

## Exploring the Methods of Assessment of General Medical Students

Neda Negahban Jouzan( MSc)<sup>1</sup> , Hossein Karimi Moonaghi(PhD)<sup>2,3,\*</sup> , Hoorak Poorzand(PhD)<sup>4</sup> ,

Mohammad Khajedaluee(PhD)<sup>5</sup> 

1. Medical Education Department, Faculty of Medicine, Mashhad University of Medical Sciences, Mashhad, Iran.
2. Medical Science Education Research Center, Mashhad University of Medical Sciences, Mashhad, Iran.
3. Nursing and Midwifery Care Research Center, Mashhad University of Medical Sciences, Mashhad, Iran.
4. Department of Cardiology, Faculty of Medicine, Mashhad University of Medical Sciences, Mashhad, Iran.
5. Department of Social Medicine, Faculty of Medicine, Mashhad University of Medical Sciences, Mashhad, Iran.

### ARTICLE INFO

#### Article type

Review article

#### Article history

Received: 18.12.2023

Accepted: 13.05.2024

#### Keywords

Assessment,  
Medical,  
Student.



[10.22038/hmed.2024.66957.1304](https://doi.org/10.22038/hmed.2024.66957.1304)

### ABSTRACT

**Introduction:** The review of published articles shows that medical schools in Iran use similar methods and formats for assessment. While assessment is considered an important pillar in educational planning and a valuable step for the training of expert and committed doctors. The aim of the study is to examine the assessment methods in terms of evaluation formats, Miller's evaluation levels and implementation capability.

**Materials & Methods:** The required information was started by searching in Persian and international databases, magazines, curriculum of prestigious universities in the world in the last 20 years alongside with virtual conferences AMEE international virtual conferences in August 2021 and the summary booklet of medical education articles of Shahid Motahari 2021 (the 22nd national conference of medical education) were reviewed. Data analysis was done by Beredy's adaptive model. The results were provided to assessment experts during a focused group meeting for further review

**Results:** According to the extracted data, the approach of assessment is towards formative assessment format and improvement of traditional methods along with modern methods, which was clearly observed in the study of the curriculum of Harvard University in America and Oxford University in England. Opinions of experts were collected about the appropriateness of assessment methods with evaluation formats and levels of Miller's pyramid, the solution and the implementation conditions. The majority of opinions and suggestions of experts were related to the change in the way of executive policies of universities and providing a context for the emergence of new idea.

**Conclusion:** The data analysis led to the identification of the evolution of suitable methods, formats of assessment and providing educational-executive solutions. It is suggested that the obtained results be used to develop assessment models according to Miller's assessment levels.

Cite this paper as:

Negahban Jouzan N, Karimi Moonaghi H, Poorzand H, Khajedaluee M. Exploring the Methods of Assessment of General Medical Students . Horizon of Medical Education Development. 2025;16(3):119-129

\* Corresponding author: Hossein Karimi Moonaghi

Email: [karimih@mums.ac.ir](mailto:karimih@mums.ac.ir)

Address: Faculty of Nursing and midwifery, Kharazmi Complex, Mashhad.



## واکاوی روش‌های سنجش دانشجویان رشته پزشکی عمومی

ندا نگهبان جوزان (MSc)<sup>۱</sup>، حسین کریمی مونی (PhD)<sup>۲,۳\*</sup>، هورک پورزند (PhD)<sup>۴</sup>، محمد خواجه دلویی (PhD)<sup>۵</sup>

۱. گروه آموزش پزشکی، دانشکده پزشکی، دانشگاه علوم پزشکی مشهد، مشهد، ایران.
۲. مرکز تحقیقات آموزش علوم پزشکی، دانشگاه علوم پزشکی مشهد، مشهد، ایران.
۳. مرکز تحقیقات مراقبت‌های پرستاری و مامایی، دانشگاه علوم پزشکی مشهد، مشهد، ایران.
۴. گروه قلب و عروق، دانشکده پزشکی، دانشگاه علوم پزشکی مشهد، مشهد، ایران.
۵. گروه پزشکی اجتماعی، دانشکده پزشکی، دانشگاه علوم پزشکی مشهد، مشهد، ایران.

| مشخصات مقاله                                                               | چکیده                                                                                                                                                                                                                                                                                                                                                                                                                                   |
|----------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>نوع مقاله</b><br>مقاله مروری                                            | <b>مقدمه:</b> بررسی مقالات منتشره نشان می‌دهد دانشکده‌های پزشکی در ایران، روش‌ها و قالب‌های مشابهی برای سنجش استفاده می‌کنند. این در حالیست که سنجش رکنی مهم در برنامه‌ریزی آموزشی و گامی ارزشمند برای تربیت پزشکانی خبره و متعهد به شمار می‌رود. هدف مطالعه، بررسی روش‌های سنجش از نظر قالب‌های ارزیابی، سطوح ارزیابی، میلر و قابلیت اجرایی می‌باشد.                                                                                   |
| <b>پیشینه پژوهش</b><br>تاریخ دریافت: ۱۴۰۲/۰۹/۲۷<br>تاریخ پذیرش: ۱۴۰۳/۰۲/۲۴ | <b>روش کار:</b> اطلاعات مورد نیاز با جستجو در پایگاه‌های داده فارسی و بین‌المللی، مجلات، کوریکولوم دانشگاه‌های معتبر دنیا در ۲۰ سال اخیر و کنفرانس‌های مجازی (AMEE ۲۰۲۱ و شهید مطهری ۱۴۰۰) آغاز و برای مقایسه از الگوی تطبیقی Beredy، استفاده شد. نتایج برای بررسی بیشتر طی جلسه متمرکز گروهی در اختیار متخصصان سنجش قرار گرفت.                                                                                                         |
| <b>کلمات کلیدی</b><br>سنجش،<br>پزشکی،<br>دانشجو.                           | <b>نتایج:</b> نتایج حاکی از روی‌آوری بسوی قالب ارزیابی تکوینی و بهبود روش‌های سنتی در کنار روش‌های نوین بوده و نه کناره‌گیری از آنها، که در بررسی کوریکولوم دانشگاه‌های هاروارد و آکسفورد مشاهده گردید. نظر متخصصان در مورد تناسب روش‌های سنجش با قالب‌های ارزیابی و سطوح هرم میلر، راهکار و چگونگی شرایط اجرا گردآوری شد. بیشترین نظر متخصصان در رابطه با تغییر در سیاست‌گذاری‌های دانشگاه، بسترسازی، حمایت و توجه به ایده‌های نو بود. |
|                                                                            | <b>نتیجه‌گیری:</b> تحلیل داده‌ها منتج به شناسایی سیر تکاملی روش‌ها، قالب‌های مناسب سنجش و ارایه راهکارهای آموزشی-اجرایی گردید. پیشنهاد می‌شود، از نتایج بدست آمده، برای توسعه مدل‌های سنجش متناسب با سطوح ارزیابی میلر استفاده گردد.                                                                                                                                                                                                    |

