

بررسی روایی و پایایی ابزار ارزیابی نقادانه چکیده مقالات اصیل پژوهشی توسط داوران دانشجو در کنگره‌های پژوهشی دانشجویان علوم پزشکی کشور

سید محمد فرشته‌نژاد^۱، حمیدرضا برادران^۲، مازیار مرادی لاکه^۳

^۱ پژوهشگر، گروه آموزش پزشکی، دانشکده مجازی؛ مرکز تحقیقات آموزش پزشکی، دانشکده پزشکی، دانشگاه علوم پزشکی تهران، تهران، ایران

^۲ دانشیار، گروه آموزش پزشکی، مرکز تحقیقات آموزش پزشکی، دانشکده پزشکی، دانشگاه علوم پزشکی تهران، تهران، ایران

^۳ دانشیار، گروه پزشکی اجتماعی، دانشکده پزشکی، دانشگاه علوم پزشکی تهران، تهران، ایران

نشانی نویسنده مسؤل: دانشکده پزشکی، دانشگاه علوم پزشکی تهران، سید محمد فرشته‌نژاد

E-mail: sm.fereshtehnejad@ki.se

وصول: ۹۲/۲/۲۹، اصلاح: ۹۲/۴/۱۵، پذیرش: ۹۲/۵/۸

چکیده

زمینه و هدف: امروزه نکته مهم در استفاده از نتایج پژوهش‌های ارائه شده در کنفرانس‌های علمی، توانمندی استفاده‌کنندگان خواهد در ارزیابی کیفیت یافته‌هاست. استفاده از چارچوب استاندارد ارزشیابی همچون چک‌لیست‌ها به‌عنوان ابزار اصلی در داوری همتایان محسوب می‌شود. لذا پژوهش حاضر با هدف بررسی روایی و پایایی چک‌لیست طراحی شده برای ارزیابی نقادانه چکیده مقالات اصیل پژوهشی توسط داوران دانشجو در کنگره‌های پژوهشی دانشجویان علوم پزشکی کشور انجام شد.

مواد و روش‌ها: مطالعه حاضر بخشی از یک پژوهش از نوع مداخله آموزشی است که در قالب یک دوره کارگاه آموزشی و با حضور ۴۰ نفر از دانشجویان رشته پزشکی از دانشگاه‌های علوم پزشکی تهران برگزار گردید. روش نمونه‌گیری از نوع غیر احتمالی هدفمند می‌باشد. شرکت‌کنندگان قبل و پس از برگزاری کارگاه، با شبیه‌سازی داوری سه چکیده مقاله را ارزیابی کردند. پایایی چک‌لیست طراحی شده به دو روش مختلف شامل ضریب همبستگی اسپیرمن و ضریب آلفای کرونباخ ارزیابی شد. روایی ابزار نیز به‌روش دلفی توسط متخصصین تایید شد.

یافته‌ها: در تمامی بخش‌های اصلی چکیده مقاله شامل مقدمه، روش بررسی، نتایج، نتیجه‌گیری و حیطه نظر کلی -هم در پیش‌آزمون و هم در پس‌آزمون- همسانی درونی گزارش شد ($P < 0/05$). ضریب پایایی آلفای کرونباخ در بخش "روش بررسی" (۰/۹۲۶) عالی و در دو بخش "نتایج" (۰/۶۵۰) و "نتیجه‌گیری" (۰/۷۳۹) در پیش‌آزمون قابل قبول بود.

نتیجه‌گیری: ابزار معرفی شده در مطالعه حاضر از روایی و پایایی قابل قبولی برخوردار بود. با توجه به حضور فعال دانشجویان علوم پزشکی در کنگره‌های پژوهشی دانشجویی بعنوان داور، استفاده از چنین چک‌لیست‌های ساختاری می‌تواند در بهبود کیفیت ارزیابی چکیده مقالات و داوری همتایان موثر باشد.

واژه‌های کلیدی: داوری همتایان، ارزشیابی نقادانه، چک‌لیست، پایایی، روایی، کنگره‌های پژوهشی، دانشجویان علوم پزشکی

مقدمه

به دنبال گسترش روزمره تولید شواهد و یافته‌های مطالعات در حیطه علوم پزشکی، لزوم دسترسی سریع‌تر به این اطلاعات بیش از پیش برای دانشجویان علوم پزشکی احساس می‌شود. از سوی دیگر شواهد مختلف نشان می‌دهد، کاربردی شدن حجم قابل توجهی از این شواهد به آهستگی صورت می‌گیرند و یا هرگز مورد استفاده قرار نمی‌گیرند (۱). این دسته از شواهد پزشکی معمولاً در قالب چکیده مقالات ارائه شده در کنگره‌ها و کنفرانس‌های علمی و یا چاپ در مجلات علمی-پژوهشی معتبر منتشر می‌شوند. ضمن آن‌که توانمندی بالقوه دنیای الکترونیک، هر روز بیش از گذشته به نشر گسترده و سریع این یافته‌ها منجر می‌شود (۴-۲) تا جایی که تخمین زده می‌شود سالانه بیش از ۱/۵ میلیون مقاله پژوهشی در مجلات علمی علوم پزشکی به انتشار می‌رسد که علی‌رغم گذر از مرحله داوری و بررسی توسط متخصصین، برخی مقالات با کیفیت‌های نامناسب در این بین به چشم می‌خورند. نکته قابل توجه در استفاده از نتایج این پژوهش‌ها، توانمندی استفاده‌کنندگان شواهد در ارزیابی کیفیت یافته‌هاست؛ از این رو امروزه ارزیابی نقادانه به‌عنوان یکی از مصادیق تفکر نقادانه، به‌عنوان ضروری‌ترین مهارت مورد نیاز برای فارغ‌التحصیلان رشته‌های مختلف علوم پزشکی می‌باشد. با گسترش روزافزون مبانی پزشکی مبتنی بر شواهد، اصول و مفاهیم ارزیابی نقادانه بیشتر از هر زمان دیگری مورد توجه قرار گرفته است و در بسیاری از مراکز آموزشی به جزئی لاینفک از آموزش پزشکی و سایر رشته‌های علوم پزشکی مبدل گشته است.

از جمله ویژگی‌های اساسی در نقد سیستماتیک و نظام‌مند یک مطالعه پژوهشی می‌توان مواردی همچون برخورداری از تفکری نقادانه، استفاده از چارچوبی سیستماتیک برای نقد و آشنایی با قوانین استاندارد پژوهشی مورد توافق محققین علوم پزشکی را بر شمرد

(۵). بنابراین ارزیابی نقادانه مقالات علمی-پژوهشی و تفسیر قابل اعتماد این شواهد، نیازمند تسلط دانشجویان به اصول و روش‌های متدولوژیک مطالعات علوم پزشکی با نگاهی موشکافانه است (۱). آموزش رسمی این مهارت‌ها به دانشجویان پزشکی در این زمینه‌ها مفید می‌باشد: شناخت خطاهای بالقوه موجود در پژوهش‌ها، درک بهتر متدولوژی گزارش یافته‌های کمی با شاخص‌های آماری توصیفی و تحلیلی، توان تعمیم‌پذیری نتایج پژوهش‌ها و قضاوت در مورد اعتبار شواهد و نحوه به‌کارگیری آن‌ها در بالین بیمار.

