

تورش‌ها در مطالعات کارآزمایی کنترل‌دار تصادفی منتشرشده در نشریه‌های تخصصی پرستاری و مامایی ایران در سال ۱۳۸۹

مینا محمدی^۱، نرگس طوقیان چهارسوقی^۲، سامره عبدلی^۳

^۱ کارشناس ارشد پرستاری، عضو هیئت علمی دانشگاه آزاد اسلامی واحد خوراسگان (اصفهان)، دانشکده پرستاری و مامایی، ایران

^۲ کارشناس ارشد پرستاری کودکان، اصفهان، ایران

^۳ دکترای تخصصی پرستاری، استادیار مرکز تحقیقات مراقبت پرستاری و مامایی، دانشکده پرستاری و مامایی، دانشگاه علوم پزشکی اصفهان، ایران

نویسنده رابط: نرگس طوقیان، نشانی: اصفهان، خیابان هزار جریب، دانشگاه علوم پزشکی اصفهان، دانشکده پرستاری و مامایی. تلفن: ۶۶۸۰۰۴۲، پست الکترونیک:

toghiann@gmail.com

تاریخ دریافت: ۱۳۹۱/۵/۲۹؛ پذیرش: ۱۳۹۱/۱۱/۷

مقدمه و اهداف: کارآزمایی‌های کنترل‌دار تصادفی از جمله انواع مطالعاتی است که می‌توانند معتبرترین اصول را برای مقایسه مداخله‌های گوناگون در پژوهش‌های علمی فراهم سازند. تورش می‌تواند به‌شدت بر کیفیت پژوهش را تأثیر بگذارد و نتایج مطالعه را نامعتبر سازد. هدف این مطالعه بررسی تورش‌ها در مطالعات کارآزمایی کنترل‌دار تصادفی منتشرشده در نشریه‌های پرستاری و مامایی است.

روش کار: در این توصیفی، همه مقالات منتشرشده در نشریه‌های تخصصی پرستاری و مامایی در سال ۱۳۸۹ از نظر معیارهای ورود بررسی شدند. تورش‌ها در مقالات منتخب با کمک ابزار استاندارد Risk of bias گروه‌بندی، توسط دو نفر به‌صورت مستقل بررسی شد. همه داده‌های به‌دست‌آمده با کمک آمار توصیفی گزارش شدند.

نتایج: پس از بررسی مقالات در ۸ مجله، ۶۸ کارآزمایی کنترل‌دار تصادفی بررسی قرار شدند. بیشتر مطالعات دارای تورش بالا و یا مبهم بودند. تولید توالی تصادفی در ۲۲٪ مطالعات، تخصیص تصادفی در ۴/۵٪، کورسازی در ۲۲٪، گزارش داده‌های ناقص در ۳۵/۳٪، گزارش انتخابی پیامدها در ۵۱/۵٪ و دیگر تورش‌ها در ۳۶/۷٪ مطالعات تورش پایینی داشتند.

نتیجه‌گیری: از کارآزمایی‌های کنترل‌دار تصادفی در تحقیقات پرستاری و مامایی بسیار استفاده می‌شود. یافته‌های به‌دست‌آمده نشان داد که بخش زیادی از این کارآزمایی‌ها دارای تورش بالا و یا مبهم است. باید یادآور شد که گزارش‌نشدن این موارد به ابهام و سردرگمی خوانندگان مطالعه منجر می‌شود و تورش در این ابعاد به میزانی قابل توجه اعتبار داخلی نتایج مطالعه را مخدوش می‌کند. بنابراین، یافته‌های این گونه مطالعات باید با احتیاط تفسیر شوند.

واژگان کلیدی: کارآزمایی کنترل‌دار تصادفی، تورش، آنالیز تمایل به درمان، آنالیز مبتنی بر پروتکل

مقدمه

روشناسی قلب هر پژوهش است و توجه به طراحی آن تضمین‌کننده درستی داده‌های جمع‌آوری شده و الگوی تحلیل آن‌ها است (۱). بر اساس بیانیه هلسینکی، «نویسندگان وظیفه دارند نتایج مطالعات انسانی خود را در دسترس عموم مردم قرار دهند و مسئولیت درستی نتایج نیز بر عهده آنان است» (۲). کارآزمایی‌های کنترل‌دار تصادفی (RCT) از جمله انواع مطالعاتی‌اند که می‌توانند معتبرترین اصول را برای مقایسه مداخله‌های گوناگون در پژوهش‌های بهداشتی و درمانی فراهم

سازند (۳). RCT‌ها این قابلیت را دارند که بتوانند به میزانی قابل توجه روش‌های مراقبتی از بیماران را اصلاح کنند و در ارتقای سلامت بیماران تأثیرگذار باشند. در مقابل، وجود تورش در طراحی و اجرای یک مطالعه می‌تواند بر کیفیت آن به‌شدت اثر بگذارد و در نهایت، نتایج یک مطالعه را نامعتبر سازد. استفاده از نتایج نامعتبر یک مطالعه در درمان نیز می‌تواند باعث شود بیماران در معرض خطرات آن قرار گیرند، بی آن که درمان به‌کاررفته برای آن‌ها مفید باشد (۴). تورش در یک تعریف کلی به معنای نبود بی‌طرفی در مطالعه و سوگیری در آن و در واقع، انحراف از حقیقت است. تورش‌ها می‌توانند در همه مراحل یک پژوهش، اعم

^۱ Randomised controlled trial

پذیرش شرکت‌کنندگان و تداوم بقای آنان در مطالعه را با کاستن از تورش ناشی از تأثیرات آگاهی بهبود می‌بخشد (۱۰). نتایج مطالعه پیدال (۲۰۰۷) نیز نشان داده‌است که استفاده نکردن از کورسازی در مطالعات کارآزمایی بالینی به افزایش اثر تخمین‌زده به میزان ۹٪ منجر می‌شود (۱۲).

