

خطاهای رایج در تفسیر داده‌های پژوهشی

زیبا فرج‌زادگان^۱

سر مقاله

مقدمه

بیشتر پژوهشگران تمایل دارند در پژوهش‌هایی که به صورت کارآزمایی یا مشاهده‌ای انجام می‌دهند، روابط بین متغیرهای موجود را به روش‌های گوناگون بیان کنند. آنان درصدد هستند که در گزارش یافته‌ها، با استناد به مقدار P و فاصله اطمینان ۹۵ درصد، خواننده را به پذیرش رابطه‌های یافت شده در پژوهش متقاعد سازند. از سوی دیگر، مجلات نیز تمایل دارند مقالاتی با یافته‌های مثبت را منتشر نمایند و به همین دلیل، گرایش نویسندگان برای گزارش یافته‌های معنی‌دار هم بیشتر می‌شود؛ در حالی که یافتن ارتباط آماری معنی‌دار همیشه به معنای یافتن نوعی رابطه حقیقی نیست. در واقع، ممکن است عوامل گوناگونی باعث شکل‌گیری نوعی رابطه غیر واقعی بین متغیرها شوند و یا بر عکس وجود رابطه‌ای حقیقی را پنهان سازند. بنابراین، پژوهشگران باید بدانند چه نکاتی را در طراحی و تجزیه و تحلیل داده‌ها در نظر گیرند تا یافته‌های خود را از نتیجه‌گیری غیر واقعی مصون نمایند.

یکی از نکات مهم در این زمینه، حجم نمونه است؛ چرا که هر قدر نمونه‌های بیشتری در مطالعه وجود داشته باشد، احتمال یافتن نتایج مثبت و روابط معنی‌دار (هرچند ضعیف) افزایش می‌یابد و در واقع، هیچ عامل توجیه‌کننده‌ای برای این رابطه نمی‌توان یافت. نقطه مقابل این وضعیت هم آن است که نتوان رابطه‌های کاملاً حقیقی را به دلیل حجم نمونه کم کشف نمود و این موضوع به قدرت آماری مطالعه بستگی دارد؛ یعنی اگر قدرت آماری قابل قبول باشد، با در نظر گرفتن سایر عوامل، می‌توان فقدان رابطه آماری را به نداشتن تفاوت یا فقدان رابطه بین متغیرها نسبت داد.

برای پیشگیری از این خطا، باید حجم نمونه مناسب باشد؛ چرا که حجم نمونه، مقدار P و تفسیر یافته‌ها را نیز تحت تأثیر قرار می‌دهد؛ به طوری که پژوهشگران در تفسیر و تحلیل داده‌ها بر مبنای مقدار P ، مقادیر کمتر از ۰/۰۵ را تنها مبنای تفسیر و نتیجه‌گیری قرار می‌دهند؛ در حالی که این استناد به تنهایی برای قضاوت و نتیجه‌گیری درباره روابط مثبت بین متغیرها کافی نیست.

فرایند تفسیر و نتیجه‌گیری بر مبنای P یا فاصله اطمینان، باید با این پرسش همراه شود که آیا این تفاوت «معنی‌دار»، «واقعی یا غیر واقعی» است. در پاسخ به این پرسش، در صورتی می‌توان ادعا کرد که مداخله «تأثیر» گذاشته یا «تفاوت و رابطه» وجود داشته است که این تأثیر و تفاوت با خطاهای احتمالی مرتبط نباشد. در این زمینه چند احتمال برای تفسیر نادرست یافته‌ها مطرح است که در ادامه بیان می‌شود.

اولین احتمال، وجود سوگرایی در پژوهش است. این مشکل بیشتر در طراحی مطالعه رخ می‌دهد. سوگرایی انواع گوناگونی دارد و باید تا حد ممکن از بروز آن جلوگیری نمود و در صورت وقوع، در مرحله طراحی و اجرای مطالعه آن