یکی از جنبه‌های ارزیابی نقادانه مقالات پژوهشی علوم پزشکی به بررسی کیفیت مقالات و یا چکیده مقالات ارسالی برای پذیرش در مجلات علمی و یا ارائه در کنگره‌ها توسط همکاران برمی‌گردد که از آن با عنوان "داوری همتایان" یاد می‌شود. مسئولیت‌های اصلی داوران در روند ارزیابی همتایان مقالات اصیل پژوهشی عبارتند از: بررسی صحت متدولوژی و تکنیکال پژوهش انجام شده، تعیین اصیل بودن و جدید بودن یافته‌های مطالعه، تأیید حمایت نتیجه‌گیری نهایی و پیشنهادات ارائه شده در انتهای مقاله با داده‌ها و یافته‌های آن مطالعه و در نهایت قضاوت در مورد مناسب بودن آن مقاله برای چاپ در مجلات مربوطه و یا ارائه در کنگره مد نظر (۶). امروزه داوری همتایان، مرکز توجه انتشار شواهد علمی در مجلات علوم پزشکی می‌باشد و مرحله‌ای بسیار مهم در روند انتشار مقالات به‌شمار می‌رود (۷). نتایج یک پژوهش نشان داد، کیفیت داوری همتایان و ارزیابی نقادانه توسط دانشجویان پزشکی با کیفیت داوری‌های انجام شده توسط اساتید برابری می‌کند؛ به‌عبارتی دانشجویان علوم پزشکی نیز می‌توانند در عرصه داوری همتایان نقشی فعال داشته باشند (۸).

در کنار آموزش مهارت‌های لازم برای ارزیابی نقادانه شواهد، استفاده از چارچوب استاندارد ارزشیابی - همچون چک‌لیست‌ها - به‌عنوان ابزار اصلی در داوری

همتایان دارای اهمیت ویژه‌ای است. امروزه چک‌لیست‌های استاندارد ارزیابی مقالات مانند consort و strobe مورد استفاده بسیاری از مجلات و مراکز پژوهشی است. اما باید توجه داشت که بسیاری از این چک‌لیست‌ها برای ارزیابی متن کامل مقالات پژوهشی طراحی شده و می‌بایست برای ارزیابی چکیده مقالات ارسالی در کنگره‌های علمی-پژوهشی تغییراتی هرچند مختصر صورت گیرد. با توجه به اهمیت و شکوفایی پژوهش‌های دانشجویان علوم پزشکی کشور در طی سال‌های اخیر، برگزاری کنگره‌های متعدد سالیانه و دوره‌ای دانشجویی و انجام مراحل داوری همتایان توسط دانشجویان، طراحی و بومی‌سازی ابزاری استاندارد برای ارزیابی چکیده‌های ارسالی از اهمیتی دوچندان برخوردار است. از این‌رو پژوهش حاضر با هدف بررسی روایی و پایایی چک‌لیست طراحی شده برای ارزیابی نقادانه چکیده مقالات اصیل پژوهشی توسط داوران دانشجو در کنگره‌های پژوهشی دانشجویان علوم پزشکی کشور انجام شده است.

مواد و روش‌ها

مطالعه حاضر بخشی از یک پژوهش از نوع مداخله آموزشی است که در قالب یک دوره کارگاه آموزشی به میزبانی دانشگاه علوم پزشکی شاهد در تاریخ ۱۳۸۹/۷/۱۵ و با حضور ۴۰ نفر از دانشجویان رشته پزشکی از دانشگاه‌های علوم پزشکی شهر تهران، از طرف دبیرخانه دائمی کنگره‌های ملی و بین‌المللی پژوهشی دانشجویان علوم پزشکی کشور به صورت ۱ روزه و به مدت ۱۰ ساعت برگزار شد. داده‌های مربوط به این کارگاه آموزشی برای بررسی روایی و پایایی چک‌لیست ارزیابی چکیده مقالات مورد استفاده قرار گرفت. پایایی چک‌لیست به دو روش مختلف ضریب همبستگی اسپیرمن و ضریب آلفای کرونباخ ارزیابی شد و روایی ابزار نیز به روش دلفی مورد تایید کارشناسان و متخصصین قرار گرفت.

دانشجویان مورد مطالعه از بین اعضای فعال کمیته‌های تحقیقات دانشجویی در دانشگاه‌های علوم پزشکی تهران و بر اساس نمرات اخذ شده در شناسه پژوهشی آنان به روش نمونه‌گیری غیر احتمالی هدفمند انتخاب شدند. معیارهای ورود عبارت بودند از: تحصیل در رشته پزشکی در نیمسال پنجم و بالاتر، نمره شناسه پژوهشی در حد مناسب، آشنایی اولیه با اصول روش تحقیق، متدولوژی پژوهش در علوم پزشکی و ساختار مقالات علمی. در ابتدای کارگاه فرم اطلاعات زمینه‌ای توسط هر یک از دانشجویان شرکت‌کننده تکمیل شد. سوالات شامل این موارد بود: نیمسال تحصیلی دانشجو، سابقه حضور در کارگاه آموزشی مشابه، سابقه داوری همتایان در کنگره‌های پژوهشی دانشجویی و یا مجلات علمی-پژوهشی، سابقه انتشار مقاله، تعداد مقالات پژوهشی در مجلات علمی-پژوهشی داخلی و یا مجلات بین‌المللی نمایه شده در pubmed/isi. در این کارگاه مطالبی در مورد چک‌لیست‌های استاندارد داوری مقالات (مانند consort و strobe) و نکات داوری اجزاء مختلف مقالات اصیل پژوهشی به دانشجویان ارائه می‌شد. شرکت‌کنندگان قبل و پس از برگزاری کارگاه، ۳ چکیده مقاله را با شبیه‌سازی ارزیابی کردند؛ یک چکیده مقاله کارآزمایی بالینی با کیفیت متوسط، یک چکیده مقاله توصیفی-مقطعی با کیفیت ضعیف و یک چکیده مقاله مقطعی تحلیلی با کیفیت بالا. با استفاده از چک‌لیست‌های استاندارد، قسمت‌های مختلف ۳ چکیده یک بار پیش از برگزاری کارگاه (پیش‌آزمون) و بار دیگر در انتهای کارگاه (پس‌آزمون) توسط دانشجویان نمره‌گذاری شد.

در این مطالعه از چک‌لیست‌های بومی شده و مورد تایید دبیرخانه دائمی کنگره‌های ملی و بین‌المللی تحقیقات دانشجویی دانشگاه‌های علوم پزشکی کشور استفاده شده است. این چک‌لیست‌ها با استناد به موارد مشابه استاندارد بین‌المللی شامل چک‌لیست consort (۹) برای ارزیابی نقادانه مطالعات مداخله‌ای و کارآزمایی

اصول عهدنامه هلسینکی (۱۱) در پژوهش پایبند بوده و اطلاعات شرکت کنندگان در مطالعه به صورت محرمانه حفظ شد.

