in implementation, limited standardization, and inconsistencies in psychometric reporting.

By applying no linguistic restrictions, this review ensures a broad and inclusive selection of studies published in English, Spanish, and Portuguese across multiple databases (Scopus, LILACS, SciELO), thereby capturing diverse regional perspectives. The study adheres to PRISMA-ScR guidelines, ensuring methodological rigor through systematic screening, data extraction, and quality assessment using the MMERSQI tool. Moreover, it identifies key bibliometric, geographic, and disciplinary trends, highlighting improvements in OSCE quality and standardization while offering practical recommendations to advance clinical competence assessment in resource-limited settings.

A fundamental contribution of this review is its emphasis on strengthening faculty development programmes to improve examiner training and ensure more reliable and standardized assessments. Additionally, it identifies the pressing need for greater investment in simulation infrastructure and the integration of OSCE into national accreditation frameworks, facilitating its widespread and sustainable adoption across the region.

By addressing these critical areas, this review provides a foundation for evidence-based policy development and targeted interventions that support the equitable and effective implementation of OSCEs in diverse educational contexts.

These findings emphasize the need for regional policies that support OSCE standardization across institutions. Strengthening faculty development programs, particularly in examiner training and psychometric validation, could enhance assessment reliability. Encouraging cross-institutional collaborations for shared OSCE blueprints and standard-setting protocols may also help address inconsistencies in station design and evaluation criteria.

Despite its broad scope, this review has certain limitations. Its reliance on selected databases may have led to the exclusion of grey literature and non-indexed regional publications, thus limiting the scope of available evidence. While no linguistic restrictions were applied, variability in reporting quality—particularly concerning station design, examiner training, and psychometric validation—posed challenges in the analysis of OSCE methodologies.

Furthermore, the heterogeneity in implementation across countries complicates direct comparisons, requiring caution when generalizing findings. Additionally, as the selection process was based on indexing and availability, unpublished institutional OSCE experiences may have been overlooked, affecting the completeness of the regional landscape.

Our findings align with global trends in OSCE implementation, confirming its predominant use in medical education and its focus on clinical and communication skills assessment. The observed increase in methodological rigor is consistent with international improvements in OSCE standardization and quality assurance. However, as the first OSCE review focused on Latin America, this study highlights unique regional challenges, including linguistic barriers, resource constraints, and disparities in implementation. These aspects are less frequently addressed in studies from high-income settings, suggesting that while Latin America follows global OSCE trends, specific contextual adaptations are required. Future research should further investigate these challenges to support the equitable and effective adoption of OSCEs across diverse educational contexts.

Conclusion

Despite the significant progress in the implementation of the OSCEs within medical education, significant gaps remain that warrant urgent attention. There is an evident need for greater standardization and broader adoption of OSCEs across various health disciplines. Although advancements in organizational structures, such as the establishment of OSCE committees, and slow but steady improvements in instrument validation have been noted, challenges persist, including the prevalence of inactive stations, inadequate training for simulated patients, and insufficient reporting on the reliability and validity of assessment instruments.

Future research must prioritize addressing the gaps identified in this analysis, especially in areas such as instrument validation, participant briefing, and the consistent application of quality measures. By building on the advances of recent years, Latin American institutions can enhance the reliability, validity, and educational impact of OSCEs, making a valuable contribution to the global dialogue on best practices in health professions education. Addressing these gaps is crucial for enhancing the effectiveness of OSCEs in resource-limited settings and advancing health professional education across the region.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12909-025-07151-5.

Supplementary Material 1

Data availability
All data generated or analyzed during this study are included in this published article [and its supplementary information files].

Declarations

Ethics approval and consent to participate
Not applicable.