را برطرف کرد. البته برخی از سوگرایی‌ها در مرحله تجزیه و تحلیل داده‌ها شناسایی می‌شود و با روش‌های پیشرفته آماری می‌توان آن‌ها را کنترل کرد. از جمله سوگرایی‌های عمده، در مرحله انتخاب نمونه (Selection bias) می‌باشد؛ بدین معنی که انتخاب نمونه تحقیق، در واقع نماینده کل افراد جامعه نباشد و این بر تعمیم‌پذیری نتایج تأثیرگذار است. نوع دیگر، سوگرایی اطلاعات (Information bias) است و زمانی رخ می‌دهد که امکان جمع‌آوری اطلاعات کافی از نمونه تحقیق ممکن نباشد و یا اطلاعات نمونه‌ها با ابزارها و روش‌های مختلفی جمع‌آوری گردد. در این حالت، اطلاعات به طور ناقص یا نادرست جمع‌آوری می‌شود و مبنای تفسیر قرار می‌گیرد. نوع دیگر، مربوط به عوامل مخدوشگر (Confounder bias) است. در این وضعیت، یافته‌های مطالعه تحت تأثیر متغیرهای متعددی قرار می‌گیرد که این متغیرهای اضافی، رابطه بین دو متغیر اصلی را تشدید یا تعدیل کرده‌اند. بیشتر پژوهشگران در تلاش هستند تا با تخصیص تصادفی و همسان‌سازی، عوامل مخدوشگر را کنترل نمایند و اثر آن‌ها را خنثی سازند. این راهکار را می‌توان برای کنترل کردن و خنثی‌سازی عوامل مخدوشگر شناخته شده انجام داد، اما باید توجه کرد که عوامل مخدوشگر ناشناخته همچنان وجود دارند و به نتیجه‌گیری خدشه وارد می‌کنند.

احتمال بعدی که باید آن را در تفسیر نتایج بررسی کرد، تصادفی بودن یافته به دست آمده است. اگر چنین باشد، با نمونه‌های گوناگون و حجم نمونه مختلف، نتایج متفاوتی به دست می‌آید. این خطا، «خطای تصادفی» محسوب می‌شود و با حجم نمونه کافی، اثر خطای تصادفی کمتر می‌شود. حتی در صورتی که محقق کنترل لازم را برای جلوگیری از سوگیری و خطای تصادفی انجام دهد، باز هم این احتمال وجود خواهد داشت که یافته به دست آمده فقط تفاوت معنی‌دار آماری را نشان دهد و مقدار تفاوت به لحاظ بالینی ارزشمند نباشد. بنابراین، لازم است پژوهشگر به تفاوت بین معنی‌داری آماری و معنی‌داری بالینی توجه نماید و در بحث و نتیجه‌گیری به آن بپردازد.

نتیجه‌گیری

در مطالعه حاضر این موضوع بررسی شد که نباید یافته‌های مثبت را فقط بر مبنای یک شاخص آماری تفسیر و نتیجه‌گیری کرد؛ چرا که درباره یافته‌های مثبت همیشه چند وضعیت وجود دارد. اول این که یافته ناشی از خطای تصادفی

مقاله حاصل تحقیق مستقل بدون حمایت مالی و سازمانی است.

۱- استاد، پزشکی اجتماعی، گروه پزشکی اجتماعی و خانواده، دانشکده پزشکی، دانشگاه علوم پزشکی اصفهان، اصفهان، ایران (نویسنده مسؤول)

Email: farajzadegan@med.mui.ac.ir

در نهایت، با این که همه خطاها را نمی‌توان در مطالعه کنترل یا حذف کرد، پژوهشگران باید با رعایت اصول و قواعد مربوط به تفسیر علمی یافته‌ها و بیان محدودیت‌های تعمیم‌پذیری یافته‌ها در مطالعه، نتایج معتبری به خوانندگان گزارش دهند.

است؛ دوم، یافته ناشی از خطای سیستماتیک یا سوگرایی است و سوم، یافته به طور قطع رابطه‌ای حقیقی را نشان می‌دهد. اگر یافته در اثر خطای تصادفی و سیستماتیک نباشد و از نظر بالینی مهم باشد، تفسیر آن بر مبنای رابطه یا اثری حقیقی انجام می‌شود.

ارجاع: فرج‌زادگان زیبا. **خطاهای رایج در تفسیر داده‌های پژوهشی.** مدیریت اطلاعات سلامت ۱۳۹۵؛ ۱۳ (۵): ۳۱۶-۳۱۷