[10.22038/hmed.2024.66957.1304](https://doi.org/10.22038/hmed.2024.66957.1304)

نحوه ارجاع به این مقاله ▶

Negahban Jouzan N, Karimi Moonaghi H, Poorzand H, Khajedaluee M. Exploring of Methods of Assessment of General Medical Students . Horizon of Medical Education Development. 2025;16(3):119-129

ایمیل: [karimih@mums.ac.ir](mailto:karimih@mums.ac.ir)

\*نویسنده مسئول: حسین کریمی مونی

آدرس: مشهد، مجتمع خوارزمی، دانشکده پرستاری و مامایی.



This work is licensed under a Creative Commons Attribution-Non Commercial 4.0 International License

## مقدمه

اطلاعات موردنیاز ارزشیابی آموزشی به کمک اندازه‌گیری (عموماً آزمودن)، یا سنجش گردآوری می‌شود و هرچه اطلاعات گردآوری شده توسط سنجش جامع‌تر و کامل‌تر باشد، می‌توان ارزشیابی درست‌تری انجام داد (۱). سه نوع رویکرد مهم سنجش شامل سنجش یادگیری، سنجش به عنوان یادگیری و سنجش برای یادگیری از جمله رویکردهایی هستند که امروزه مورد توجه دانشکده‌ها به‌ویژه دانشکده‌های پزشکی از حیث ارتقای توانمندی یادگیری و یاددهی و تربیت فارغ التحصیلان متخصص پزشکی و نیز رفع نیازهای بهداشت و درمان جامعه قرار گرفته است (۲). شواهد نشان می‌دهد که کیفیت آزمون‌ها به طور قابل توجهی کیفیت آموزش و یاددهی-یادگیری را تحت تاثیر قرار می‌دهد (۳).

براساس مشاهدات پژوهشگر در دانشکده پزشکی روش‌های مختلف اما مشابهی برای سنجش دانشجویان پزشکی استفاده می‌شود که می‌توان به آزمون چندگزینه‌ای (MCQ)<sup>۱</sup> و انشایی (ترکیب بازپاسخ-کوتاه پاسخ)، آزمون بالینی عینی ساختارمند (OSCE)<sup>۲</sup>، کارپوشه (Portfolio)، مشاهده مستقیم دانشجو در حین انجام پروسیجر (DOPS)<sup>۳</sup>، ارزیابی به روش بررسی پرونده بیماران ویزیت شده (CBD)<sup>۴</sup> یا (CSR)<sup>۵</sup> اشاره نمود (۴). در دهه گذشته دانشکده‌های پزشکی، تلاش‌های جدیدی برای طراحی و اجرای روش‌های ارزیابی صحیح و قابل اعتماد در زمینه توانمندی کارآموزان و پزشکان شاغل انجام داده‌اند.

طبق بررسی‌ها در نظام جامع ارزیابی پیشرفت تحصیلی دانشجویان در دانشکده‌های پزشکی مشخص گردید، در حال حاضر، بیش‌تر موارد اهداف حوزه شناختی مورد سنجش قرار گرفته و به اهداف حوزه روانی-حرکتی و عاطفی (مهارت‌های نرم‌تر مانند برقراری ارتباط، انجام کار تیمی، همدلی و...)، کم‌تر توجه می‌شود، یا اهمیت و طریقه پرداختن و سنجش این حوزه‌ها ناشناخته باقی مانده و آزمون‌ها تنها برای نمره‌دهی به دانشجویان بوده و کاربرد چندانی برای دادن

بازخورد به آنان ندارد. برای اطمینان از کیفیت یادگیری و دادن بازخورد به موقع به دانشجویان می‌توان از نوع سنجش تکوینی استفاده نمود. هلنا فریس و همکاران (۲۰۱۵)، ارزیابی تراکمی را تا سال‌های اولیه یادگیری پاسخگوی نیازهای سنجش دانسته که با ورود به بالین نیازمند سنجش تکوینی خواهیم بود (۵).

بر اساس اطلاعات موجود، روش‌های ارزیابی در آموزش پزشکی به‌ویژه آموزش بالینی در غالب دانشکده‌های پزشکی یکنواخت بوده و علاوه بر عدم تناسب کافی با اهداف آموزشی، در سنجش مهارت‌های عملی دانشجویان از کارآمدی لازم برخوردار نمی‌باشند (۶).

در پژوهش عابدی (۲۰۱۸)، در شهرستان بیرجند غالب روش‌های مورد استفاده اساتید، (۴۰٪ آزمون‌های شفاهی، ۳۳٪ ارزیابی ۳۶۰ درجه یا چند منبعی بوده و هیچ یک نیز از شبیه‌ساز بالینی استفاده نکرده‌اند)، برای سنجش مهارت‌های حوزه عاطفی و روانی-حرکتی مناسب نبوده‌اند (۷). چالش‌های موجود در زمینه ارزیابی در حیطه آموزش پزشکی و شیوه انتخاب روش مناسب ارزیابی و اثرات ناشی از انتخاب صحیح این روش‌ها متناسب با مقاطع رشته‌های مختلف علوم پزشکی، ضرورت پرداختن به این موضوع به عنوان یکی از اولویت‌های فعلی آموزش پزشکی، مطالعه و پژوهش در زمینه ارزیابی در حیطه آموزش پزشکی را بیش از هر زمان دیگر مورد تاکید قرار می‌دهد (۸). انتظارات از برنامه درسی مبتنی بر شایستگی باید سطح حرفه‌ای دانشجویان پزشکی را ارتقا بخشد. این بدیهی است که کسب تجربیات و هدف‌گذاری به آموزش و یادگیری حرفه‌ای بودن به تنهایی بسنده نمی‌شود، و حرفه‌ای بودن باید به صورت سازنده و درست نیز ارزیابی شود (۹). کریمی (۲۰۱۲)، برخی دانشجویان در رابطه با تجربه یادگیری بالینی اظهار داشتند نوع ارزیابی‌ها یکی از موانع توسعه حرفه‌ای بودن است و تنها انگیزه برای یادگیری را کسب نمره دانسته‌اند (۱۰). با توجه به ضرورت‌های بیان شده و برای پاسخ به دو سوال پژوهشی،