باید توجه داشت که برای کاهش تورش‌ها در مطالعات RCT، افزون بر استفاده از روش‌های پیش‌گفته لازم است که با کمک راهکارهایی مؤثر در کاهش تورش ریزش نمونه‌ها^۴ گام برداشت. داده‌های گمشده که در اثر ریزش نمونه‌ها از مطالعه و یا انصراف و خروج از مطالعه به وجود می‌آیند، باعث ایجاد تورش در میزان اثر تخمین‌زده می‌شوند (۱۲). در واقع، حذف نتایج حاصل از نمونه‌هایی با داده‌های گمشده در مرحله تجزیه و تحلیل، به نسبت زمانی که همه داده‌ها تجزیه و تحلیل شدند، موجب می‌شود که نتایج به‌طور کاذب به نفع گروه مداخله رقم بخورند (۱۳). از این رو همه RCTها باید به علت ریزش حذف و خروج نمونه‌ها از مطالعه اشاره کنند و یا در صورتی که از آنالیز تمایل به‌درمان^۵ (ITT) در مطالعه استفاده کرده‌اند، آن را گزارش نمایند.

از آنجا که معمولاً مطالعات RCT بیش از دیگر انواع مطالعات می‌توانند در امر مراقبت و درمان بیماران تأثیرگذار باشند، گزارش‌های این مطالعات نیز باید این توانایی را در خواننده ایجاد کنند که با توجه به اطلاعات گزارش‌شده، خ قادر به قضاوت در مورد اعتبار داخلی و خارجی کارآزمایی باشند. از این رو، در انتشار مطالعات RCT باید به بالاترین سطوح استاندارد توجه کرد (۱۴). از سال ۱۳۸۳ تاکنون حدود ۷۵ مجله علمی و پژوهشی فعالیت خود را شروع کرده‌اند که در این میان ۱۰٪ آن‌ها نشریه‌های تخصصی پرستاری و مامایی‌اند. با توجه به رشد روزافزون مجلات تخصصی پرستاری و مامایی در ایران و تأثیر آن‌ها در ارتقاء کیفیت مراقبت‌های پرستاری و مامایی، گزارش‌دهی درست مقالات منتشر شده در آن‌ها بسیار مهم است. اما تاکنون مطالعه‌ای که به بررسی تورش‌ها در مقالات چاپ شده در مجلات پرستاری و مامایی ایران نپرداخته است. از این رو، بر آن شدیم تا به مطالعه‌ای با این هدف بپردازیم.

از طراحی تا انتشار مطالعه رخ دهند. انواع گوناگونی از تورش‌ها در مطالعات وجود دارند که یکی از مهم‌ترین آن‌ها تورش انتخاب است. بهترین روش برای کاهش تورش انتخاب استفاده از تصادفی‌سازی است (۵). تصادفی‌سازی در مطالعات کارآزمایی بالینی فرایند تقسیم شرکت‌کنندگان به گروه‌های مختلف است. تصادفی‌سازی به هر شرکت‌کننده شانس برابر برای قرار گیری در هر گروه را می‌دهد و شامل دو مرحله تولید توالی تصادفی و تخصیص آن بر نمونه‌های واجد شرایط است (۶،۷). به‌رغم اهمیت قابل توجه تصادفی‌سازی در مطالعات RCT، نتایج مطالعه شولز (۲۰۰۲) نشان داد که تنها ۳۲٪ مقالات منتشرشده در ۴ مجله ژنیکولوژی در سال‌های ۱۹۹۰-۱۹۹۱ از روشی درست برای تولید توالی تصادفی استفاده کرده بودند. همچنین، ۲۳٪ مطالعات اطلاعات کافی درباره مراحل تخصیص تصادفی را ارائه داده بودند (۳). چگونگی و مراحل انجام صحیح تصادفی‌سازی در دستیابی به نتایج معتبر و قابل استناد آن قدر اهمیت دارد که نتایج مطالعه‌ای توسط موهر (۱۹۹۸)، نشان داد مطالعاتی که از روش نادرست در تولید توالی تصادفی و تخصیص آن استفاده کرده بودند، در مقایسه با مطالعاتی که از روش درست استفاده کرده بودند، به میزان ۵۰-۳۰٪ اثر تخمین‌زده را بیش از مقدار واقعی گزارش می‌کنند (۸). در حقیقت، می‌توان گفت که مطالعات با کیفیت پایین تمایل به اغراق نتایج به‌دست‌آمده دارند (۹).

دومین تورش مهم در کارآزمایی‌های بالینی تورش اجرا^۱ است. استفاده از کورسازی می‌تواند در کاهش تورش اجرا کمک‌کننده باشد، به‌گونه‌ای که استفاده از کورسازی دستیابی به نتایج واقعی را ارتقا می‌بخشد و از تأثیرات سوء ناشی از آگاهی و نگرش پژوهشگران و شرکت‌کنندگان بر نتایج مطالعه می‌کاهد. همچنین با کورسازی ارزیاب، می‌توان تا حد امکان از تورش اندازه‌گیری^۲ در مطالعه پیشگیری کرد (۱۰). واژه‌های یک سو کور، دو سو کور و سه سو کور ممکن است معانی متفاوتی برای افراد مختلف داشته باشند. افزون بر آن، بسیاری از محققان علوم پزشکی کورسازی را با پنهان‌سازی تخصیص تصادفی اشتباه می‌گیرند. چنین اشتباهاتی نشان‌دهنده فقدان درک افراد از هر دو مفهوم است (۱۱). کورسازی معمولاً بررسی ناهمسان و متفاوت نتایج به‌دست‌آمده را کاهش می‌دهد (تورش اطلاعات^۳) و همچنین،