Author details
[1]Facultad de Medicina, Universidad del Desarrollo, Las Condes, Chile
[2]Unidad de Simulación e Innovación en Salud, Universidad San Sebastián, Los Ríos, Chile
[3]Residente Hematología y Hemoterapia Hospital Universitario de Basurto, Bilbao, España
[4]Universidad Peruana Cayetano Heredia , Lima, Peru
[5]Departamento Medicina Interna Occidente, Facultad de Medicina, Universidad de Chile, Santiago, Chile
[6]Universidad de La Serena, La Serena, Chile
[7]Hospital Nacional Cayetano Heredia, Lima, Perú
[8]Hospital Nacional Arzobispo Loayza, Lima, Perú
[9]Unidad Institucional de Simulación, Universidad Peruana Cayetano Heredia, Lima, Perú

References
1. Harden RM. Revisiting 'Assessment of clinical competence using an objective structured clinical examination (OSCE)'. Med Educ. 2016;50(4):376–9. https://doi.org/10.1111/medu.12801. PMID: 26995470.
2. Boursicot K, Roberts T, Burdick W. Structured assessments of clinical competence. Understanding medical education. London: Wiley; 2018. pp. 335–45.
3. Khan KZ, Gaunt K, Ramachandran S, Pushkar P. The Objective Structured Clinical Examination (OSCE): AMEE Guide No. 81. Part ll: organisation & administration. Med Teach. 2013;35(9): e1447–63. doi: 10.3109/0142159X.2013.818635. PMID: 23968324.
4. Daniels VJ, Pugh D. Twelve tips for developing an OSCE that measures what you want. Med Teach. 2018;40(12):1208–13. Epub 2017 Oct 25. PMID: 29069965.
5. Pell G, Fuller R, Homer M, Roberts T, International Association for Medical Education. How to measure the quality of the OSCE: A review of metrics - AMEE guide no. 49. Med Teach. 2010;32(10):802–11. doi: 10.3109/0142159X.2010.507716. PMID: 20854155.
6. Mitchell ML, Henderson A, Groves M, Dalton M, Nulty D. The objective structured clinical examination (OSCE): optimising its value in the undergraduate nursing curriculum. Nurse Educ Today. 2009;29(4):398–404. https://doi.org/10.1016/j.nedt.2008.10.007. Epub 2008 Dec 3. PMID: 19056152.
7. Kane MT. Validating the interpretations and uses of test scores. J Educ Meas. 2013;50(1):1–73. https://doi.org/10.1111/jedm.12000.
8. Boursicot K, Kemp S, Wilkinson T, Findyartini A, Canning C, Cilliers F, Fuller R. Performance assessment: Consensus statement and recommendations from the 2020 Ottawa Conference. Med Teach. 2021;43(1):58–67. doi: 10.1080/0142159X.2020.1830052. Epub 2020 Oct 14. PMID: 33054524.
9. Downing SM. Validity: on meaningful interpretation of assessment data. Med Educ. 2003;37(9):830–7. https://doi.org/10.1046/j.1365-2923.2003.01594.x. PMID: 14506816.
10. Raymond MR, Grande JP. A practical guide to test Blueprinting. Med Teach. 2019;41(8):854–61. https://doi.org/10.1080/0142159X.2019.1595556. Epub 2019 Apr 24. PMID: 31017518.
11. Kogan JR, Conforti L, Bernabeo E, Iobst W, Holmboe E. Opening the black box of clinical skills assessment via observation: a conceptual model. Med Educ. 2011;45(10):1048–60. https://doi.org/10.1111/j.1365-2923.2011.04025.x. PMID: 21916943.
12. Foy JP, Serresse L, Decavèle M, Allaire M, Nathan N, Renaud MC, Sabourdin N, Souala-Chalet Y, Tamzali Y, Taytard J, Tran M, Cohen F, Bottemanne H, Monsel A. Clues for improvement of research in objective structured clinical examination. Med Educ Online. 2024;29(1):2370617. Epub 2024 Jun 27. PMID: 38934534; PMCID: PMC11212575.
13. Troncon LEA, Rodrigues MLV, Piccinato CE, Figueiredo JFC, Peres LC, Cianflone ARL. Overcoming difficulties in the introduction of a summative assessment of clinical competence in a Brazilian medical school. In: Scherpbier AJJA, van der Vleuten CPM, Rethans JJ, van der Steeg AFW, editors. Advances in