آنالیز آماری

داده‌ها با استفاده از نرم‌افزار آماری SPSS v.17 (Chicago, IL, USA) تحلیل شد. در توصیف یافته‌ها از میانگین، میانه، مد، درصد فراوانی، محدوده، انحراف معیار و خطای معیار میانگین استفاده شد. در آنالیز تحلیلی داده‌ها، از شاخص‌های همبستگی برای ارزیابی پایایی نمرات بخش‌های مختلف چک‌لیست ارزیابی چکیده مقالات اصیل پژوهشی استفاده شد. برای این منظور، همبستگی اسپیرمن بین نمرات هر آیتم با کل نمره‌ی مربوط به آن حیطه، ثبات درونی سوالات پرسشنامه را نشان می‌داد. همچنین ضریب آلفای کرونباخ برای ارزیابی پایایی پرسشنامه محاسبه شد. بر اساس منابع موجود (۱۲)، ضریب آلفای کرونباخ بزرگتر از ۰/۷ معادل پایایی قابل قبول و بزرگتر از ۰/۹ بیانگر پایایی عالی در نظر گرفته شد. البته در برخی متون دیگر نیز مقادیر بالاتر از ۰/۶ بیانگر پایایی مورد قبول معرفی شده‌اند (۱۳). در تمامی تحلیل‌های آماری فرضیات دو طرفه در نظر گرفته شده و مقدار P-value کمتر از ۰/۰۵ نشانگر اختلاف یا ارتباط آماری معنی‌داری بود.

یافته‌ها

در این کارگاه آموزشی تعداد ۴۰ نفر از دانشجویان پزشکی شامل ۲۴ (۶۰٪) مرد و ۱۶ (۴۰٪) زن با میانگین سنی (SD=۳/۲) ۲۵/۶ سال شرکت داشتند. میانه نیم‌سال تحصیلی دانشجویان ۷ بود و ۶ (۱۵٪) نفر از آنان پیشینه داوری در کنگره‌های پژوهشی دانشجویان داشتند. جدول ۲ وضعیت متغیرهای دموگرافیک و زمینه‌ای دانشجویان مورد مطالعه در کارگاه آموزشی نخست را نشان می‌دهد.

بالینی و چک‌لیست strobe (۱۰) برای بررسی کیفیت مطالعات مشاهده‌ای شامل مطالعات مقطعی، مورد-شاهدی و هم‌گروهی طراحی و بومی‌سازی شده است. چک‌لیست طراحی شده در ۳ مرحله و توسط کارگروه‌های متفاوت مورد بررسی کارشناسان در حیطه کنگره‌های پژوهشی دانشجویی قرار گرفت. لازم به ذکر است این افراد در زمینه روش‌شناسی پژوهش در علوم پزشکی متخصص بودند. نسخه موجود در جدول ۱، پس از برگزاری جلسات انفرادی و گروهی با کارشناسان و به روش دلفی از نظر اعتبار ساختاری و محتوایی و با توجه به نظر متخصصین مورد تایید نهایی قرار گرفت. چک‌لیست نهایی در مجموع شامل ۳۱ سوال در حیطه‌های مختلف، برای ارزیابی چکیده مقالات اصیل پژوهشی می‌باشد که عبارتند از: عنوان (۱ سوال)، ساختار کلی چکیده (۱ سوال)، مقدمه (۳ سوال)، روش بررسی (۱۵ سوال)، نتایج (۴ سوال)، نتیجه‌گیری (۳ سوال)، واژگان کلیدی (۲ سوال) و نظر کلی (۲ سوال).

نحوه امتیازدهی بخش‌های مختلف که بر اساس اجماع نظر متخصصین تعیین شده است نیز بدین شرح می‌باشد: عنوان (۳ امتیاز)، ساختار کلی چکیده (۱ امتیاز)، مقدمه (۶ امتیاز)، روش بررسی (۲۷ امتیاز)، نتایج (۱۳ امتیاز)، نتیجه‌گیری (۱۲ امتیاز)، واژگان کلیدی (۳ امتیاز) و نظر کلی (۲۰ امتیاز). قابل ذکر است که در مورد چکیده مقالات مطالعات مشاهده‌ای، سوالات شماره ۶ الی ۱۱ موجود در چک‌لیست کاربرد نداشتند. در مجموع حداکثر امتیاز چکیده مقالات کارآزمایی بالینی ۸۵ مقالات مشاهده‌ای توصیفی ۷۵ و مشاهده‌ای تحلیلی ۸۰ می‌شد. همچنین دقت نمره‌دهی در تمامی موارد ۰/۵ واحد بوده است.

مطالعه حاضر مورد تایید معاونت تحقیقات و فناوری وزارت بهداشت، درمان و آموزش پزشکی بوده و طرح تحقیقاتی آن نیز به مصوب گروه آموزش پزشکی دانشگاه علوم پزشکی تهران می‌باشد. محققین به کلیه

جدول ۱: چک‌لیست مورد استفاده جهت ارزیابی چکیده مقالات اصیل پژوهشی

الف	ساختار چکیده، عنوان	ع
۱	آیا عنوان مناسبی دارد؟ (جامع، مانع، گویا و بیانگر هدف است)	۳
۲	آیا چکیده در قالب ساختار خواسته شده است؟ (مقدمه، روش بررسی، نتایج، نتیجه گیری)	۱
ب	مقدمه	۶
۳	آیا در مقدمه background خوبی از موضوع مورد مطالعه بیان شده است؟	۲
۴	آیا اهمیت و ضرورت انجام مطالعه به خوبی بیان شده؟	۲
۵	آیا فرضیه و پرسش مطالعه برای خواننده روشن است؟ و آیا نوع طراحی این مطالعه قادر به پاسخ گویی به این پرسش می باشد؟	۲
پ	روش بررسی	۲۷
I	طراحی مطالعه	۱۴
۶	آیا نوع مطالعه به درستی بیان گردیده است؟ (یا روش مورد استفاده در Lab Research)	۲
۷	آیا به زمان و مکان مطالعه اشاره شده است؟ (در مطالعات تجربی، حیوانی و آزمایشگاهی می‌تواند ذکر نشود)	۱
۸	آیا جمعیت مورد مطالعه به خوبی تعریف شده است؟ (معیارهای پذیرش و عدم پذیرش و یا معیارهای واجد شرایط بودن) (و یا نوع میکروب/سلول و محیط کشت در Lab Research)	۲
۹	آیا روش نمونه‌گیری ذکر شده است؟ آیا این روش متناسب با هدف مطالعه انتخاب شده است؟	۲
۱۰	آیا حجم نمونه (Sample Size) در کل و به تفکیک گروه‌های مطالعاتی (در صورت وجود) بیان شده است؟ و آیا به تعداد کافی انتخاب شده است؟	۲
۱۱	آیا در مطالعات Cohort و case-control، عامل مواجهه (Exposure) به خوبی بیان شده است؟ آیا در مطالعات مداخله‌ای نحوه مداخله به وضوح بیان شده است؟	۲
۱۲	آیا به زمان پیگیری مطالعه اشاره شده است؟ آیا این زمان متناسب با هدف مطالعه انتخاب شده است؟	۲
۱۳	آیا کد ثبت مطالعه RCT ذکر شده است؟ (Trial Registry Number، صرفاً در مطالعات کارآزمایی بالینی)	۱
II	گمارش (فقط در مورد مطالعات Analytical و RCT)	۵
۱۴	آیا نحوه اختصاص افراد به گروه‌های مطالعه و شاهد مناسب بوده است؟ (نحوه تصادفی سازی در مطالعات RCT بیان شده است؟) یا (در مطالعات تحلیلی گمارش مشاهده‌ای بدون bias در انتخاب بوده است؟)	۲
۱۵	آیا تفاوت‌هایی بین گروه مطالعه و شاهد بجز خصوصیت تحت بررسی وجود دارد که بتواند نتایج را تحت تاثیر قرار دهد (متغیر مخدوش کننده) و در صورت وجود آیا راهکاری برای حل این مشکل در نظر گرفته است؟ (مثلاً Matching و ...)	۲
۱۶	در صورت نیاز به انجام Blinding آیا به درستی صورت گرفته است؟	۱
III	ارزیابی	۸
۱۷	آیا متغیرهای مورد سنجش به خوبی بیان شده و متناسب با نتایج ارائه شده است؟	۲
۱۸	آیا پیامد‌های اصلی مطالعه بیان گردیده و به درستی سنجیده شده است؟ (متناسب با هدف بوده، از دقت کافی برخوردار بوده و بدون تاثیر پذیری از مشاهده است)	۲
۱۹	آیا روش‌های آماری مورد استفاده برای برآورد وسعت یا قوت ارتباط یا وابستگی‌های مشاهده شده در پژوهش بیان شده است؟	۲
۲۰	کیفیت کلی بیان روش کار و روانی مطلب	۲