^۱ Performance bias

^۲ Measurement bias

^۳ Information bias

^۴ Attrition bias

^۵ Intention to treat analysis

روش کار

این مطالعه یک مطالعه توصیفی است که به بررسی کیفیت RCTهای منتشرشده در مجلات تخصصی پرستاری و مامایی کشور در سال ۱۳۸۹ پرداخته است. پس از تهیه فهرست مجلات تخصصی پرستاری و مامایی که دارای رتبه علمی و پژوهشی از کمیسیون نشریات کشور در سال ۱۳۸۹ بودند، همه مقالات منتشرشده آن مجلات در سال ۱۳۸۹ از طریق بانک اطلاعاتی Magiran بررسی شدند. در ابتدا، عنوان و چکیده مطالعات به دست آمده از جست و جو، توسط دو نفر، از نظر دارا بودن معیارهای ورود، بررسی شدند. مطالعات مورد توافق وارد مرحله بعدی شدند و مطالعات رد شده از سوی هر دو نفر، کنار گذاشته شدند. در مورد مطالعاتی که درباره ورود آن‌ها بین دو محقق اختلاف نظر وجود داشت، با بحث و گفت و گو توافق به دست آمد. همچنین، برای بررسی کیفیت مطالعات، دو کارشناس به طور مستقل، هر مطالعه را با کمک یک فهرست کنترل استاندارد، ارزیابی کردند. برای بررسی کیفیت مطالعات از ابزار استاندارد استفاده شد. معیارهای ورود به مطالعه شامل مطالعاتی بود که در قسمت روش پژوهش به نوع مطالعه کارآزمایی، کنترل دار، تصادفی اشاره کرده بودند، بدون توجه به این که آیا روش تصادفی سازی درست بود یا خیر. همچنین، مطالعاتی که در آن کورسازی انجام شده یا نشده بود، نیز وارد شدند. مقاله‌های منتخب در مجلات دارای رتبه علمی و پژوهشی از کمیسیون نشریات کشور در حیطه پرستاری و مامایی، در سال ۱۳۸۹، منتشر شده بودند. برای استخراج داده‌ها از یک فرم دارای ساختار بندی و استاندارد استفاده شد. در ابتدا، از هر مطالعه اطلاعات کلی شامل عنوان، سال انتشار، نویسندگان، مکان انجام مطالعه، مجله و دیگر اطلاعات لازم، جمع‌آوری شد. برای بررسی تورش‌ها در مطالعات از ابزار استاندارد Risk of bias مورد استفاده در مطالعات مرور نظام‌مند^۱ که توسط گروه کارکن^۲ تدوین شده است (آخرین به‌روزرسانی ۲۰۱۱)، استفاده شد. این ابزار شامل شش بعد است. ابعاد شامل روش ایجاد توالی تصادفی، نحوه تخصیص تصادفی نمونه‌ها و گزارش انتخابی پیامدها^۳ کورسازی و وجود هر گونه تورش احتمالی تأثیرگذار بر نتایج و گزارش داده‌های ناکامل است (۱۰). هر یک از موارد مورد بررسی در ابزار به سه صورت کم تورش، پُر تورش و مبهم از نظر تورش‌ها، گزارش شد. ابزار

استاندارد Risk of bias، یک ابزار معتبر و پایا در ارزیابی همه مطالعات کارآزمایی بالینی تصادفی، بدون توجه به زبان و زمان و مکان چاپ مقاله است (جدول شماره ۱). از آمار توصیفی برای تجزیه و تحلیل داده‌ها استفاده شد و هر یک گزینه‌ها به صورت توزیع فراوانی مطلق و نسبی گزارش شدند.

یافته‌ها

از میان مجلات پرستاری و مامایی، ۸ مجله دارای رتبه علمی و پژوهشی در سال ۱۳۸۹ بودند. پس از جست‌وجو در همه مجلات، ۲۸۶ مقاله به دست آمد که پس از بررسی توسط دو کارشناس، ۷۰ مقاله معیارهای ورود را داشتند. در میان مطالعات به دست آمده، دو مطالعه در دو مقاله منتشر شده بود. از این رو، ۶۸ کارآزمایی به لحاظ تورش بررسی شدند. در بررسی مجلات پرستاری و مامایی در سال، ۱۳۸۹ به طور میانگین ۲۴٪ مقالات منتشرشده از نوع RCT بودند که کمترین درصد مربوط به فصلنامه پرستاری ایران با ۱۴/۵٪ و بیشترین درصد مربوط به مجله حیات با ۵۳٪ بود (جدول شماره ۲). تعداد نویسندگان مقالات به طور میانگین ۴ نویسنده (حداقل ۱ و حداکثر ۷ نفر) بود. از ۶۸ مطالعه در ۱۹ (۲۸٪) مطالعه یک متخصص آمار و یا اپیدمیولوژیست از جمله نویسندگان مطالعه بودند. نوع مداخله به کاررفته در ۲۹ مطالعه رفتاری، ۱۵ مطالعه آموزشی، ۱۴ مطالعه دارویی و ۹ مطالعه بررسی تأثیر ماده‌ای خاص بود. نتایج ۵۲ (۷۶/۵٪) مطالعه از نظر آماری معنی‌دار گزارش شده بود. تعداد گروه‌ها در ۹۱٪ مطالعات دو گروهه، در ۶٪ سه گروهه و در ۳٪ چهار گروهه بود. نود و چهار درصد مطالعات حجم نمونه‌ای کمتر از ۱۰۰ نفر و ۵/۴٪ مطالعات حجم نمونه‌ای متوسط، یعنی بین ۲۰۰-۱۰۰ نفر، در هر گروه داشتند. در این میان یک مطالعه (۱/۵٪) حجم نمونه ۲۰۰ نفر در هر گروه را گزارش کرده بود. از میان ۶۸ مطالعه بررسی شده بیشتر مطالعات از بیان صریح نوع طرح پژوهشی به کاررفته خودداری کرده بودند، به گونه‌ای که نوع طرح به کاررفته در روش مطالعه با عباراتی مانند «یک طرح مداخله‌ای» و یا «یک کارآزمایی بالینی» گزارش شده بود. اما در بررسی انجام شده، سه جزء اصلی یک کارآزمایی، مانند مداخله، گروه کنترل و تصادفی سازی نمونه‌ها در مطالعه گزارش شده بود. بنابراین، این مطالعات RCT در نظر گرفته شدند و کیفیت آن‌ها بررسی شد. شش مطالعه دارای طرح خوشه‌ای بودند که ۳ مطالعه از بیان صریح خوشه‌ای بودن نمونه‌گیری خود خودداری کرده بودند و سه مطالعه نیز از طرح متقاطع (crossover) استفاده کرده بود.

