



ارزیابی کیفی پژوهش

دکتر ونداد شریفی

یک را در چهار سطح رده‌بندی می‌کند که بالاترین آن از نظر اهمیت و توان علمی واجد کیفیتی پیشرو در سطح جهان است. از یافته‌های این نهاد برای درجه‌بندی مراکز آموزش علمی و اختصاص بودجه به آنها استفاده می‌شود. ارزیابی تحقیقات روانپزشکی، نوروساینس و روانشناسی بالینی را کمیته‌ای به سرپرستی پروفسور پیتر مک‌گافین^۱ از کینگز کالج لندن بر عهده دارد.

برای سنجش کیفیت پژوهش، روش‌های مختلفی به کار می‌رود. یک روش معمول، استفاده از شاخص‌های رایج کتاب‌سنجی مانند شاخص اثرگذاری^۲ و میزان استناد^۳ است (به علم‌سنجی در شماره‌های قبلی مراجعه کنید). فرض این است که اگر مقاله‌ای در مجله‌ای با شاخص اثرگذاری بالا چاپ شود یا میزان استناد به آن مقاله زیاد باشد، می‌توان گفت که کیفیت آن پژوهش نیز بالاست. البته به استفاده از این شاخص‌ها برای سنجش کیفیت انتقادهایی شده است که برخی از آنها به محدودیت‌ها و سوگیری‌های خود شاخص‌ها معطوف است، ولی برخی دیگر به این مسئله مرتبط می‌باشد که ممکن است این شاخص برای سنجش کیفیت چندان مناسب نباشد. برای مثال، ممکن است مقاله محقق «الف»^{۲۵} استناد داشته باشد و مقاله محقق «ب» نیز^{۲۵} استناد، اما^{۲۵} استناد مربوط به محقق «الف»، مقاله را تحسین و از یافته‌های آن در پژوهش خود استفاده کرده باشند، ولی در مورد محقق «ب»،^{۲۰} استناد به انتقاد اساسی از کیفیت پژوهش پرداخته و فقط ۵ مورد به مقاله نگرش مثبت داشته باشند. در این حالت، هر دو محقق میزان استناد و بالطبع ارزش‌گذاری یکسانی خواهند داشت. از سوی

مطالعات نشان داده‌اند که رشد علمی کشور در سال‌های اخیر از نظر تعداد مقالات منتشر شده در رشته‌های مختلف علمی، حتی در سطح بین‌المللی، چشمگیر بوده است. اما همواره این سؤال مطرح شده که آیا رشد کمی پژوهش‌ها با رشد کیفی آنها همراه بوده و اساساً زیاد بودن مقالات یک محقق، دانشگاه و حتی یک کشور، لزوماً با کیفیت بیشتر آنها مترادف است. روشن است که هرچه کیفیت علمی پژوهش‌ها بیشتر باشد، آن پژوهش‌ها در پیشبرد و نیز کار بست علم مؤثرترند. از سوی دیگر، کیفیت پایین پژوهش می‌تواند آثار منفی داشته باشد و به نتیجه‌گیری‌های نادرست منجر گردد، به طوری که یک پژوهش نشان داده است که هرچه کیفیت مقالات بررسی اثرات درمانی پایین‌تر باشد، میزان اثرمندی گزارش شده بیشتر است. زمانی که از کیفیت پژوهش صحبت می‌شود، ضرورتاً مسئله ارزش مطرح و بنابراین سنجش کیفیت، نوعی ارزش‌گذاری به پژوهش‌ها نیز تلقی می‌گردد (البته منظور از پژوهش، مقاله یا گزارش پژوهشی است و در اینجا از کنترل کیفیت حین پژوهش بحث نمی‌شود).