در مورد دو چکیده مربوط به مطالعات مشاهده‌ای، ضریب همبستگی اسپیرمن در آیت‌های مربوط به بخش "نتایج" و "روش بررسی" پس از برگزاری کارگاه آموزشی در پس-آزمون نسبت به پیش‌آزمون افزایش داشته است. برای نمونه ضریب همبستگی آیت‌های مربوط به "حجم نمونه" در ارزیابی چکیده‌ی شماره ۱ از ۰/۱۹۵ ($P=0/254$) در پیش‌آزمون به ۰/۴۴۵ ($P=0/005$) در پس‌آزمون افزایش یافته است. همچنین ضریب همبستگی آیت‌های ارزیابی "روش‌های آماری" در ارزیابی چکیده‌ی شماره ۲ از

جدول ۳ نتایج بررسی همسانی درونی آیت‌های مختلف موجود در چک‌لیست ارزیابی چکیده مقالات اصیل پژوهشی با محاسبه‌ی همبستگی اسپیرمن میان هر آیت با مجموع امتیاز هر حیطة را در ارزشیابی سه چکیده مختلف نقد شده توسط دانشجویان در پیش‌آزمون و پس-آزمون در این کارگاه آموزشی نشان می‌دهد. همان‌طور که مشاهده می‌شود در بیشتر موارد همسانی درونی معنی‌داری بین آیت‌های مختلف چک‌لیست و امتیاز مجموع آن بخش گزارش شده است. به‌ویژه

ادامه جدول ۱

۱۳	.	یافته‌ها	ت
۲	.	بیان مناسب اطلاعات زمینه ای (Baseline characteristics)	۲۱
۲	.	ترتیب گزارش نتایج (توصیفی بعد در صورت وجود تحلیلی اولیه و پیشرفته) روش آنالیز و بیان شاخص های مرکزی و پراکندگی مثل میانگین و انحراف معیار (در آنالیز Descriptive) و بیان شاخص های مناسب در آنالیز های تحلیلی شامل شاخص های معنی داری و شدت ارتباط (P-value یا power و یا ضریب همبستگی و CI ذکر شده باشد).	۲۲
۶	.	آیا نتایج ذکر شده همراستا با اهداف مطالعه می باشد؟	۲۳
۳	.	نتیجه گیری	۲۴
۱۲	.	تفسیر	ث
۸	.	آیا نتیجه نهایی هدف اصلی را بیان کرده است؟	I
۴	.	آیا تفسیر اطلاعات بدست آمده متناسب با یافته ها بوده و با توجه به نوع مطالعه منطقی است؟ (مثلاً در صورت ذکر رابطه علت و معلولی آیا سه معیار ارتباط، تقدم ارتباط، تغییر معلول بدنبال علت در مطالعه اثبات می شود؟)	۲۵
۴	.	تعمیم	۲۶
۴	.	تعمیم نتایج مطالعه به گروه های مشابه و یا جمعیت های دیگر به درستی و با در نظر گرفتن مفروضات لازم انجام گرفته است؟ (مثلاً اغراق آمیز و به دور از یافته های مطالعه نباشد).	II
۳	.	واژگان کلیدی	۲۷
۲	.	آیا واژگان کلیدی مرتبط می باشد؟	ج
۱	.	آیا واژگان کلیدی بین ۳ تا ۱۰ عدد می باشد؟	۲۸
۲۰	.	نظر کلی داور	۲۹
۵	.	آیا پیوستگی لازم در بخشهای مختلف وجود داشت؟	ج
۱۵	.	نظر کلی داور در مورد این چکیده مقاله (ابتکار داشتن، جذاب و بدیع بودن، ...)	۳۰
۸۵	.	مجموع امتیازات	۳۱

جدول ۲: وضعیت متغیرهای دموگرافیک و زمینه‌ای در دانشجویان (n=۴۰) مورد مطالعه در کارگاه آموزشی

جنسیت تعداد (درصد)	
زن	۱۶ (۴۰٪)
مرد	۲۴ (۶۰٪)
سن (سال)	۲۵/۶ (۳/۲۰)
میانگین (انحراف معیار)	
نیمسال تحصیلی	۷ (۶)
میان (فاصله بین چارگی)	
سابقه قبلی داوری در کنگره‌های پژوهشی	۶ (۱۵٪)
تعداد (درصد)	
سابقه قبلی داوری برای مجلات علمی-پژوهشی	۲ (۵٪)
تعداد (درصد)	
نمره آزمون	۶/۷ (۳/۲)
میانگین (انحراف معیار)	
نمره شناسه علمی-پژوهشی	۵۸/۲ (۲۴/۵)
میانگین (انحراف معیار)	

در تمامی بخش‌های اصلی چکیده مقاله (IMRAD) و حیطه نظر کلی -هم در پیش‌آزمون و هم در پس‌آزمون- همسانی درونی از نظر آماری معنی‌دار گزارش شده‌است ($P < 0/05$). در چکیده مقاله کارآزمایی بالینی، بیشترین ضریب همبستگی همسانی درونی مربوط به بخش "روش بررسی" در پس‌آزمون می‌باشد ($P < 0/001$), Spearman rho = 0/860. نتایج بررسی پایایی کلی چک-

۰/۳۴۸ ($P = 0/030$) در پیش‌آزمون به ۰/۵۳۷ ($P < 0/001$) در پس‌آزمون ارتقا یافته‌است.

جدول ۴ نتایج بررسی همسانی درونی حیطه‌های مختلف موجود در چکلیست ارزیابی چکیده مقالات اصیل پژوهشی با محاسبه همبستگی اسپیرمن میان مجموع نمره هر حیطه با امتیاز کل چکلیست را در ارزشیابی سه چکیده مختلف در پیش‌آزمون و پس‌آزمون نشان می‌دهد.