Case Based Discussion<sup>۴</sup>  
Chart Stimulated Recall Oral Examination<sup>۵</sup>  
Helena Ferris<sup>۱</sup>

Multiple Choice Questions<sup>۱</sup>  
Objective structural clinical examination<sup>۲</sup>  
Direct Observation of Procedural Skills<sup>۳</sup>

ISI, ERIC, AMEE, BEME و نشریه‌های معتبر دانشگاه‌های ایران ( بر اساس رتبه‌بندی شانگهای ۱۳۹۹ دانشگاه علوم پزشکی تهران و شهید بهشتی) در ۲۰ سال اخیر و برای آگاهی از بروزترین رخدادهای روش‌های سنجش در سال ۲۰۲۱ از چکیده مقاله‌های کنفرانس مجازی بین‌المللی AMEE<sup>۷</sup> در آگوست ۲۰۲۱ و کتابچه خلاصه مقاله‌های آموزش پزشکی شهید مطهری ۱۴۰۰ (بیست و دومین همایش کشوری آموزش پزشکی)، ادامه یافت. برای آگاهی از انواع روش و قالب ارزیابی در دانشکده‌های معتبر دنیا، کوریکولوم رشته پزشکی عمومی دانشگاه‌های معتبر جهان را (براساس رتبه‌بندی QS ۲۰۲۰، دانشگاه هاروارد آمریکا و آکسفورد انگلستان)، برای بررسی انتخاب گردید. مقالات به‌دست‌آمده پس از انجام فیلتر عنوان، چکیده و متن کامل و با توجه به معیارهای ورودی و خروجی غربالگری انجام شد. اولین معیار ورودی شامل ورود مطالعات انجام شده در زمینه سنجش در ۲۰ سال اخیر بود، چراکه صرف سنتی بودن یک روش دلیلی بر کنار گذاشتن آن نبوده و هدف تیم تحقیق این بود که مدل شامل روش‌های سنتی کارآمد و نوین سنجش باشد و دومین معیار، ورود تمام مطالعات با هر نوع روش انجام و مقاطع رشته پزشکی عمومی (ذکر شده در نمودار ۱) و معیار خروج مطالعاتی غیر از رشته پزشکی عمومی می‌باشد. تعداد ۱۱۶ مقاله برای بررسی انتخاب گردید. در مرحله چهارم و تحلیل، از الگوی تطبیقی نویسنده لهستانی (Beredy)، دارای چهار مرحله توصیف، تفسیر، هم‌جواری و مقایسه استفاده شد. در مرحله توصیف، پدیده‌های تحقیق بر اساس شواهد و اطلاعات، یادداشت‌برداری و با تدارک یافته‌های کافی برای بررسی و نقادی در مرحله بعد آماده می‌شوند. در مرحله تفسیر، اطلاعات توصیف شده در مرحله اول واری و تحلیل می‌شود. در مرحله هم‌جواری، اطلاعاتی که در دو مرحله قبل آماده شده‌اند به منظور ایجاد چارچوبی برای مقایسه شباهتها و تفاوتها، طبقه‌بندی و در کنار هم قرار می‌گیرند. در مرحله

۱. کدام روش‌ها برای سنجش دانشجویان پزشکی عمومی در ایران و دانشکده‌های معتبر دنیا استفاده می‌شود؟ و ۲. آیا می‌توان از نتایج برای توسعه مدل سنجش استفاده نمود؟، این پژوهش در سال ۱۳۹۹، به صورت مطالعه کتابخانه‌ای، در قالب پژوهش تطبیقی با هدف کلی "بررسی روش‌های سنجش دانشجویان پزشکی عمومی" انجام شد تا بدین وسیله بتوان انواع روش‌ها و قالب‌های سنجش متناسب با هرم میلر را مورد بررسی و توجه اساتید و برنامه‌ریزان درسی قرار داد و نگاهی نو در نحوه به‌کارگیری روش‌ها برای سنجش دانش و توانمندی‌های دانشجویان پزشکی عمومی و چه بسا توسعه مدل سنجش، ایجاد کرد.

## روش کار

با توجه به اهداف، پژوهش بصورت تطبیقی-کتابخانه‌ای شامل پنج مرحله تنظیم سوال و هدف پژوهش، جستجو برای تکراری نبودن مطالعه در ایران، در پایگاه‌های بین‌المللی و کنفرانس‌های مجازی، کوریکولوم پزشکی مصوب دانشگاه‌های معتبر خارجی و تحلیل داده و با توجه به الگو تطبیقی Beredy جستجوی گسترده‌ای آغاز گردید. برای آگاهی از ضرورت و تکراری نبودن پژوهش، باکلیدواژه‌های فارسی "روش‌های سنجش" و "دانشجویان پزشکی" و معادل انگلیسی مش ترم "Assessment methods" و "Medical students" در پایگاه‌های داده ایرانی: SID, MAGIRAN, Civilica و Gray literature، جستجو ادامه یافت. مطالعه‌ای در ایران که روش‌های سنجش دانشجویان پزشکی عمومی را از نظر سطوح هرم میلر و قابلیت اجرایی بررسی کند، مشاهده نگردید. در مرحله دوم، برای داشتن اطلاعات جامع، جستجو را به طور گسترده و بین‌المللی با کلیدواژه‌های مش ترم، شامل "Assessment Methods" و "Medical Students" و جستجوی دقیق‌تر و مرتبط‌تر با کلیدواژه "Undergraduate"، در Medline به وسیله موتور جستجوگر پابمد، ProQuest

<sup>7</sup> Association for Medical Education in Europe

## یافته‌ها

در این پژوهش، اطلاعات ۱۱۶ مقاله در جدولی شامل ستون‌های نوع روش سنجش، نویسندگان، سال پژوهش، هدف از انجام و نتیجه آن قرار داده شد، تا با توجه به الگوی تطبیقی Beredy مقایسه و درک شباهت‌ها و تفاوت‌ها صورت گیرد. پس از مقایسه، نتایج حاکی از پیشروی سنجش آموزش پزشکی به سمت قالب ارزیابی تکوینی و استفاده از روش‌های نوین سنجش در کنار روش‌های سنتی کارآمد بوده که به وضوح در بررسی کوریکولوم دانشگاه‌های معتبر هاروارد آمریکا و آکسفورد انگلستان، نیز به چشم می‌خورد (جدول شماره ۱).