^۱ Systematic review^۲ Cochrane group^۳ Selective outcome reporting

همچنین، درصد بسیار کمی از مطالعات نوع طرح پژوهشی خود را در عنوان مطالعه بیان کرده بودند. از میان ۶۸ مطالعه بررسی‌شده، ۴۸ مطالعه از نوع کارآزمایی بالینی تصادفی با گروه‌های موازی بودند که تصادفی‌سازی ساده روی فرد انجام شده بود که از این مطالعات، ۸۰٪ دارای حجم نمونه مساوی در گروه‌های تخصیص داده شده بودند. پنجاه و سه درصد مطالعات از سازمان حمایت‌کننده مالی طرح پژوهشی نام برده بودند و در همه موارد دانشگاه‌های دولتی بودند و ۴۷٪ مطالعات نام سازمان را گزارش نکرده بودند. همچنین، ۶۷ مطالعه (۹۸/۵٪) به ثبت طرح در مرکز کارآزمایی‌های بالینی ملی و یا بین‌المللی اشاره‌ای نکرده بودند. هیچ یک از مطالعات به انحراف از طرح پژوهشی تدوین‌شده^۱ در مطالعه خود اشاره‌ای نکرده بودند.

بررسی تورش‌ها

نمای یک به گزارش شش بعد تورش‌ها به صورت کم تورش، پُر تورش و مبهم پرداخته‌است (نمودار شماره ۱). نتایج مطالعه نشان داد که ۲۲٪ مطالعات از روشی کم تورش در تولید توالی تصادفی استفاده کرده بودند. بیشترین روش مورد استفاده، جدول اعداد تصادفی در ۵ مطالعه و کمترین روش مورد استفاده، کمینه‌سازی در یک مطالعه بود (جدول شماره ۳). از نظر کورسازی، ۲۲٪ مطالعات تورش پایین و ۵/۴٪ تورش بالا داشتند. بیشترین کورسازی اعمال‌شده، به‌ترتیب بر شرکت‌کنندگان در ۸ مطالعه، ارزیاب در ۷ مطالعه، پرسنل و محققان در ۵ مطالعه و متخصص آمار در یک مطالعه گزارش شده بود (جدول شماره ۴). همچنین، درصد افرادی که در مطالعات بین تصادفی‌سازی تا پیگیری نهایی ریزش یافته بودند، دامنه‌ای بین ۰٪ تا ۴۰٪ داشت. ۵ به نظر می‌رسید که ۲۶٪ مطالعات هیچ نوع ریزشی نداشتند که از آن میان، ۱۸ مطالعه (۲۲٪) پیامدها را بلافاصله پس از مداخله اندازه‌گیری کرده بودند، بنابراین فرصت کمی برای ریزش نمونه‌ها وجود داشت.

از ۲۲ مطالعه‌ای که ریزش نمونه‌ها را گزارش کرده بودند، در ۵ مطالعه میزان ریزش بیش از ۲۰٪ بود. در این میان، ۱۶ (۲۳/۵٪) مطالعه تعداد و علت ریزش نمونه‌ها، به تفکیک هر گروه گزارش نکرده بودند. از این رو، این مطالعات در ارزیابی انجام شده دارای تورش بالا برآورد شدند. تنها ۶ مطالعه از مطالعات دارای ریزش دارای علت و تعداد ریزش مشابه در گروه‌ها بودند. بنابراین، در

بررسی گزارش انتخابی پیامدها، از ۶۸ مطالعه ۳۵ مطالعه (۵۱/۵٪) همه پیامدها را ظاهراً گزارش کرده بودند. اما ۱۷ مطالعه (۲۵٪) برخی از پیامدهایی را که در روش پژوهش به آن اشاره کرده بودند، یا کلاً گزارش نکردند و یا به‌صورت ناقص گزارش کردند. همچنین، ۱۶ مطالعه نیز در بررسی گزارش انتخابی پیامدها در وضعیت مبهم قرار داشتند (جدول شماره ۵). جدول شماره ۵ به گزارش دیگر تورش‌ها، به‌تفکیک علت، پرداخته‌است. به‌طور کلی ۳۶/۷٪ مطالعات تورش پایین و ۳۵/۳٪ تورش بالا داشتند و ۲۸٪ مطالعات از نظر دیگر تورش‌ها مبهم بودند. شایع‌ترین علل تورش بالا، ناهمگنی در نتایج پایه^۲ بود، به‌گونه‌ای که ۱۳٪ مطالعات دارای این تورش بودند. ابزار مورد استفاده نیز در ۲۳/۵٪ موارد دارای تورش مبهم و بالا بود که این مسئله به‌ویژه در مورد ابزارهای پژوهشگرساخته که محققان هیچ‌گونه اطلاعاتی در مورد اعتبار و پایایی ابزار گزارش نکرده بودند، بیشتر بود (جدول شماره ۶).