جدول ۳: بررسی همسانی درونی (Internal Consistency) آیت‌های مختلف موجود در چک‌لیست ارزیابی چکیده مقالات اصیل پژوهشی با محاسبه‌ی همبستگی اسپیرمن میان هر آیت‌ها با مجموع امتیاز هر حیطة (Domain) توسط دانشجویان (n=۴۰) مورد مطالعه در پیش‌آزمون (Pre-Test) و پس‌آزمون (Post-test)

حیطه	شماره	آیتم	مطالعه مقطعی-توصیفی با کیفیت ضعیف				مطالعه مقطعی-تحلیلی با کیفیت بالا				مطالعه کارآزمایی بالینی با کیفیت متوسط			
			پیش‌آزمون	پس‌آزمون	پس‌آزمون	پیش‌آزمون	پیش‌آزمون	پس‌آزمون	پس‌آزمون	پیش‌آزمون	پیش‌آزمون	پس‌آزمون	پس‌آزمون	
			P-value	r Spearman	P-value	r Spearman	P-value	r Spearman	P-value	r Spearman	P-value	r Spearman	P-value	r Spearman
ساختار	T.1	عنوان	<./۰۰۱	۱	<./۰۰۱	۱	<./۰۰۱	۱	<./۰۰۱	۱	<./۰۰۱	۱	<./۰۰۱	۱
چکیده، عنوان	T.2	ساختار	<./۰۰۱	۱	<./۰۰۱	۱	<./۰۰۱	۱	۰/۰۳۹	۰/۳۳	<./۰۰۱	۱	<./۰۰۱	۱
مقدمه	I.1	پیش‌زمینه و معرفی	<./۰۰۱	۱	<./۰۰۱	۱	<./۰۰۱	۱	۰/۶۰	۰/۶۰	<./۰۰۱	۱	<./۰۰۱	۱
	I.2	ضرورت و اهمیت	<./۰۰۱	۱	<./۰۰۱	۱	<./۰۰۱	۱	۰/۷۳	۰/۷۳	<./۰۰۱	۱	<./۰۰۱	۱
	I.3	اهداف و فرضیه	<./۰۰۱	۱	<./۰۰۱	۱	<./۰۰۱	۱	۰/۶۳	۰/۶۳	<./۰۰۱	۱	<./۰۰۱	۱
	M.1	طراحی	<./۰۰۱	۱	<./۰۰۱	۱	<./۰۰۱	۱	۰/۳۷	۰/۳۷	<./۰۰۱	۱	<./۰۰۱	۱
	M.2	زمان و مکان	<./۰۰۱	۱	<./۰۰۱	۱	<./۰۰۱	۱	۰/۳۷	۰/۳۷	<./۰۰۱	۱	<./۰۰۱	۱
	M.3	معیارهای ورود	<./۰۰۱	۱	<./۰۰۱	۱	<./۰۰۱	۱	۰/۴۴	۰/۴۴	<./۰۰۱	۱	<./۰۰۱	۱
	M.4	روش نمونه‌گیری	<./۰۰۱	۱	<./۰۰۱	۱	<./۰۰۱	۱	۰/۵۵	۰/۵۵	<./۰۰۱	۱	<./۰۰۱	۱
	M.5	حجم نمونه	<./۰۰۱	۱	<./۰۰۱	۱	<./۰۰۱	۱	۰/۵۵	۰/۵۵	<./۰۰۱	۱	<./۰۰۱	۱
	M.6	مداخله	<./۰۰۱	۱	<./۰۰۱	۱	<./۰۰۱	۱	۰/۴۴	۰/۴۴	<./۰۰۱	۱	<./۰۰۱	۱
	M.7	پیگیری	<./۰۰۱	۱	<./۰۰۱	۱	<./۰۰۱	۱	۰/۴۴	۰/۴۴	<./۰۰۱	۱	<./۰۰۱	۱
	M.8	کد ثبت کارآزمایی بالینی	<./۰۰۱	۱	<./۰۰۱	۱	<./۰۰۱	۱	۰/۴۴	۰/۴۴	<./۰۰۱	۱	<./۰۰۱	۱
	M.9	جایگذاری	<./۰۰۱	۱	<./۰۰۱	۱	<./۰۰۱	۱	۰/۴۴	۰/۴۴	<./۰۰۱	۱	<./۰۰۱	۱
	M.10	جورسازی	<./۰۰۱	۱	<./۰۰۱	۱	<./۰۰۱	۱	۰/۴۴	۰/۴۴	<./۰۰۱	۱	<./۰۰۱	۱
	M.11	کورسازی	<./۰۰۱	۱	<./۰۰۱	۱	<./۰۰۱	۱	۰/۴۴	۰/۴۴	<./۰۰۱	۱	<./۰۰۱	۱
	M.12	متغیرها	<./۰۰۱	۱	<./۰۰۱	۱	<./۰۰۱	۱	۰/۵۷	۰/۵۷	<./۰۰۱	۱	<./۰۰۱	۱
M.13	پیامد اصلی	<./۰۰۱	۱	<./۰۰۱	۱	<./۰۰۱	۱	۰/۳۸	۰/۳۸	<./۰۰۱	۱	<./۰۰۱	۱	
M.14	روش‌های آماری	<./۰۰۱	۱	<./۰۰۱	۱	<./۰۰۱	۱	۰/۳۱	۰/۳۱	<./۰۰۱	۱	<./۰۰۱	۱	
M.15	پیوستگی کلی	<./۰۰۱	۱	<./۰۰۱	۱	<./۰۰۱	۱	۰/۲۱	۰/۲۱	<./۰۰۱	۱	<./۰۰۱	۱	
R.1	توصیفات زمینه‌ای	<./۰۰۱	۱	<./۰۰۱	۱	<./۰۰۱	۱	۰/۳۳	۰/۳۳	<./۰۰۱	۱	<./۰۰۱	۱	
R.2	ترتیب گزارش	<./۰۰۱	۱	<./۰۰۱	۱	<./۰۰۱	۱	۰/۵۶	۰/۵۶	<./۰۰۱	۱	<./۰۰۱	۱	
R.3	تحلیل‌های آماری	<./۰۰۱	۱	<./۰۰۱	۱	<./۰۰۱	۱	۰/۷۶	۰/۷۶	<./۰۰۱	۱	<./۰۰۱	۱	
R.4	همراستایی با اهداف	<./۰۰۱	۱	<./۰۰۱	۱	<./۰۰۱	۱	۰/۵۹	۰/۵۹	<./۰۰۱	۱	<./۰۰۱	۱	
نتیجه گیری	C.1	نتیجه‌گیری نهایی	<./۰۰۱	۱	<./۰۰۱	۱	<./۰۰۱	۱	۰/۸۷	۰/۸۷	<./۰۰۱	۱	<./۰۰۱	۱
	C.2	تفسیر	<./۰۰۱	۱	<./۰۰۱	۱	<./۰۰۱	۱	۰/۶۱	۰/۶۱	<./۰۰۱	۱	<./۰۰۱	۱
	C.3	تعمیم	<./۰۰۱	۱	<./۰۰۱	۱	<./۰۰۱	۱	۰/۷۴	۰/۷۴	<./۰۰۱	۱	<./۰۰۱	۱
واژگان کلیدی	K.1	تعداد	<./۰۰۱	۱	<./۰۰۱	۱	<./۰۰۱	۱	۰/۹۵	۰/۹۵	<./۰۰۱	۱	<./۰۰۱	۱
	K.2	مناسبت	<./۰۰۱	۱	<./۰۰۱	۱	<./۰۰۱	۱	۰/۱۴	۰/۱۴	<./۰۰۱	۱	<./۰۰۱	۱
نظر کلی داور	O.1	پیوستگی کلی	<./۰۰۱	۱	<./۰۰۱	۱	<./۰۰۱	۱	۰/۷۳	۰/۷۳	<./۰۰۱	۱	<./۰۰۱	۱
	O.2	نوآوری، ارزش علمی	<./۰۰۱	۱	<./۰۰۱	۱	<./۰۰۱	۱	۰/۹۶	۰/۹۶	<./۰۰۱	۱	<./۰۰۱	۱

کارآزمایی بالینی، ضریب پایایی آلفای کرونباخ در بخش "روش بررسی" عالی (۰/۹۲۶) و در دو بخش "نتایج" (۰/۶۵۰) و "نتیجه‌گیری" (۰/۷۳۹) در پیش‌آزمون قابل-قبول بود.