جدول ۱. برنامه درسی دانشگاه هاروارد و آکسفورد

| برنامه درسی                       | محتوای آموزشی                                                                            | ارزیابی                                                                                                                                                                                                                                                                                                                                                                                   |
|-----------------------------------|------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| برنامه درسی دانشکده پزشکی هاروارد | دارای دو بخش Pathways و تکنولوژی علم سلامت (HST) ۸                                       | ارزیابی فرآیندی مداوم و تقریباً مستمر است. ارزیابی در مراحل پیش‌کارورزی-کارورزی و بالین هر یک جداگانه دارای سیاستهای ارزیابی تکوینی و تراکم مدون است (۱۲).                                                                                                                                                                                                                                |
| برنامه درسی دانشکده پزشکی آکسفورد | در دو بخش پیش بالینی (سه سال اول دوره شش ساله) شامل BM۹ و FHS ۱۰ و بالینی (از سال چهارم) | ارزیابی رسمی اصلی با مقالات کتبی در هر موضوع و سابقه رضایت بخش از حضور در جلسات تدریس اجباری و موفقیت در کار عملی. بیشتر تمرکز بر روی پرورش تفکر انتقادی می‌باشد. ارزیابی در بالین در سالهای چهار و پنج ارزیابی تکوینی مداوم و سپس ارزیابی تراکمی در انتها بعلاوه حضور رضایت‌بخش در طول دوره و انجام پروژه تحقیقاتی بوده و در سال ششم بصورت تراکمی ارزیابی کتبی و عملی انجام می‌شود (۱۳). |

روشن است که صرف سنتی و قدیمی بودن یک روش نباید آن را کنار گذارد. برای نمونه آزمون چندگزینه‌ای با وجود قدمت چندساله‌اش اما همچنان یکی از روش‌های سنجش معقول و به‌صرفه برای اساتید و موسسه‌ها به‌شمار می‌رود. گرچه باید این نکته را در نظر داشت که تلاش در مسیر بهبود

مقایسه، مسأله تحقیق با توجه به جزییات در زمینه شباهتها و تفاوتها و دادن پاسخ به سؤالات تحقیق بررسی و مقایسه می‌شود (۱۱). اطلاعات در قالب پرسشنامه‌ای شامل قبول یا رد هر روش در سطح موردنظر میلر و قابلیت اجرایی در سطح کشور، در جلسه متمرکز گروهی مطرح شد. همچنین بصورت آنلاین (به‌دلیل شرایط کرونا)، برای ۱۷ متخصص ارسال گردید که ۱۱ عضو (سه متخصص پزشکی اجتماعی، یک جراح و فلوشیپ زانو، یک متخصص در علوم تشریح، یک متخصص و فلوشیپ قلب و عروق، یک متخصص زنان و زایمان، یک دکتری آموزش بهداشت و ارتقای سلامت و سه دکتری در زمینه آموزش پزشکی) در این زمینه همکاری نمودند. بررسی و پاسخگویی طی دو ماه و پیگیری از راه ایمیل، تماس تلفنی، حضوری انجام و پاسخ‌ها براساس تنظیمات پرسشنامه آنلاین درصدهی شد.



نمودار ۱. نمودار انتخاب مقاله

#### جدول ۴. نتایج نظرسنجی روش‌های سنجش در سطح HOW

| نوع روش<br>ارزیابی | سطح<br>SHOWS<br>HOW | قالب<br>تکوینی | قالب<br>تراکمی | قابلیت<br>اجرایی |
|--------------------|---------------------|----------------|----------------|------------------|
| CA-OSCE            | (٪۱۰۰)۱۱            | (٪۹/۹۰)۱۰      | -              | (٪۱۰۰)۱۱         |
| CBD                | (٪۹/۹۰)۱۰           | (٪۸۱/۸۱)۹      | (٪۹/۹۰)۱۰      | (٪۸۱/۸۱)۹        |
| DOPS               | (٪۹/۹۰)۱۰           | (٪۹/۹۰)۱۰      | -              | (٪۸۱/۸۱)۹        |
| SLICE              | (٪۱۰۰)۱۱            | (٪۵۴/۵۴)۶      | -              | (٪۶۳/۶۳)۷        |
| M-PAT              | (٪۷۲/۷۲)۸           | (٪۷۲/۷۲)۸      | -              | (٪۷۲/۷۲)۸        |
| SP(VPS)            | (٪۱۰۰)۱۱            | (٪۱۰۰)۱۱       | -              | (٪۶۳/۶۳)۷        |
| OSPE               | (٪۱۰۰)۱۱            | (٪۷۲/۷۲)۸      | -              | (٪۸۱/۸۱)۹        |
| OSLER              | (٪۱۰۰)۱۱            | -              | (٪۹/۹۰)۱۰      | (٪۷۲/۷۲)۸        |
| OSCE               | (٪۱۰۰)۱۱            | -              | (٪۱۰۰)۱۱       | (٪۱۰۰)۱۱         |
| SP                 | (٪۱۰۰)۱۱            | -              | (٪۱۰۰)۱۱       | (٪۸۱/۸۱)۹        |