^۲ Baseline imbalance

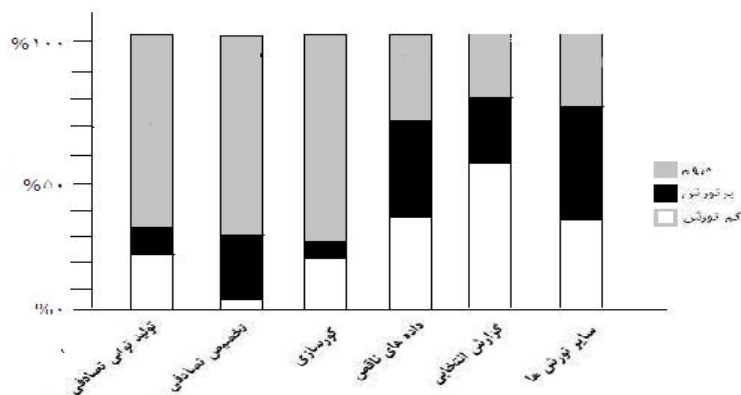
^۱ Deviation from the protocol

جدول شماره ۱- ابزار بررسی تورش‌ها در مطالعات کارآزمایی بالینی

آیتم	کم تورش	پُر تورش	مبهم
آیا روش تصادفی کردن درست بوده‌است؟ (توالی تصادفی غیر قابل پیش‌بینی)	تولید توالی بر اساس جدول اعداد تصادفی، نرم‌افزار تولید توالی تصادفی، پرتاب سکه، کشیدن کارت، قرعه‌کشی، کمینه‌سازی و دیگر موارد	تولید توالی بر اساس روزهای زوج و فرد تولد، روز پذیرش، هفته‌های زوج و فرد، شماره تخت و یا شماره پرونده بیمار، قضاوت پژوهشگر، تمایل شرکت‌کننده و دیگر موارد	نبود اطلاعات کافی برای قضاوت مثبت یا منفی
آیا تخصیص به مقدار کافی پنهان‌سازی شده است؟	شرکت‌کنندگان و پژوهشگر قادر به پیش‌بینی نحوه تخصیص افراد به گروه‌ها نبوده‌است، با روش‌هایی مانند استفاده از جعبه‌های مشابه و دارای اعداد تصادفی، نامه‌های مهر و موم شده دارای اعداد متوالی و دیگر موارد	شرکت‌کنندگان و یا پژوهشگر احتمالاً توانسته‌اند نحوه قرارگیری افراد را پیش‌بینی کنند، مانند استفاده از نامه‌های باز و شفاف، تخصیص یک در میان افراد به گروه‌ها و یا بر اساس تاریخ تولد و غیره	نبود وجود اطلاعات کافی برای قضاوت مثبت یا منفی
آیا کورسازی به خوبی اجراء شده است؟ (شرکت‌کنندگان، پرسنل درمانی، ارزیاب)	از کورسازی برای افراد کلیدی در مطالعه استفاده شده و یا برای افراد کلیدی در مطالعه استفاده نشده‌است اما ارزیاب در مطالعه کورسازی‌شده و افرادی که کورسازی نشده‌اند، تورشی را ایجاد نکرده‌است.	نتایج و مقادیر به‌دست‌آمده احتمالاً تحت تاثیر عدم کورسازی قرار گرفته‌اند و یا برای افراد کلیدی مطالعه کورسازی انجام شده‌است، اما به احتمال بسیار کورسازی فاش شده و با شکست روبه‌رو بوده است.	کمبود اطلاعات کافی برای قضاوت مثبت یا منفی
آیا نتایج داده‌های ناقص بیان شده‌اند؟	داده‌های گمشده وجود ندارد و یا داده‌های گمشده در مطالعه وجود دارد، اما در ارتباط با متغیرهای اصلی مطالعه نیستند و یا تعداد داده‌های گمشده بین گروه‌های مداخله و کنترل به یک میزان و با علل مشابه است.	مواردی مانند دلایل گزارش داده‌ها به‌صورت ناقص احتمالاً با درستی نتایج مطالعه مرتبط است. در گروه‌های مداخله و کنترل تعداد و دلیل ایجاد داده‌های گمشده یکسان نیست. میزان داده‌های گمشده به‌حدی زیاد است که می‌توانسته روی نتایج تاثیرگذار باشد. ITT صورت نگرفته است.	کمبود اطلاعات کافی برای قضاوت مثبت یا منفی
آیا مطالعه همه پیامدهای مورد نظر را بررسی و گزارش کرده‌است؟ (گزارش انتخابی پیامدها)	فرم پیشنهاد مطالعه در دسترس است و همه پیامدهای از پیش تعیین‌شده گزارش شده‌اند و یا فرم پیشنهاد مطالعه در دسترس نیست اما گزارش‌های منتشرشده آشکارا دربردارنده همه پیامدهای مورد انتظار است.	همه پیامدهای اولیه گزارش نشده‌اند و یا تعدادی از پیامدهای گزارش‌شده از ابزار و یا مقیاس‌هایی استفاده کرده‌اند که از پیش در پروپوزال بیان نشده‌اند. تعدادی از پیامدها به‌صورت ناقص گزارش شده‌اند و یا مقاله پیامدهای کلیدی مورد انتظار را گزارش نکرده است.	کمبود اطلاعات برای قضاوت مثبت یا منفی
آیا مطالعه از دیگر مسائلی که می‌تواند باعث سوگیری شود، مبرا است؟ (دیگر تورش‌ها)	تورشی در مطالعه وجود ندارد.	حداقل یک تورش قابل توجه در مطالعه وجود دارد، از جمله استفاده از طرح مطالعه و یا ابزار غلط، توقف زودتر از موعد مطالعه، ناهمگنی در اطلاعات پایه و وجود دیگر مشکلات در مطالعه	ممکن است تورش وجود داشته باشد، اما اطلاعات کافی در دسترس نیست.