بحث

پژوهش حاضر یکی از معدود مطالعاتی است که به بررسی روایی و پایایی چک‌لیست ارزیابی چکیده

لیست با روش محاسبه ضریب آلفای کرونباخ در ارزشیابی سه چکیده مورد بررسی توسط دانشجویان ضرایب را به ترتیب زیر نشان داد:

ضرایب ۰/۷۲ و ۰/۷۹ در ارزیابی پیش‌آزمون و پس‌آزمون چکیده شماره ۱، ضرایب ۰/۶۷ و ۰/۶۹ در ارزیابی پیش‌آزمون و پس‌آزمون چکیده شماره ۲ و ضرایب ۰/۵۹ و ۰/۶۷ در ارزیابی پیش‌آزمون و پس‌آزمون چکیده شماره ۳ همچنین در ارزیابی چکیده مطالعه

جدول ۴: بررسی همسانی درونی (Internal Consistency) حیطه‌های (Domain) مختلف موجود در چک لیست ارزیابی چکیده مقالات اصیل پژوهشی با محاسبه همبستگی اسپیرمن میان هر حیطه با مجموع امتیاز کل در ارزشیابی سه چکیده توسط دانشجویان (n=۴۰) مورد مطالعه در پیش آزمون (Pre-Test) و پس آزمون (Post-test)

پس آزمون		پیش آزمون		حیطه	شماره چکیده
P-value	r Spearman	P-value	r Spearman		
۰/۰۰۸	۰/۴۲	۰/۴۷۹	۰/۱۲	ساختار چکیده، عنوان	چکیده شماره ۱ (مطالعه مقطعی- توصیفی با کیفیت ضعیف)
<۰/۰۰۱	۰/۶۳	<۰/۰۰۱	۰/۶۰	مقدمه	
<۰/۰۰۱	۰/۷۲	<۰/۰۰۱	۰/۵۷	روش بررسی	
<۰/۰۰۱	۰/۸۲	<۰/۰۰۱	۰/۶۵	یافته‌ها	
<۰/۰۰۱	۰/۷۷	<۰/۰۰۱	۰/۸۳	نتیجه گیری	
۰/۷۰۲	-۰/۰۶	۰/۱۷۹	۰/۲۲	واژگان کلیدی	چکیده شماره ۲ (مطالعه مقطعی- تحلیلی با کیفیت بالا)
<۰/۰۰۱	۰/۷۹	<۰/۰۰۱	۰/۸۷	نظر کلی داور	
۰/۰۰۸	۰/۴۲	۰/۰۸۴	۰/۲۸	ساختار چکیده، عنوان	
۰/۰۰۲	۰/۴۸	<۰/۰۰۱	۰/۵۴	مقدمه	
<۰/۰۰۱	۰/۶۹	<۰/۰۰۱	۰/۶۶	روش بررسی	
<۰/۰۰۱	۰/۶۱	<۰/۰۰۱	۰/۶۲	یافته‌ها	چکیده شماره ۳ (مطالعه کارآزمایی بالینی با کیفیت متوسط)
<۰/۰۰۱	۰/۷۱	<۰/۰۰۱	۰/۷۳	نتیجه گیری	
۰/۷۳۹	۰/۰۶	۰/۲۹۸	۰/۱۷	واژگان کلیدی	
<۰/۰۰۱	۰/۷۴	<۰/۰۰۱	۰/۷۴	نظر کلی داور	
۰/۰۰۹	۰/۴۲	۰/۱۹۸	۰/۲۱	ساختار چکیده، عنوان	
<۰/۰۰۱	۰/۶۶	۰/۰۰۱	۰/۵۱	مقدمه	چکیده شماره ۳ (مطالعه کارآزمایی بالینی با کیفیت متوسط)
<۰/۰۰۱	۰/۸۶	<۰/۰۰۱	۰/۷۱	روش بررسی	
<۰/۰۰۱	۰/۵۷	<۰/۰۰۱	۰/۶۳	یافته‌ها	
<۰/۰۰۱	۰/۶۴	<۰/۰۰۱	۰/۷۱	نتیجه گیری	
۰/۹۱۰	-۰/۰۲	۰/۰۰۹	۰/۴۱	واژگان کلیدی	
<۰/۰۰۱	۰/۶۱	<۰/۰۰۱	۰/۵۸	نظر کلی داور	

از چک لیست‌های استاندارد در ارزیابی نقادانه شواهد علوم پزشکی از جمله داوری هم‌تایان می‌باشد. به عبارتی پس از برگزاری کارگاه، دانشجویان با شناخت بیشتری نسبت به آیتم‌های مورد پرسش مقالات را ارزیابی کردند. در بررسی شاخص‌های استانداردسازی پرسش-نامه، بالاترین شاخص همسانی درونی در بررسی چکیده مقاله کارآزمایی بالینی مربوط به بخش "روش بررسی" بود که این نکته اهمیت توجه به فاکتورهای روش‌شناسی در مطالعات مداخله‌ای را نشان می‌دهد. در بررسی نتایج آنالیز پایایی پرسشنامه در این پژوهش می‌بایست به تغییر دانش، نگرش و عملکرد دانشجویان نسبت به ارزیابی نقادانه چکیده‌ها پس از برگزاری کارگاه، توجه ویژه داشت. این امر سبب می‌شود که برخی از شاخص‌های مربوط به پایایی همچون همبستگی درون-گروهی قابل

مقالات پرداخته است. با توجه به بررسی متون انجام شده، تاکنون در ایران چنین پژوهشی انجام نشده و مطالعات بین‌المللی مشابه نیز بیشتر در گروه اساتید، اعضای هیات علمی و یا داوران حرفه‌ای مجلات علوم پزشکی صورت گرفته است. با توجه به نتایج مطالعه انجام شده، پرسش-نامه مورد استفاده در این پژوهش برای ارزیابی چکیده مقالات اصیل پژوهشی در همه حیطه‌های ساختاری مقاله (IMRaD) از پایایی معنادار آماری برخوردار بود. نکته قابل توجه این است که پس از برگزاری کارگاه و تشریح آیتم‌های موجود در پرسش‌نامه برای دانشجویان، شاخص‌های مربوط به پایایی و تکرارپذیری سوالات چک لیست به‌طور قابل توجهی در پس‌آزمون افزایش یافته است. این نکته نشان دهنده اهمیت شفاف سازی و ارائه راهنما و یا گایدلاین‌هایی برای توضیح نکات لازم و نحوه استفاده

گزارش نباشد؛ چرا که بسیاری از دانشجویان با توجه به مطالب ارائه شده در کارگاه سعی در تغییر نحوه ارزیابی خود با استفاده از چک‌لیست‌ها داشته‌اند که خود سبب کاهش تکرارپذیری نتایج در بعضی از آیت‌ها در طی آزمون مجدد می‌شد. در مجموع می‌توان چنین نتیجه گرفت که چک‌لیست مورد استفاده در پژوهش حاضر یک پرسشنامه ساختاری شامل آیت‌های مهم در ارزیابی قسمت‌های مختلف مقاله بوده و بر اساس چک‌لیست‌های ارزیابی نقادانه‌ی استاندارد تهیه شده و قابلیت استفاده در داوری همتایان را دارد. در هنگام استفاده از این چک‌لیست‌ها توجه داوران به نکات مهم برای ارزیابی هر یک از قسمت‌های مقاله شامل مقدمه، روش بررسی، یافته‌ها و نتیجه‌گیری معطوف می‌شود. در مقابل، بسیاری از مجلات و کنگره‌های علمی-پژوهشی از پرسشنامه‌هایی با سوالات باز و بدون اشاره به جزئیات آیت‌های موثر در کیفیت هر بخش مقاله استفاده می‌کنند. واضح است در چنین مواردی احتمال عدم توجه به بسیاری از نکات روش‌شناسی پژوهش در قسمت‌های مختلف مقاله افزایش می‌یابد که به نوبه خود منجر به کاهش کیفیت داوری همتایان می‌شود.