#### جدول ۵. نتایج نظرسنجی روش‌های سنجش در سطح Does

| نوع روش<br>ارزیابی | سطح<br>DOES | قالب<br>تکوینی | قالب<br>تراکمی | قابلیت<br>اجرایی |
|--------------------|-------------|----------------|----------------|------------------|
| VTCS               | (٪۸۱/۸۱)۹   | (٪۸۱/۸۱)۹      | -              | (٪۸۱/۸۱)۹        |
| MINI-CEX           | (٪۱۰۰)۱۱    | (٪۱۰۰)۱۱       | (٪۷۲/۷۲)۸      | (٪۱۰۰)۱۱         |
| Degree۳۶۰          | (٪۱۰۰)۱۱    | (٪۶۳/۶۳)۷      | -              | (٪۸۱/۸۱)۹        |
| PA                 | (٪۷۲/۷۲)۸   | (٪۷۲/۷۲)۸      | -              | (٪۷۲/۷۲)۸        |
| Portfolio          | (٪۱۰۰)۱۱    | -              | (٪۹/۹۰)۱۰      | (٪۱۰۰)۱۱         |
| Logbook            | (٪۹/۹۰)۱۰   | -              | (٪۸۱/۸۱)۹      | (٪۹/۹۰)۱۰        |
| SRPE               | (٪۹/۹۰)۱۰   | -              | (٪۸۱/۸۱)۹      | (٪۸۱/۸۱)۹        |
| SJTS               | (٪۸۱/۸۱)۹   | -              | (٪۷۲/۷۲)۸      | (٪۸۱/۸۱)۹        |
| DOPS               | (٪۱۰۰)۱۱    | (٪۹/۹۰)۱۰      | -              | (٪۸۱/۸۱)۹        |

مشروح نظرات:

روش ارزیابی Context Rich Short Answer Question (CR-SAQ) را تا حدودی برای ارزیابی دانش در سطح KNOWS و برای سنجش در حد محفوظات ضروری انتخاب نمودند اما از نظر قابلیت اجرایی به دلیل محدودیت زمانی در طراحی مناسب نمی دانند.

کیفیت و اعتبار روش‌ها را نباید فراموش کرد. طبق اطلاعات استخراج شده و نظر متخصصان، مهم‌ترین یافته در این پژوهش که می‌توان بیان نمود این است که برای استفاده از انواع روش‌های ارزیابی در هر یک از قالب‌های تکوینی و تراکمی اجماعی وجود نداشته و اجماع به سمت استفاده درست از انواع روش‌ها و تاکید بر به‌کارگیری قالب ارزیابی تکوینی در کنار تراکمی بوده در حالیکه غالب پژوهش‌های داخل کشور نشان داده که برخلاف آنچه دربرنامه‌درسی پزشکی ذکر شده در نهایت معیار اصلی ارزیابی فقط آزمون پایان ترم و روش‌های متداول سنجش بوده است. (نتایج پیوست گردید).

در مطالعه حاضر نتایج از نظر قابلیت اجرایی و تناسب با هرم میلر بصورت چارچوب ارزیابی، در اختیار متخصصان قرار گرفته و پاسخ‌ها و نظرات جمع‌آوری گردید (جدول شماره ۲-۳-۴-۵).

#### جدول ۲. نتایج نظرسنجی روش‌های سنجش در سطح Knows

| نوع روش<br>ارزیابی | سطح<br>KNOWS | قالب<br>تکوینی | قالب<br>تراکمی | قابلیت<br>اجرایی |
|--------------------|--------------|----------------|----------------|------------------|
| CR-SAQ             | (٪۹/۹۰)۱۰    | (٪۹/۹۰)۱۰      | (٪۹/۹۰)۱۰      | (٪۸۱/۸۱)۹        |
| MCQ                | (٪۹/۹۰)۱۰    | (٪۹/۹۰)۱۰      | (٪۱۰۰)۱۱       | (٪۱۰۰)۱۱         |
| EMI/EMQ            | (٪۱۰۰)۱۱     | (٪۱۰۰)۱۱       | (٪۹/۹۰)۱۰      | (٪۹/۹۰)۱۰        |

#### جدول ۳. نتایج نظرسنجی روش‌های سنجش در سطح Knows

| نوع روش<br>ارزیابی | سطح<br>KNOWS<br>HOW | قالب<br>تکوینی | قالب<br>تراکمی | قابلیت<br>اجرایی |
|--------------------|---------------------|----------------|----------------|------------------|
| Essay              | (٪۸۱/۸۱)۹           | (٪۷۲/۷۲)۸      | (٪۹/۹۰)۱۰      | (٪۵۴/۵۴)۶        |
| KF                 | (٪۱۰۰)۱۱            | (٪۹/۹۰)۱۰      | (٪۱۰۰)۱۱       | (٪۸۱/۸۱)۹        |
| E-PMP              | (٪۸۱/۸۱)۹           | (٪۹/۹۰)۱۰      | -              | (٪۸۱/۸۱)۹        |
| POE                | (٪۹/۹۰)۱۰           | (٪۸۱/۸۱)۹      | (٪۷۲/۷۲)۸      | (٪۵۴/۵۴)۶        |
| PMP                | (٪۱۰۰)۱۱            | -              | (٪۹/۹۰)۱۰      | (٪۸۱/۸۱)۹        |
| SCT                | (٪۹/۹۰)۱۰           | -              | (٪۹/۹۰)۱۰      | (٪۹/۹۰)۱۰        |
| CBWE               | (٪۹/۹۰)۱۰           | -              | (٪۹/۹۰)۱۰      | (٪۷۲/۷۲)۸        |



نیز با توجه به زمان بر بودن روند انجام آزمون به عنوان تراکمی با نظر پنج نفر از اساتید مناسب‌تر است و می‌تواند هم در غالب تکوینی- تراکمی استفاده شود. از طرفی طرح آن نیازمند فرصت کافی اساتید بوده که همچنان راهکار آن سیاست تشویقی است.

روش‌های Computer Assisted Objective Structured Clinical Examination (CA-Virtual Patient (VPS), OSCE) Standardized (SP), Simulation Structured Real-Patient, Patient Computer- (CBWE), Examination (SRPE) Video- Examination Based Written based Test of Communication Skills (VTCS) از نظر اجرایی نیازمند تهیه سخت‌افزار و فضای کافی فیزیکی چه برای انجام آزمون و چه نگهداری از وسایل و تجهیزات برای روش‌های که نیازمند تجهیزات هزینه‌بر بوده و نیز شرایط مدیریتی مناسب برای نگهداری آن‌هاست و در واقع نیازمند پذیرش ضرورت تهیه و صرف بودجه از جانب دانشگاه آموزشی می‌باشد. برخی اساتید نیز اظهار داشته‌اند که با نحوه استفاده برخی از روش‌ها آشنا نبوده‌اند و به این خاطر برای سنجش دانشجویان تنها به تعداد محدودی از روش‌ها که طبق روال استفاده می‌شده استناد نموده‌اند.