جدول شماره ۲- فراوانی مطلق و نسبی مطالعات کارآزمایی به تفکیک مجلات پرستاری و مامایی در سال ۱۳۸۹

مجله	شماره در سال	مقالات منتشرشده در سال ۱۳۸۹	کارآزمایی بالینی (درصد)	درصد کارآزمایی به کل مقالات
دو ماهنامه زنان مامایی و نازایی ایران	۶	۴۸	۱۳	٪۲۷
Iranian journal of nursing and midwifery research	۴	۴۵	۱۴	٪۳۱
پژوهش پرستاری ایران	۴	۳۳	۵	٪۱۵
فصلنامه حیات	۴	۲۶	۱۴	٪۵۳
مجله دانشکده پرستاری و مامایی دانشگاه شهید بهشتی	۴	۳۴	۵	٪۱۵
فصلنامه پرستاری مراقبت ویژه	۴	۳۱	۵	٪۱۵
دو ماهنامه پرستاری ایران	۶	۴۸	۷	٪۱۴/۵
مجله دانشکده پرستاری و مامایی مشهد	۴	۲۱	۵	٪۲۴
جمع	۳۶	۲۸۶	۶۸	٪۲۴



نمودار شماره ۱- توزیع فراوانی نسبی هر یک از پارامترهای کیفی

جدول شماره ۳- توزیع فراوانی مطلق و نسبی انواع روش‌های تولید توالی تصادفی در مطالعات کارآزمایی منتشره در مجلات تخصصی پرستاری و مامایی در سال ۱۳۸۹

انواع روش‌های تولید توالی تصادفی	تعداد (n = ۶۸)	درصد
کشیدن کارت	۴	۶
قرعه‌کشی	۳	۴/۵
جدول اعداد تصادفی	۵	۷/۳
پرتاب سکه	۲	۳
کمینه‌سازی	۱	۱/۵
جمع	۱۵	٪۲۲
شماره تخت زوج و فرد	۱	۱/۵
شماره پذیرش زوج و فرد	۲	۳
هفته و یا روزهای زوج و فرد	۳	۴/۵
توالی یکی در میان	۱	۱/۵
جمع	۷	٪۱۰/۱۵
مبهم (فاقد توضیح مشکل یا عدم توضیح)	۴۶	٪۶۷/۱۶
جمع	۴۶	٪۶۷/۱۶

مبهم

جدول شماره ۴- توزیع فراوانی مطلق و نسبی انواع کورسازی در مطالعات، به تفکیک تورش اجرا و تورش اندازه‌گیری

کورسازی	کم تورش	پُر تورش	مبهم
کورسازی شرکت کنندگان، کارکنان و محققان	۱۳	۲	۵۳
کورسازی ارزیاب	۷	۱	۶۰
کورسازی* (درصد)	۱۵ (۰.۲۲)	۳ (۰.۴/۵)	۵۰ (۰.۷۳/۵)

* درصدی از مطالعات به صورت یک سوکور، دو سوکور و سه سوکور بودند، از این رو سطر آخر جدول جمع دو سطر بالای آن نیست.

جدول شماره ۵- توزیع فراوانی مطلق و نسبی گزارش انتخابی پیامدها در مطالعات کارآزمایی منتشره در مجلات تخصصی پرستاری و مامایی در سال ۱۳۸۹

کم تورش	پر تورش	مبهم	
۵۱	۵	۱۲	
۵۸	۶	۴	
۵۷	۱۰	۱	
گزارش انتخابی پیامدها* (درصد)	۳۵ (۰.۵۱/۵)	۱۷ (۰.۲۵)	۱۶ (۰.۲۳/۵)

* نتیجه کلی جدول نشانگر وضعیت کلی مطالعات است، به گونه‌ای که برخی از مطالعات دارای بیش از یک تورش در مطالعه بودند و از این رو، سطر آخر جدول جمع سطرهای بالا نیست.

جدول شماره ۶- توزیع فراوانی مطلق و نسبی دیگر تورش‌ها در مطالعات کارآزمایی منتشره در مجلات تخصصی پرستاری و مامایی در سال ۱۳۸۹

دیگر تورش‌ها	کم تورش	پُر تورش	مبهم	کل
Design- specific	۶۱	۶	۱	۶۸
Early stopping	۶۴	۰	۴	۶۸
Baseline imbalance	۴۳	۹	۱۶	۶۸
Blocked randomization in unblinded groups	۶۰	۸	۰	۶۸
Co- intervention	۶۸	۰	۰	۶۸
Contamination	۵۷	۲	۹	۶۸
Non-sensitive tool	۵۲	۴	۱۲	۶۸
Deviation from protocol	۰	۰	۶۸	۶۸
* (%) OTHER BIAS	۲۵	۲۴	۱۹	۶۸
	(۰.۳۶/۷)	(۰.۳۴/۳)	(۰.۲۸)	(۰.۱۰۰)

* نتیجه کلی جدول نشانگر وضعیت کلی مطالعات است، به گونه‌ای که برخی از مطالعات بیش از یک تورش در مطالعه داشتند و بنابراین، سطر آخر جدول جمع سطرهای بالا نیست.

مانند تصادفی‌سازی بر اساس روزهای زوج و فرد استفاده کرده بودند. هیگانهز یک ضعف عمده در همه روش‌های سیستماتیک را قابل پیش‌بینی بودن برنامه توالی تصادفی می‌داند، به گونه‌ای که روش مورد استفاده باعث شود تا اعمال یک برنامه تخصیص تصادفی ناممکن شود (۱۰).

نتایج مطالعه حاضر با نتایج مطالعه شولتز (۱۹۹۴) همخوانی

بحث

تولید توالی تصادفی

نتایج مطالعه نشان داد که ۶۷٪ مطالعات بررسی‌شده، روش مورد استفاده برای تولید توالی تصادفی را گزارش نکردند و تنها ۲۲٪ مطالعات از روش درست و ۱۰٪ از روش‌های سیستماتیک،

روش تخصیص تصادفی خود را به‌خوبی شرح داده بودند (۱۷)، (۱۶). همچنین، به نظر می‌رسد که نوع مداخله به‌کاررفته در کاربرد و به‌دنبال آن، گزارش دهی تخصیص تصادفی بسیار مؤثر است. کروسیتی (۲۰۱۰) در مطالعه خود می‌گوید احتمال آن که تخصیص تصادفی در مطالعاتی با مداخله آموزشی/رفتاری مبهم و یا با تورش بالا گزارش شود، چهار برابر مطالعاتی با مداخله دارویی است (۴). از آنجا که بیشتر مداخله‌ها (۶۵٪) در مطالعات بررسی‌شده از نوع مداخله‌های آموزشی/رفتاری بود و مداخله‌های دارویی و یا بررسی اثربخشی یک ماده خاص درصد کمتری را تشکیل می‌دادند. بنابراین، یکی از علل گزارش‌دهی ضعیف تخصیص تصادفی در مطالعات بررسی‌شده را نوع مداخله به‌کاررفته در این گونه مطالعات می‌توان دانست.