تاکنون در چند مطالعه مشابه نیز ابزارهای متفاوتی در قالب چک‌لیست برای ارزیابی مقالات اصیل پژوهشی معرفی شده‌اند که بیشتر به بررسی متن کامل مقالات مربوط بوده و کمتر مورد آنالیز آماری برای بررسی دقیق روایی و پایایی قرار گرفته‌اند. در سال ۱۹۹۵، رانگاکچاری و میرسون (۱۴) چک‌لیست خلاصه‌ای را برای ارزیابی بهتر مقالات منتشر شده توسط دانشجویان علوم پزشکی طراحی کردند. هدف نهایی این ابزار بررسی و ارزیابی نقادانه پژوهش‌های منتشر شده بود و علی‌رغم شباهت برخی از آیت‌ها، کنگره‌های پژوهشی و داوری همتایان را در نظر نگرفته بود. در سال ۲۰۰۰، سیلز و تاناکا (۱۵) چک‌لیست کامل‌تری را برای استفاده در داوری همتایان مقالات پژوهشی اصیل ارسال شده به مجلات علوم پزشکی ارائه کردند. مشابه مطالعه حاضر، در این چک-

لیست نیز آیت‌های ساختاری بخش‌های مختلف ساختار IMRaD مورد توجه قرار گرفته بود و در انتها نیز سوالاتی در مورد نظر عمومی و کلی داور در مورد مقاله مورد بررسی مطرح شده بود. مانند آنچه در چک‌لیست مطالعه حاضر مد نظر قرار گرفته، به نظر می‌رسد در کنار آیت‌های جزء به جزء ساختاری، وجود سوالاتی برای اخذ نظر کلی داوران حائز اهمیت می‌باشد. از طرف دیگر یکی از مشکلاتی که سیلز و تاناکا (۱۵) به‌عنوان نقطه ضعف چک‌لیست‌های ارزیابی موجود به آن اشاره می‌کنند، کیفی بودن این‌گونه پرسشنامه‌هاست؛ حال آن‌که یکی از مزایای چک‌لیست پیشنهادی در مطالعه ما وجود سیستم امتیازدهی و کمی شدن نمره‌ی چکیده مقالات مورد بررسی است که کاربردهای خاص خود را دارد. در سال ۲۰۰۸، دوچسنه و یانین (۱۶) چک‌لیست مفصلی شامل ۷۱ آیت را برای داوری متن کامل مقالات ارسالی برای مجلات علوم پزشکی معرفی کردند. در این چک‌لیست نیمه‌کمی آیت‌های مدنظر به سه حالت منفی، صفر (یا خنثی) و مثبت ارزش‌گذاری شد (۱۶). به نظر می‌رسد سیستم کاملاً کمی مورد استفاده در چک‌لیست حاضر، می‌تواند برای فضای رقابتی همچون انتخاب تعداد اندکی مقاله برای ارائه در یک کنگره از بین تعداد زیاد چکیده‌های ارسالی مناسب‌تر باشد.

اخیراً در یک مقاله مروری منتشر شده در سال ۲۰۱۲ نکاتی بر اساس بررسی شواهد موجود برای بهبود کیفیت ارزیابی نقادانه شواهد به داوران توصیه شده است که یکی از مهم‌ترین این موارد پرهیز از خطاهای سیستماتیک ذهنی همچون نگرش منفی داوران نسبت به مقالات موسوم به "نتایج منفی" می‌باشد (۱۷). به نظر می‌رسد استفاده از چک‌لیست ساختاری پژوهش حاضر می‌تواند باعث کاهش نگرش‌های منفی در داوران باشد که این مساله به‌ویژه در مهارت‌پذیری دانشجویان علوم پزشکی در مباحثی چون ارزیابی نقادانه و تفکر انتقادی علمی از اهمیتی دو چندان برخوردار است. از سوی دیگر،

برخی از شواهد منتشر شده عنوان کرده اند که کیفیت داوری همتایان در متخصصین با تجربه از پژوهشگران جوانتر پایین تر است. دلایلی برای توجیه کیفیت بهتر ارزیابی در داوران جوانتر ذکر شده است؛ معمولاً این داوران اشتیاق بیشتری برای داوری دارند و تعداد بیشتری از مقالات ارسالی را بررسی می کنند. این افراد مشتاق هستند که توسط سردبیر مجلات شناخته شوند و معمولاً به تازگی در مورد طراحی مطالعه، روش های آماری و اصول مقاله نویسی آموزش دیده اند. به علاوه داوران با سابقه بیشتر گاهی اشتیاق کمتری برای داوری دارند؛ چرا که آن-چه داور طی داوری یاد می گیرد و به دست می آورد، کمتر از وقت و انرژی است که صرف می کند و درعین حال احساس می کنند که این مقاله شبیه دست نوشته های قبلی است (۱۸).

نوالتا و همکاران (۸) در پژوهشی به مقایسه کیفیت داوری همتایان مقالات بین دو گروه دانشجویان علوم پزشکی و اعضای هیات علمی پرداختند. نتایج این مطالعه نشان داد، بین کیفیت ارزیابی گروه دانشجویان با گروه متخصصین هیات علمی تفاوت آماری معناداری وجود ندارد. یکی از دلایلی که پژوهشگران برای توجیه این یافته ها به آن استناد کرده اند وجود راهنماها و گایدلاین های آموزشی برای داوران دانشجو در مجله مورد بررسی بود که نشان می دهد در صورت استفاده از چک لیست های مستند برای ارزیابی نقادانه می توان کیفیت ارزشیابی نهایی را حتی در گروه دانشجویان ارتقاء داد.