## بحث

پژوهش حاضر با هدف کلی واکاووی روش‌های سنجش دانشجویان پزشکی عمومی، انجام شد. با بررسی در متون منتشرشده در ایران، مطالعه‌ای مبنی بر بررسی قابلیت اجرایی روش‌های سنجش دانشجویان پزشکی عمومی و تناسب آن با سطوح هرم میلر یافت نشد. پژوهش‌هایی که بدان پرداخته شده در مورد بررسی معایب و مزایای یک نوع روش سنجش، میزان استفاده از روش‌ها توسط اساتید و یا اشاره به فاصله بین آموزش‌های تیوری-عملی و سنجش بوده است. به عنوان نمونه؛ یوسفی و همکاران (۲۰۱۲)، به بررسی معایب و مزایای

یک متخصص روش MCQ را به دلیل انتخاب تصادفی یا رد گزینه‌ها توسط دانشجو مورد توافق در سطح KNOWS در قالب تکوینی انتخاب ننموده و روش Extended (EMI) Matching Item را به شرط همکاری اساتید و طراحی مناسب قابل اجرا دانسته است.

روش‌های (OSPE), Key Feature Questions (KF), Objective Structured Practical Patient Management (PMP), Examination /Electronic Patient (EPMP) / Problem Mini-Peer Management Problem (PA) Peer و Assessment Tool (MPAT) Assessment تحت شرایط همکاری اساتید و طرح سوال مناسب قابل اجراست و از طرفی چون طرح سوال مناسب زمان‌بر است استفاده از سیاست‌های تشویقی برای همکاری اساتید را می‌طلبد.

روش (POE) Problem-Based Oral Examination اگر دارای ویژگی عینی بودن باشد، قابلیت سنجش دانش در سطح به کاربردن را دارد اما با توجه به مشکلات از جمله زمان‌بر بودن و تعداد بالای دانشجویان و تعداد پایین اساتید، استفاده از این روش در قالب تکوینی را دچار مشکل می‌کند. بهتر است از ارزیابان متعدد همگن به صورت Cross Over ۱۱ استفاده و سوالات به صورت یکسان طراحی گردد تا اعمال نظر شخص ارزیاب به حداقل برسد.

مشکل اجرا در روش Long Answer Essay Question (LAQ) / Open Ended Question (OEQ) وجود دارد که ممکن است نظر ممتحن خواسته یا ناخواسته در برگه پاسخ دانشجو اعمال شود که باید نام دانشجو و پاسخ اصلی در هر سوال کدگذاری شده و از چند ارزیاب استفاده گردد.

روش‌های Structured Long Interview (SLICE) Script (SCT) /and Clinical Examination Situational Judgment Concordance Test Objective Record (OSLER) و Tests (SJTS), Structured Long Examination

۱۱ مقاطع

روش‌های رایج را عدم آگاهی نسبت به استفاده و نحوه اجرای آن‌ها دانستند؛ همخوانی دارد.

انجام ترکیبی و یا منفرد روش‌های سنجش نیازمند بستر سازی و فراهم کردن شرایط اجرایی می‌باشد؛ محدودیت‌هایی در تامین منابع مالی و زمان از موانع ارایه و بکارگیری راهکارهای مناسب به‌شمار می‌روند که در این راستا باید سیاستگذاران اهم تلاش خود را داشته باشند. طبق بررسی‌های انجام شده برای سنجش دانشجویان در نظر گرفتن چند روش با توجه به حیطه مورد ارزیابی بسیار بااهمیت است. همانطور که با بیش از نیمی از نظرات موافق متخصصان این پژوهش همراه بوده و از این رو با مطالعه طاهرزاد و همکارش (۲۰۰۸)، همخوانی دارد. وی سنجش توانمندی بالینی را کاملاً پیچیده دانسته که برای ارزیابی معتبر آن به روش‌های چندگانه، ترکیبی و سطح بالا نیاز است (۱۹).

در پژوهش حاضر، متخصصان بر این باورند که یکی از تغییرات در سیاست‌های دانشگاهی می‌تواند تشویق، سرمایه‌گذاری، توجه به نواوغ و ایجاد بستری برای ظهور اندیشه‌های نو باشد و تاکید داشتند که توجه به نواوغ می‌تواند سبب پیشرفت چشمگیر در حوزه سنجش گردد و در سخت‌ترین شرایط مانند دوران پاندمیک کرونا به کمک اساتید بیاید. که با نتایج لیوبا کونوپاسک<sup>۱۲</sup> و همکاران (۲۰۱۶)، که ارزیابی تکوینی را در کالبد "ارزیابی برای یادگیری" به جای "ارزیابی از یادگیری" نشان داده و پیشنهاد کرده‌اند، به جای صرف هزینه‌های کلان برای تغییر برنامه‌های جدید، تغییرات فرهنگی و بازپس‌گیری زیرساخت‌ها با مدیریت و رهبری قوی باید انجام شود (۲۰)، مشابهت داشت.