باید در نظر داشت که گزارش‌دهی ضعیف تخصیص تصادفی تنها در مرحله انتشار مقاله اتفاق نمی‌افتد. دوروکس (۲۰۰۴) به این نتیجه رسید که بیشتر مقالاتی که وضعیت تخصیص تصادفی در آن‌ها مبهم است، در ۶۰٪ طرح‌های پژوهشی آن‌ها نیز این وضعیت مبهم گزارش می‌شود (۱۳).

کورسازی

نتایج مطالعه نشان داد که بیشتر مطالعات (۷۳/۵٪) درباره گزارش‌دهی کورسازی در مقالات خود سکوت کرده بودند و ۲۲٪ مطالعات تورش پایینی داشتند. نتایج به‌دست‌آمده مشابه با مطالعه ریوس و همکاران (۲۰۰۸) بود، به‌گونه‌ای که ۲۳٪ مطالعات در حیطه اندوکرینولوژی به کورسازی اشاره کرده‌بودند (۱۷). اما گزارش کورسازی در مطالعات درماتولوژی بیشتر از مطالعه کنونی بود و بیش از نیمی از RCT‌ها به استفاده از کورسازی اشاره کرده بودند (۱۶). کروستی و همکاران (۲۰۱۰) به ۸۱٪ گزارش کورسازی در بررسی کارآزمایی‌های منتشرشده در مجلات اطفال اشاره کرده‌است. این تفاوت را می‌توان بر اساس استدلال کروسیتی این‌گونه بیان کرد که احتمال آن که کارآزمایی‌هایی که تأثیر یک مداخله آموزشی/رفتاری را بررسی می‌کنند، سه برابر تورش در زمینه کورسازی را مبهم و یا بالا گزارش کنند، نسبت به کارآزمایی‌هایی که تأثیر داروها را بررسی می‌کنند، سه برابر بیشتر است. احتمالاً این مسئله بدان دلیل است که کورسازی شرکت‌کنندگان و محققان در کارآزمایی‌هایی با مداخله‌های آموزشی/رفتاری به‌شدت مشکل و یا حتی به دلیل ماهیت

دارد. شولتز در بررسی کیفیت تصادفی‌سازی در مطالعات RCT منتشرشده در چهار مجله زنان و مامایی به این نتیجه رسید که به‌رغم آن که مطالعات مدعی استفاده از تصادفی‌سازی بودند، تنها ۳۲٪ مطالعات به روشی درست در تولید توالی تصادفی اشاره کرده بودند و ۱۱٪ از روش‌های غیر تصادفی (سیستماتیک متد) استفاده کرده بودند (۳).

از سوی دیگر، در کارآزمایی‌هایی با حجم نمونه بالا (معمولاً بیش از ۱۰۰ نفر در هر گروه)، تصادفی‌سازی ساده باعث ایجاد گروه‌هایی با حجم نمونه نسبتاً مشابه می‌شود، اما در کارآزمایی‌هایی با حجم نمونه کم استفاده از تصادفی‌سازی می‌تواند باعث توزیع ناهمسان نمونه‌ها به گروه‌ها شود (۱۰). ۸۰٪ مطالعات بررسی‌شده دارای گروه‌هایی با حجم نمونه مساوی بودند و این میزان بیش از میزانی بود که بتوان به‌تنهایی به شانس نسبت داد. این در حالی است که انتظار می‌رود حداکثر ۵۰٪ از کارآزمایی‌ها با تصادفی‌سازی ساده، دارای حجم نمونه‌ای برابر در هر گروه باشند. این مشاهده پیشنهاد می‌کند که در برخی از کارآزمایی‌ها ممکن است تخصیص افراد به گروه‌ها صرفاً به‌دنبال یک توالی تصادفی غیر قابل پیش‌بینی انجام نشده‌باشد و ممکن است محققان به اشتباه تصور کنند که یک توزیع همسان در گروه‌ها به افزایش اعتبار و مقبولیت بیشتر کارآزمایی آن‌ها منجر خواهد شد (۱۵، ۱۶) که این مسئله خود دلیلی بر ضعف در روش‌های مورد استفاده برای تولید توالی تصادفی در بیشتر مطالعات بررسی‌شده است.

تخصیص تصادفی

از ۶۸ مطالعه بررسی‌شده ۷ مطالعه از روش سیستماتیک برای تصادفی‌سازی نمونه‌ها استفاده کرده‌بودند. روش‌هایی همچون شماره تخت بیمار، یک هفته در میان و غیره که امکان تخصیص تصادفی به‌دلیل آشکارشدن توالی تصادفی وجود ندارد. سه مطالعه (۴/۵٪) به استفاده از نامه‌های در بسته با توالی تصادفی برای تخصیص تصادفی اشاره کرده‌بودند و دیگر مطالعات (۷۲٪) نیز ا به نحوه تخصیص نمونه‌ها به گروه‌ها اشاره نکرده‌بودند. البته، گزارش‌دهی ضعیف تخصیص تصادفی فقط به مجلات پرستاری و مامایی اختصاص ندارد، بلکه گزارش‌دهی درست تخصیص تصادفی در دیگر مطالعات نیز بسیار ضعیف است؛ تنها ۷٪ مطالعات در حیطه درماتولوژی و ۱۱/۲٪ مطالعات در حیطه اندوکرینولوژی

کوروس و همکاران (۲۰۰۲) نیز متوجه شدند که از ۱۰۰ مطالعه‌ای که مدعی استفاده از رویکرد ITT در آنالیز داده‌هایشان بودند، تنها ۴۲ مطالعه واقعاً ITT را به کار برده بودند (۲۰).