medical education. Dordrecht: Springer; 1997. https://doi.org/10.1007/978-94-011-4886-3_58.
14. Bustamante M, Carvajal C, Gottlieb B, Contreras JE, Uribe M, Melkonian E, Cárdenas P, Amadori A, Parra JA. Hacia Un Nuevo instrumento de evaluación En La Carrera de medicina. Uso Del Método OSCE [A new instrument for the evaluation of the medical profession. Use of the OSCE method]. Rev Med Chil. 2000;128(9):1039–44. Spanish. PMID: 11349493.
15. Mak S, Thomas A. Steps for Conducting a Scoping Review. J Grad Med Educ. 2022;14(5):565–567. doi: 10.4300/JGME-D–22–00621.1. PMID: 36274762; PMCID: PMC9580325.
16. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, Shamseer L, Tetzlaff JM, Akl EA, Brennan SE, Chou R, Glanville J, Grimshaw JM, Hróbjartsson A, Lalu MM, Li T, Loder EW, Mayo-Wilson E, McDonald S, McGuinness LA, Stewart LA, Thomas J, Tricco AC, Welch VA, Whiting P, Moher D. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 2021;372:n71. https://doi.org/10.1136/bmj.n71. PMID: 33782057; PMCID: PMC8005924.
17. Naciones, Unidas. Grupos regionales de estados miembros de las Naciones Unidas. [internet] [cited 2025 Feb 15]. Available from: https://www.un.org/dgacm/es/content/regional-groups
18. Al Asmri M, Haque MS, Parle JA, Modified Medical Education Research Study Quality. Instrument (MMERSQI) developed by Delphi consensus. BMC Med Educ. 2023;23(1):63. https://doi.org/10.1186/s12909-023-04033-6.
19. World Bank Grouo [Internet], Washington DC. USA; [Cited October 7 2024]. Retrieved from https://datos.bancomundial.org/indicador/NY.GDP.PCAP.CD?locations=ZJ
20. Martinez FT, Soto JP, Valenzuela D, González N, Corsi J, Sepúlveda P. Virtual clinical simulation for training amongst undergraduate medical students: A pilot randomised trial (VIRTUE-Pilot). Cureus. 2023;15(10):e47527. https://doi.org/10.7759/cureus.47527.
21. Porto FR, Ribeiro MA, Ferreira LA, Oliveira RG, Devito KL. In-person and virtual assessment of oral radiology skills and competences by the objective structured clinical examination. J Dent Educ. 2023;87(4):505–13. https://doi.org/10.1002/jdd.13138.
22. Swanson DB, Clauser BE, Case SM. Clinical skills assessment with standardized patients in High-Stakes tests: A framework for thinking about score precision, equating, and security. Adv Health Sci Educ Theory Pract. 1999;4(1):67–106. https://doi.org/10.1023/A:1009862220473. PMID: 12386436.
23. Patricia BAS, Wright Navarrete Ana Cecilia. Adaptación interactiva computacional de un examen clínico objetivo estructurado para Enfermería. Educ Med Super [Internet]. 2014 Dic [citado 2024 Sep 21]; 28(4): 667–676. Disponible en: http://scielo.sld.cu/scielo.php?script=sci_arttext%26;pid=S0864–21412014000400006%26;lng=es
24. Costa LCSD, Avelino CCV, Freitas LA, Agostinho AAM, Andrade MBT, Goyatá SLT. Undergraduates performance on vaccine administration in simulated scenario. Rev Bras Enferm. 2019 Mar-Apr;72(2):345–53. https://doi.org/10.1590/0034-7167-2018-0486. English, Portuguese.
25. Silva CC, Lunardi AC, Mendes FA, Souza FF, Carvalho CR. Objective structured clinical evaluation as an assessment method for undergraduate chest physical therapy students: a cross-sectional study. Rev Bras Fisioter. 2011 Nov-Dec;15(6):481–6. Epub 2011 Nov 21. PMID: 22094547.
26. de la Barra-Ortiz HA, Gómez-Miranda LA, de la Fuente-Astroza JI. Objective structured clinical examination (OSCE) to assess the clinical skills of physical therapy students when using physical agents. Revista De La Facultad De Med. 2021;69(3):55–63.
27. Cobo-Mejía EA, Sandoval-Cuellar C, Villarraga-Nieto ADP, Alfonso-Mora ML, Castellanos-Garrido AL, Acosta-Otálora ML, Rondón-Villamil YA, Goyeneche-Ortegón RL, Castellanos-Vega RDP. Validity and reliability of an Osce for clinical reasoning in physiotherapy. Turk J Physiother Rehabil. 2022;33(1):11–5. https://doi.org/10.21653/tjpr.839006.
28. de la Barra-Ortiz HA, Gómez-Miranda LA, de la Fuente-Astroza JI. Nivel de satisfacción y correlación Entre El Desempeño y La autoevaluación de Los estudiantes de fisioterapia En El examen clínico objetivo estructurado (ECOE) al utilizar agentes Físicos. Revista De La Facultad De Med. 2022;70(3):31–57.
29. Figueroa-Arce N, Figueroa-González P, Gómez-Miranda L, Gútierrez-Arias R, Contreras-Pizarro V. Implementación de Un examen clínico objetivo estructurado (ECOE) Como Herramienta Para evaluar El desarrollo Del Razonamiento clínico En estudiantes de fisioterapia. Revista De La Facultad De Med. 2022;70(2):53–65.
30. Bustos M, Arancibia C, Muñoz N, Azócar J. La simulación clínica En Atención primaria de Salud En Contexto de docencia: Una experiencia Con estudiantes de fonoaudiología. Rev Chil Fonoaudiol [Internet] 22 de noviembre de 2018