ارزیابی کیفی با مقاصد و روش‌های مختلف صورت می‌گیرد؛ مانند داوری و پذیرش مقالات، بررسی طرح‌های تحقیقاتی، بررسی عملکرد محققان، مراکز علمی و دانشگاه‌ها، نحوه تخصیص بودجه و تسهیلات و ... گاه به منظور سنجش کیفیت مقالات پژوهشی حوزه‌های علمی، پژوهش‌هایی با روش‌شناسی مشخص انجام می‌شود و نهادهایی هم وجود دارند که مسئولیت اصلی آنها ارزیابی کیفیت پژوهش‌هاست. برای مثال، Research Assessment Exercise در انگلستان یک برنامه ملی peer review برای افزایش کیفیت پژوهش‌هاست و پژوهش‌های انجام شده در سطوح مختلف دانشگاهی و آموزش عالی را بررسی و کیفیت هر

1- Peter McGuffin
3- citation

2- impact factor

Cochrane در مورد اثربندی روش‌های درمانی است، با روش ساختار یافته، نوعی ارزیابی کیفی از مجموعه مقالات مختلف به عمل می‌آورد تا بتواند از آن مجموعه جمع‌بندی صحیحی داشته باشند. البته گاه دیده شده که حتی در مرورهای Cochrane، این ارزیابی کیفیت که در واقع نقدی به پژوهش‌های انجام شده است، در تفسیر نتایج و جمع‌بندی کلی استفاده نمی‌شود. اگرچه به این نکته هم باید توجه کرد که ممکن است تحقیقی با کیفیت مطلوب و رعایت ضوابط علمی انجام شده باشد، ولی به دلیل گزارش‌دهی نامطلوب، کیفیت گزارش یا مقاله پژوهشی پایین ارزیابی شود. از این رو توصیه می‌شود محققان برای گزارش مقاله علمی خود از الگوهای استاندارد، مانند راهنمای CONSORT (که برای گزارش کارآزمایی‌های بالینی پیشنهاد شده است) استفاده کنند. بسیاری از پژوهش‌ها از روش‌های فوق برای بررسی کیفیت مقالات پژوهشی به ویژه در حوزه پزشکی بهره برده‌اند. در اینجا به چند پژوهش منتشر شده در زمینه کیفیت مقالات پژوهشی روانپزشکی اشاره می‌شود: سلدانی^۳ و همکاران، کیفیت روش‌شناسی ۱۰۰ مقاله حوزه درمان اختلال دو قطبی را بررسی و مشاهده کردند که در ۴۳ درصد از مقالات که میزان ریزش بیماران^۴ در حد زیاد (بیش از ۴۰ درصد) بوده است، فقط دو مقاله فاصله اطمینان^۵ را گزارش کرده و اکثراً به ذکر *p value* اکتفا نموده‌اند و یک مقاله هم توان آماری مطالعه را آنالیز کرده است. تورنلی^۶ و همکاران نیز با بررسی ۲۰۰ کارآزمایی بالینی در اسکیزوفرنیا مشاهده کردند که اکثر مطالعات کوتاه مدت^۱ (کمتر از شش هفته) و تعداد نمونه کم (میانگین ۶۵ نفر) بوده است. فقط چهار درصد در مورد نحوه اختصاص نمونه به گروه‌های درمانی^۷ و دو درصد در مورد نحوه *blinding* به اندازه کافی توضیح داده‌اند. ۳۵ درصد از هیچ ابزار درجه‌بندی شده برای سنجش اثربخشی درمان استفاده نکرده و باقی مقالات بیش از ۶۴۰ ابزار مختلف را به کار برده‌اند.

به نظر می‌رسد که مطالعات مربوط به سنجش کیفیت پژوهش‌ها در ایران محدود باشد. چند مقاله کیفیت پایان‌نامه‌های

دیگر، برخی، شاخص اثرگذاری را بیشتر شاخص نوعی نظرسنجی و رأی دادن سایر محققان می‌دانند نه کیفیت؛ مثلاً، ممکن است مقاله‌ای به این دلیل بیشتر مورد استفاده و استناد قرار گیرد که در مجله تخصصی معتبرتر و شناخته شده‌تری نزد اهل علم چاپ شده است که این لزوماً به معنی کیفیت بالاتر آن نیست. با این همه، شاخص‌هایی مانند اثرگذاری و میزان استناد یکی از رایج‌ترین شاخص‌های سنجش کیفیت است. در چند مطالعه هم مشاهده شده است که این دو شاخص با ارزشیابی جداگانه کیفیت به وسیله خبرگان، ارتباطی قوی داشته‌اند.