نتایج این مطالعه از نظر توجه به نواوغ و استفاده از ایده‌هایشان، مشابهت‌هایی با نتایج پژوهش شایان (۲۰۱۱)، داشت. وی نرم‌افزاری برای سنجش قدرت استدلال بالینی، به نام EPMP طراحی نمود که با استقبال دانشجویان و اساتید همراه بود (۲۱). مهم است که در نظر داشته باشیم انتخاب روش‌های آموزش و سنجش مناسب باید منجر به ارتقای عملکرد و افزایش انگیزه گردد. به خصوص در حیطه پزشکی

روش ارزیابی همتایان پرداخته و آن را روشی ارزان و در دسترس برای ارزیابی حیطه عاطفی قلمداد کردند (۱۴). میرزایی و همکاران (۲۰۱۵)، به میزان استفاده اعضای هیات علمی دانشگاه علوم پزشکی ایلام از روش‌های سنجش چندگزینه‌ای، آسکی و ۳۶۰ درجه اشاره نمودند (۱۵). قادری و همکاران (۲۰۱۵)، نگرش دانشجویان بالینی دانشگاه علوم پزشکی بیرجند را در مورد تناسب محتوای دوره علوم پایه و آزمون‌ها بررسی کردند. غالب دانشجویان اظهار داشتند که بین سوال‌های آزمون جامع علوم پایه و دروس مورد مطالعه در دوره علوم پایه رابطه معنی‌داری وجود ندارد (۱۶). بنابراین غالب پژوهش‌ها در ایران در سطح اهداف گفته شده بوده و نه فراتر که از این جهت با اهداف مطالعه حاضر مغایرت دارد؛ چراکه برای تربیت نیروهای متخصص، حرفه‌ای و متعهد نیازمند پژوهش‌های جامع در مورد انتخاب و نحوه بکارگیری صحیح روش و قالب سنجش در حیطه مرتبط با در نظرگیری شرایط اجرایی آنهاست.

ذاکری‌مقدم و همکارانش (۲۰۲۱)، مروری بر روش‌های ارزیابی نوین برای تمام دانشجویان علوم پزشکی در ده سال اخیر، انجام داده و پیشنهاد دادند که باتوجه به هرم میلر روش مناسب سنجش انتخاب گردد (۱۷). نتایج این مطالعه از این نظر که ارزیابی در یک حوزه و مطابق مطالعات پیشین تعیین میزان استفاده از روش‌ها بوده با نتایج مطالعه حاضر مغایرت داشت اما مشابهت‌هایی از نظر بررسی در ده سال اخیر و پیشنهاد نویسندگان، توجه به هرم ارزیابی میلر، یافت شد. چرا که تیم تحقیق نیز روش‌های سنجش (در حوزه‌های شناختی روانی-حرکتی و عاطفی به‌واسطه مرور متون ۲۰ساله) را واکاوی و علاوه بر تناسب با هرم ارزیابی میلر، شرایط اجرا را نیز بررسی نمودند.

کمیلی و همکارانش (۲۰۰۲)، در دانشگاه علوم پزشکی زاهدان، نشان دادند؛ برخلاف استفاده بیشتر اساتید از قالب ارزیابی مرحله‌ای در طول ترم اما بازهم ملاک مهم ارزیابی برپایه آزمون پایان ترم بوده است (۱۸). که با نتایج در پژوهش حاضر از این نظر که متخصصان نیز ملاک استفاده از برخی

<sup>۱۲</sup> Lyuba Konopasek



دانشگاه‌های کشور و دیگر دانشگاه‌های علوم پزشکی بواسطه مرور متون، مورد بررسی قرار داد تا پیشرفت در بکارگیری روش‌ها را در توانمندسازی و کارآمدی دانشجویان پزشکی عمومی که در خط اول بهداشت-درمان قرار دارند، روشن سازد. طبق نظر متخصصان در این زمینه، نباید تلاش و دغدغه پژوهشگران برای ساماندهی در بخش آموزش و سنجش بی‌جواب باقی بماند؛ بلکه پیشنهاد می‌شود تا سیاست‌های تشویقی متناسب با روند های اجرایی در هر دانشکده و نیز بازنگری دوره‌ای و دریافت بازخورد از اساتید و دانشجویان انجام گردد. همچنین می‌توان با ترکیب روش‌های مذکور در قالب‌های سنجش و متناسب با سطوح هرم میلر گامی در جهت توسعه مدل‌های سنجش برداشت.

### تقدیر و تشکر

بدین‌وسیله از دانشگاه علوم پزشکی مشهد به دلیل همکاری در روند انجام تحقیق، تشکر و قدردانی می‌شود.

### تضاد منافع

تمام نویسندگان میزان مشارکت برابری را در تهیه دست نوشته حاضر داشته‌اند و به صراحت اعلام می‌نمایند که هیچگونه تضاد منافع از نظر مالی، اداری، سیاسی، نژادی، قومیتی، با مندرجات مقاله ندارند.

### حمایت مالی

این مقاله از طرح پژوهشی مصوب با کد ۹۹۲۱۷۸ استخراج شده است بدین‌وسیله از دانشگاه علوم پزشکی مشهد به دلیل تأمین اعتبار مالی تحقیق، تشکر و قدردانی می‌شود.

### ملاحظات اخلاقی

این مقاله حاصل بخشی از پایان نامه تحت عنوان "واکاوی روش‌های سنجش دانشجویان رشته پزشکی عمومی به منظور ارایه مدلی برای سنجش" در مقطع کارشناسی ارشد آموزش پزشکی با شماره پایان نامه ۱۷۶۰ آ و کد اخلاقی IR.MUMS.MEDICAL.REC.1400.291 می‌باشد.

که رشته‌ای ارزشمند در حفظ و بهبود سلامت جامعه دارد. بنابراین این سوال مطرح می‌شود که آیا شیوه ارزیابی ما روحیه پرسشگری، تفکر انتقادی و خودکفایی علمی را در آینده سازان سلامت مردمانمان ایجاد می‌کند؟ (۲۲). رسیدن به اهداف بهداشت-درمان و آموزش پزشکی، اعم از ارتقای سطح دانش و به‌کارگیری آن در جهت افزایش و بهبود خدمات به بیماران برای داشتن جامعه‌ای سالم، عزم جدی را می‌طلبد و باید از تحقیق‌های انجام شده در این‌باره با مشورت کارشناسان مربوطه بهره جست و تنها به سخنرانی یا تاکیدات در سطح کتاب و دستورالعمل بسنده نکرد و پیگیری لازم نیز انجام گردد. تیم تحقیق پیشنهاد می‌کند که می‌توان از نظرات کارشناسان با تجربه آموزشی برای جایگزینی روش‌های قبلی یا اصلاح برگزاری و انتخاب روش و قالب ارزیابی مناسب بهره برد. همچنین با برگزاری جلسه‌های هم‌فکری و مشاوره‌ای با گروه‌ها برای آشنایی با روش‌های جدید ارزیابی؛ چرا که لازم است با توجه به محتواهای آموزشی، گروه‌ها در زمینه انتخاب آزمون‌ها راهنمایی گردند. مشابه سنجه‌هایی که برای آزمون‌های چهارگزینه‌ای (ضریب تمیز یا درجه دشواری و...) انجام می‌گیرد.