یکی از مشکلات اجرایی و محدودیت های پژوهش حاضر خطای سیستماتیک در نمونه گیری می باشد. در این مطالعه از نمونه گیری غیر احتمالی و افراد برگزیده استفاده شد؛ به طوری که کلیه دانشجویان شرکت کننده ملزم به گذراندن کارگاه های روش تحقیق عمومی و مقاله نویسی پیش از شرکت در این کارگاه بودند. با توجه به هدف و ماهیت برگزاری این کارگاه ها،

در عمل امکان استفاده از روش های تصادفی و احتمالی برای انتخاب شرکت کنندگان از میان تمامی دانشجویان علوم پزشکی وجود نداشت. با توجه به این مساله در تعمیم پذیری نتایج این مطالعه می بایست با احتیاط بیشتری عمل کرد. به عبارتی نتایج این پژوهش در استفاده از چک لیست پیشنهادی بیشتر به گروهی از دانشجویان که آشنایی بیشتری با پژوهش و تحقیق دارند برمی گردد. یکی دیگر از مواردی که می بایست در استفاده از چک لیست پیشنهادی در این مطالعه مدنظر قرار گیرد، حذف برخی از سوالات چک لیست هایی همچون consort و strobe می باشد. از آنجایی که هدف نهایی از طراحی و معرفی چک لیست مطالعه حاضر، ارزیابی و داوری چکیده مقالات اصیل پژوهشی می باشد؛ لذا برخی از آیتم هایی که انتظار می رود صرفاً در متن کامل مقالات بدان ها اشاره شود، در این جا حذف شده اند. مسلماً در چکیده مقالات پژوهشی با در نظر گرفتن محدودیت تعداد لغات امکان بیان تمامی نکات اصلی پژوهش وجود ندارد و تنها اشاره به مهم ترین موارد کفایت می کند. در مجموع، مطالعه حاضر یکی از معدود پژوهش هایی است که به معرفی یک ابزار استاندارد برای ارزیابی چکیده مقالات اصیل پژوهشی ارسال شده برای کنگره های علوم پزشکی پرداخته است و سوالات مفصل و کامل چک لیست های استاندارد فعلی را برای ارزیابی چکیده مقالات ارسالی به کنگره ها خلاصه سازی و بومی کرده است. با توجه به استقبال گسترده و حضور فعال دانشجویان علوم پزشکی در کنگره های پژوهشی دانشجویی در قالب داور، می توان انتظار داشت که استفاده از چک لیست های ساختاری استاندارد شده با امکان ثبت نمره کمی، می تواند در بهبود کیفیت ارزیابی چکیده مقالات و داوری همتایان در کنگره های علوم پزشکی به ویژه کنفرانس های دانشجویی مفید واقع شود.

References

1. Parkes J, Hyde C, Deeks J, Milne R. Teaching critical appraisal skills in healthcare settings. *Cochrane*

- Database Syst Rev, 2011;3.
2. Glowniak JV. Medical resources on the Internet. *Ann Intern Med* 1995; 123(2): 123-31.
 3. Glowniak J, Bushway Mk. Computer networks as a medical resource: accessing and using the Internet. *JAMA*.1994; 272: 1934-9.
 4. Goldwein JW, Benjamin I. Internet-based medical information: time to take charge. *Ann Intern Med*, 1995; 123: 152-3.
 5. Brown J, Libberton P. *Principles of Professional Studies in Nursing*, Hampshire, Palgrave Macmilla. 2007.
 6. Gough NR. Training for peer review. *Sci. Signal* 2009; 2.(85): tr2.
 7. Bordage G, Caelleigh A. A tool for reviewers: review criteria for research manuscripts. *Acad Med*. 2001; 76(9):904-8.
 8. Navalta JW, Lyons TS. Student peer review decisions on submitted manuscripts are as stringent as faculty peer reviewers. *Adv Physiol Educ*. 2010; 34: 170-3.
 9. Schulz KF, Altman DG, Moher D. CONSORT 2010 Statement: Updated guidelines for reporting parallel group randomised trials. *BMJ*. 2010; 63: 834-40.
 10. Knottnerus A, Tugwell P. STORE--a checklist to Strengthen the Reporting of Observational Studies in Epidemiology. *J Clin Epidemio.*, 2008; 61(4): 323.
 11. World Medical Association. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. *J Postgrad Me*. 2002; 48: 206-8.
 12. George D, Mallery P. *SPSS for Windows step by step: A simple guide and reference.*, Boston, Allyn & Bacon. 2003.
 13. Gerevich J, Bacskai E, Ko J, Rozsa S. Reliability and validity of the Hungarian version of the European Addiction Severity Index. *Psychopathology*. 2005; 38(6): 301-9.
 14. Rangachari PK, Mierson S. A checklist to help students analyze published articles in basic medical sciences. *Am J Physiol Adv Physiol*. 1995; 268(6Pt 3): S21-5.
 15. Seals DR, Tanaka H. Manuscript peer review: a helpful checklist for students and novice referees. *Adv Physiol Educ*. 2000; 23(1): 52-8.
 16. Duchesne S, Jannin P. Proposing a manuscript peer-review checklist. *Neuroimage*. 2008; 39(4): 1783-7.
 17. Azer SA, Ramani S, Peterson R. Becoming a peer reviewer to medical education journals. *Med Teach*. 2012; 34(9): 698-704.
 18. Kliewer MA, Freed KS, DeLong DM, Pickhardt PJ, Provenzale JM. Reviewing the reviewers: comparison of review quality and reviewer characteristics at the American Journal of Roentgenology. *AJR Am J Roentgenol*.2005; 184(6): 1731-5

Archive of SID

Validity and reliability of a checklist for critical appraisal of the original research abstracts by student peer reviewers in medical student research congresses

Seyed-Mohammad Fereshtehnejad

Researcher, Medical Science Department, Virtual Faculty, Tehran University of Medical Sciences.

Hamid Reza Baradaran,

Associate Professor, Medical Science Department, Tehran University of Medical Sciences.

Maziar Moradi Lakeh

Associate Professor of Social Medicine, Tehran University of Medical Sciences.

Received:19/05/2013, Revised:06/07/2013, Accepted:30/07/2013

Corresponding Author:

Seyed-Mohammad Fereshtehnejad,
Tehran University of Medical
Sciences.
E-mail: sm.fereshtehnejad@ki.se

Abstract

Background: Nowadays, it is of utmost importance to critically appraise the research evidences presented in scientific congresses by the audiences. In addition to improvement in scientific and practical skills of critical appraisal, it is important to use a standard framework as the major tool for peer reviewing. We aimed to assess the validity and reliability of one proposed checklist for critical appraisal of the original research abstracts by student peer reviewers.

Methods and Materials: This study was a part of an educational interventional project that was performed in a setting of a workshop where 40 medical students from the medical faculties of the universities in Tehran were recruited. Participants were selected using a non-probability purposive sampling method. Educational curriculum of the workshop included clarifying explanations on the 31-item checklist for peer reviewing of the abstracts as well as several tips about each item using lectures, simulations and group discussions during 10 hours. Medical students used the checklist twice, at beginning and the end of the workshop to score three sample abstracts. Data were collected and analyzed using Spearman correlation (internal consistency) and Cronbach's' alpha methods to calculate the reliability of different items and domains of the introduced peer reviewing checklist by SPSS software. Moreover, Delphi method was applied to confirm the validity of the instrument by experts' opinion.

Results: A group of experts finally confirmed the validity of this checklist by means of Delphi method. Moreover, internal consistency of the main domains of the checklist consisting of "Introduction", "Methods", "Results" and "Conclusion" were statistically significant ($P < 0.05$). The coefficient of Cronbach's alpha was calculated as excellent (0.926) in "Methods", and acceptable in two other domains of "Results" (0.650) and "Conclusion" (0.739) in pre-test data.

Conclusion: The introduced checklist was shown to be a structured, valid and reliable tool, which could be used in research congresses to improve the quality of peer reviewing.

Keywords: Peer reviewing, Critical appraisal, Checklist, Reliability, Validity, Research Congress, Medical students