روش‌های دیگری هم وجود دارند که مانند شاخص اثرگذاری، بر این فرض استوارند که مقاله با کیفیت بالا در مجله با کیفیت بالا چاپ می‌شود؛ مانند مجلاتی که مشترکین زیادی دارند، در پایگاه‌های اطلاعاتی مهمی نمایه می‌شوند، بررسی و داوری موشکافانه‌تری دارند و بالطبع میزان پذیرش مقاله^۱ آنها نیز پایین است.

روش رایج دیگر برای ارزیابی کیفیت مقالات و گزارش‌های پژوهشی، استفاده از معیارهای از پیش تعیین شده است. این ارزیابی را عمدتاً گروهی از خبرگان و صاحب‌نظران علمی انجام می‌دهند. آنها بررسی می‌کنند که معیارها و ضوابط (مانند مفهوم‌سازی، اهداف، توصیف و نحوه انتخاب نمونه، روش‌های آماری و مانند آنها) تا چه حد رعایت و از خطا^۲ اجتناب شده است. برای پایایی بیشتر اندازه‌گیری، ابزارهایی مانند *Jadad scale* و برای ارزیابی کیفیت کارآزمایی‌های بالینی، ابزاری به نام *READER* طراحی شده است. رویکردهای استفاده شده در این روش، هم کمی (مانند روش‌های نمره‌گذاری) و هم کیفی است. یک مسئله مهم در این ارزیابی‌ها، پایایی و روایی اندازه‌گیری است. در برخی مطالعات دیده شده که در یک مقاله، حتی میزان توافق خبرگان یک رشته در یک شاخص خاص هم بالا نبوده است. از همین رو پیشنهاد شده است که صاحب‌نظران کیفیت را با یک روش‌شناسی مشخص و پس از طی یک دوره آموزش ارزیابی کنند.

Peer review در واقع نوعی ارزیابی کیفیت است که با روش فوق انجام می‌گیرد و به ویژه در مجلات علمی به کار برده می‌شود. در مقالات مروری نظام‌مند نیز که معتبرترین آنها مرورهای

1- acceptance rate
3- Soldani
5- confidence interval
7- treatment allocation

2 - error
4 - dropout
6- Thornley

در انتها بر این نکته تأکید می‌شود که سنجش علم فقط به سنجش تعداد مقالات محدود نمی‌شود و شاید کشور ما به خصوص نیازمند ارزشیابی کیفیت پژوهش‌هایش نیز باشد و با هر روش ممکن باید روشن سازد که کیفیت تحقیقات تا چه حد بالا و تا چه حد در پیشبرد علم و کاربرد آن مؤثر بوده است.

پزشکی عمومی را بررسی کرده‌اند و یک مقاله نیز روش‌های آماری ۹۰ مقاله منتشر شده در مجلات پزشکی سال ۱۳۷۷ را بررسی و مقاله اخیر نیز نواقصی را در مقالات مشاهده کرده است. برای مثال، در ۴۹ درصد موارد به نوع تحقیق اشاره نشده، در ۳۸ درصد توصیف روش‌های آماری یا وجود نداشته یا بسیار ناقص بوده، در نیمی از مقالات حداقل یک خطای واضح آماری وجود داشته و فقط در ۱۱ درصد صحت کلیه آزمون‌های آماری تأیید شده است.

منابعی برای مطالعه بیشتر

- Clyde, L. A. (2004). Evaluating the quality of research publications: A pilot study of school librarianship. *Journal of the American Society for Information Science and Technology*, 55(13), 1119-1130.
- Morrison, J., & Bligh, G. (2003). Assessing the quality of research. *Medical Education*, 39, 542-544.
- Thornley, B., & Andams, C. (1998). Content and quality of 2000 controlled trials in schizophrenia over 50 years. *BMJ*, 317, 1181-1184.
- Lee, K. P., Schotland, M., Bachhetti, P., & Bero L. (2002). Association of journal quality indicators with methodological quality of clinical research articles. *JAMA*, 287, 2805-2808.
- Goldani, F., Ghaemi, S.N., & Balessarihi, R. J. (2003). Research reports on treatments for bipolar disorder: Preliminary assessment of methodological quality. *Acta Psychiatrica Scandinavica*, 112, 72-74.