[citado 22 de septiembre de 2024];17:1–14. Disponible en: https://revfono.uc hile.cl/index.php/RCDF/article/view/51599

31. Ilgen JS, Ma IW, Hatala R, Cook DA. A systematic review of validity evidence for checklists versus global rating scales in simulation-based assessment. Med Educ. 2015;49(2):161–73. https://doi.org/10.1111/medu.12621. PMID: 25626747.

32. Wood TJ, Pugh D. Are rating scales really better than checklists for measuring increasing levels of expertise? Med Teach. 2020;42(1):46–51. Epub 2019 Aug 20. PMID: 31429366.

33. Armijo-Rivera S, Behrens CC, Giaconi ME, Hurtado AS, Fernandez MR, Parra P, Morales MV, Makoul G. Validation of the Spanish version of a patient-centered communication assessment instrument in osces|Validación de La versión En Español de Un instrumento de evaluación de La Comunicación Centrada En El Paciente En OSCE. Educ Med. 2021;22(4):193–8.

34. Ortiz C, Vela J, Contreras C, Belmar F, Paul I, Zinco A, Ramos JP, Ottolino P, Achurra P, Jarufe N, Alseidi A, Varas J. A new approach for the acquisition of trauma surgical skills: an OSCE type of simulation training program. Surg Endosc. 2022;36(11):8441–50. Epub 2022 Mar 2. PMID: 35237901; PMCID: PMC8890468.

35. Calderón MJM, Pérez SIA, Becerra N, Suarez JD. Validation of an instrument for the evaluation of exchange transfusion (INEXTUS) via an OSCE. BMC Med Educ. 2022;22(1):480. https://doi.org/10.1186/s12909-022-03546-w. PMID: 35725443; PMCID: PMC9210713.

36. Homer M, Fuller R, Hallam J, Pell G. Setting defensible standards in small cohort osces: Understanding better when borderline regression can 'work'. Med Teach. 2020;42(3):306–15. Epub 2019 Oct 26. PMID: 31657266.

37. Moreno-López R, Hope D. Can borderline regression method be used to standard set osces in small cohorts? Eur J Dent Educ. 2022;26(4):686–91. https://doi.org/10.1111/eje.12747. Epub 2022 Jan 7. PMID: 34921711.

38. Yeates P, Maluf A, McCray G, Kinston R, Cope N, Cullen K, O'Neill V, Cole A, Chung CW, Goodfellow R, Vallender R, Ensaff S, Goddard-Fuller R, McKinley R. Inter-school variations in the standard of examiners' graduation-level OSCE judgements. Med Teach. 2024 Jul 8:1–9. doi: 10.1080/0142159X.2024.2372087. Epub ahead of print. PMID: 38976711.

39. Ossenberg C, Henderson A, Mitchell M. What attributes guide best practice for effective feedback? A scoping review. Adv Health Sci Educ Theory Pract. 2019;24(2):383–401. https://doi.org/10.1007/s10459-018-9854-x. Epub 2018 Oct 3. PMID: 30284067.

Publisher's note