#### • پیشنهادهای اجرایی متخصصان:

با توجه به نظر متخصصان و صاحب‌نظران در زمینه سنجش دانشجویان پزشکی راهکارهایی برای بهبود روند سنجش ارایه می‌گردد. امید است مورد توجه سیاست‌گذاران و برنامه‌ریزان آموزش و سنجش قرار گیرد:

۱. اجرای سیاست‌های تشویقی متناسب با روند های اجرایی در هر دانشکده
۲. بازنگری دوره‌ای روش‌های ارزیابی گروه‌های مختلف بالینی و پایه
۳. ارزیابی آزمون‌های انجام‌شده قبلی دیگر مانند OSCE، KF یا PMP که توسط کارشناسان مورد تایید بوده و همچنین دادن بازخورد در مورد اشکالات طراحی آن‌ها.

### نتیجه گیری

پژوهش حاضر، روش‌های ارزیابی پیشرفت تحصیلی، نحوه کاربرد و قابلیت اجرایی هر روش در قالب صحیح را در سطح

## مشارکت نویسندگان

| مفهوم سازی      | طراحی طرح | جمع آوری و/یا پردازش داده ها | تجزیه و تحلیل و/یا تفسیر | نگارش پیش نویس اصلی | بررسی انتقادی و ویرایش | نهایی سازی | پاسخگویی به تمام جنبه های کار |
|-----------------|-----------|------------------------------|--------------------------|---------------------|------------------------|------------|-------------------------------|
| ن. نگهبان جوزان |           |                              |                          |                     |                        |            |                               |
| ح. کریمی مونیقی |           |                              |                          |                     |                        |            |                               |
| ه. پورزند       |           |                              |                          |                     |                        |            |                               |
| م. خواجه دلویی  |           |                              |                          |                     |                        |            |                               |

\* راهنمای رنگ‌ها:



## References

1. Seif a. Educational measurement, assessment, and evaluation. Seven, editor. . 7, editor. Tehran: Tehran, Douran; 2018. 8-36 p.
2. Shumway JM, Harden RM. AMEE Guide No. 25: The assessment of learning outcomes for the competent and reflective physician. Medical teacher. 2003;25(6):569-84.
3. Koohpayezade Isfahani jalil, Kamran HspSAS. Investigating the use of clinical assessment methods in medical sciences universities in Tehran from the point of view of clinical professors. SID. 2012;16(5):395-402.
4. Negahban Jouzan N, Karimi Moonaghi H, Poorzand H, Khajedaloue M. Presenting a model for assessing general medical students: a brief report. Tehran University Medical Journal. 2023;81(1):86-92.
5. Ferris H, O'Flynn D. Assessment in Medical Education; What Are We Trying to Achieve? International Journal of Higher Education. 2015;4(2):139-44.
6. Brady AM. Assessment of learning with multiple-choice questions. Nurse Educ Pract. 2005;5(4):238-42.
7. Abedi F. An Investigation of the Application of Different Methods of Student evaluation by clinical education groups of Birjand University of Medical Sciences. Future of Medical Education Journal. 2018;8(1):22-6.
8. Medicine Io. Committee on Quality of Health Care in America. Crossing the quality chasm: a new health system for the 21st century. National Academies Press. 2001.
9. Van Mook WN, van Luijk SJ, O'Sullivan H, Wass V, Schuwirth LW, van der Vleuten CP. General considerations regarding assessment of professional behaviour. European Journal of Internal Medicine. 2009;20(4):e90-e5.
10. Karimi moonaghi H, Derakhshan A, Khajedaloue M, Dashti Rahmat abadi MB, Tahereh. Lived Clinical Learning Experiences of Medical Students: A Qualitative Approach. Iranian Journal of Medical Education. 2012;11(6):635-47.
11. Karimi Moonaghi H, Montazeri R. A comparative study of the curriculum of masters degree in medical education in Iran and some other countries. Strides in Development of Medical Education. 2015;11(4):420-33.
12. Harvard Medical School Curriculum America: Harvard Medical School; 2021/12/08 [Available from: Available at: <https://meded.hms.harvard.edu/>].
13. Curriculum Oxford School of Medical Sciences: Oxford School of Medical Sciences; 2021 [Available from: Available at: <https://www.medscl.ox.ac.uk/study/medicine/pre-clinical/structure>].
14. Yousefy A. Peer assessment in evaluation of medical sciences students. Iranian Journal of Medical Education. 2012;11(9):1443-52.
15. Mirzaei A, Kawarizadeh F, Lohrabian V, Yegane Z. Evaluation methods of the academic achievement of students Ilam University of Medical Sciences. Education strategies in medical sciences. 2015;8(2):91-7.
16. Ghaderi R. Medical Students' Attitudes towards Comprehensive and Basic Sciences Exam Questions in Birjand University of Medical Sciences. Journal of Medical Education. 2015;14(2):35-92.
17. Zakerimoghdam M, Yazdanparast E, Hosseiny SF, Chenari HA. A Review of New Methods Assessment in Clinical Education of Medical Science Students. Iranian Bimonthly of Education Strategies In Medical Sciences. 2021;14(3):92-102.
18. KOMELI G, REZAI G. Methods of student assessment used by faculty members of basic medical sciences in Medical University of Zahedan. 2001,1(4):52-57.
19. Tahernezhad K, Javidan F. Advanced Assessment of Medical Students Clinical Performance: Challenges Methods and Approaches. Strides in development of medical education. 2008;5(1):58-70.
20. Konopasek L, Norcini J, Krupat E. Focusing on the formative: building an assessment system aimed at student growth and development. Academic Medicine. 2016;91(11):1492-7.
21. Shayan S. Using Patient Management Problem (EPMP) in Assessment of Clinical Competency. Iranian Journal of Medical Education. 2011;10(5):1087-92.
22. Alizadeh M, Mazouchian H. Letter to editor. Iranian Journal of Medical Education. 2015;15(0):505-7.