پولیت و همکاران (۲۰۰۹) به بررسی ITT در مطالعات کارآزمایی بالینی منتشرشده در مجلات پرستاری پرداخت. از ۱۲۴ مطالعه بررسی‌شده ۱۵/۳٪ مطالعات استفاده از رویکرد ITT در آنالیز داده‌ها را گزارش کرده‌بودند، اما با در نظر گرفتن مطالعاتی که هیچ‌گونه ریزشی نداشتند، در کل نزدیک به نیمی از مطالعات (۴۸/۸٪) از ITT استفاده کرده بودند (۱۹). ادتگیو و همکاران (۲۰۰۰) متوجه شدند که تنها ۱۲٪ از مقالات در زمینه درماتولوژی از ITT استفاده کرده بودند (۱۶) که این مسئله نزدیک به یک‌چهارم مقالات را در مطالعه حاضر شامل می‌شود. در واقع، باید گفت که محققان پرستاری و مامایی از رویکرد ITT بیش از آنچه که از آن اطلاع داشته باشند و یا ادعا کنند، در مطالعات خود استفاده می‌کنند.

از سوی دیگر، حذف افرادی که از مطالعه انصراف می‌دهند و یا مداخله مورد نظر را به‌خوبی انجام نمی‌دهند، در عمل حذف افرادی است که معمولاً از مداخله مورد نظر به‌خوبی نتیجه نگرفته‌اند و این مسئله باعث می‌شود تا نتایج به‌دست‌آمده به نفع گروه مداخله رقم خورد و به معنی داری هر چه بیشتر نتایج منجر شده‌است، به‌گونه‌ای که ۲۸٪ مطالعات مواردی مانند انصراف نمونه‌ها، مرگ و اجرای ناکامل مداخله و یا شرکت ناقص در جلسات را به‌عنوان معیار خروج مطالعه خود گزارش کرده بودند و این عامل باعث شده‌بود تا نتایج ۸۰٪ مطالعات از نظر آماری به نفع گروه مداخله معنی‌دار باشد و تنها ۲۰٪ از مطالعات به معنی‌دار نبودن به‌لحاظ آماری در نتایج خود اشاره کنند.

پولیت یکی از عوامل مؤثر در به‌کارگیری ITT در مطالعات را تعداد نویسندگان برمی‌شمرد. زیرا هرچه تعداد نویسندگان یک مطالعه بیشتر باشد، احتمال استفاده از متخصص آمار و یا اپیدمیولوژی در مطالعه، بیشتر است (۱۹). اگرچه میانگین تعداد نویسندگان در مطالعه کنونی ۴ نفر (حداقل ۱ نفر و حداکثر ۷ نفر) بود، اما تنها در یک‌سوم از مطالعات متخصص آمار و یا اپیدمیولوژیست از جمله نویسندگان مقاله بودند.

هیچ‌یک از ۸ مجله ایرانی پرستاری و مامایی در راهنمای نویسندگان به استفاده از CONSORT اشاره‌ای نکرده بودند و تنها یک مطالعه از دی‌گرام CONSORT در گزارش مقاله خود استفاده کرده بود. استفاده از CONSORT بر انجام ITT و این که باید مراحل نمونه‌گیری در قالب یک دی‌گرام ترسیم شود، تأکید دارد.

مطالعات در برخی موارد، ناممکن است (۴). از آنجا که بیشتر مداخله‌ها (۶۵٪) در حیطه پرستاری و مامایی از نوع مداخله‌های آموزشی/ رفتاری است و مداخله‌های دارویی یا اثربخشی یک ماده خاص درصد کمتری را (۳۵٪) تشکیل می‌دهد. از این رو یکی از علل گزارش‌نشدن کورسازی در ۷۳٪ مطالعات به دلیل ناممکن‌بودن اعمال کورسازی بر شرکت‌کنندگان، محققان و کارکنان بوده‌است. اگرچه این مطالعات در زمینه کورسازی متخصص آمار و یا حتی نویسنده مقاله، محدودیتی نداشتند، اما تنها یک مطالعه به بی‌اطلاعی متخصص آمار از نوع گروه‌ها اشاره کرده‌بود. کورسازی در مطالعات کارآزمایی بالینی آن قدر اهمیت دارد که نبود آن به اغراق در میزان اثر تخمین‌زده به میزان ۹٪ منجر می‌شود (۱۲). البته باید در نظر داشت که گزارش‌نشدن کورسازی را نمی‌توان صرفاً به معنای انجام‌نشدن آن انگاشت. دیویوکس (۲۰۰۴) متوجه شد که نویسندگان به‌طور پیاپی از کورسازی استفاده می‌کنند، هرچند در گزارش‌دهی آن کوتاهی می‌کنند، به‌گونه‌ای که در بررسی ۱۰۵ کارآزمایی بالینی و به‌دنبال آن، تماس با نویسندگان آن مقالات متوجه شد که از ۲۶٪ مطالعاتی که اشاره‌ای به کورسازی شرکت‌کنندگان نکرده بودند، ۲۰٪ عملاً از کورسازی استفاده کرده بودند و یا از ۸۳٪ مطالعاتی که به کورسازی ارزیاب اشاره‌ای نداشتند، ۷۹٪ از کورسازی استفاده کرده بودند (۱۸).

داده‌های ناقص

نتایج مطالعه نشان داد که درصد نمونه‌هایی که در فاصله بین تصادفی‌سازی تا پیگیری نهایی ریزش یافته‌اند، دامنه‌ای بین ۰٪ تا ۴۰٪ داشته‌است. همچنین، ۲۶/۵٪ مطالعات به نظر می‌رسد که هیچ‌گونه ریزشی نداشتند که در این میان ۱۸ (۲۲٪) مطالعه پیامدها را بلافاصله پس از مداخله اندازه‌گیری کرده‌بودند. بنابراین، فرصت کمی برای ریزش نمونه‌ها وجود داشته‌است. همسو با مطالعه حاضر پولیت و همکاران (۲۰۰۹) میزان ریزش نمونه‌ها را بین ۰٪ تا ۶۲٪ در مطالعات کارآزمایی بالینی پرستاری گزارش کردند که همانند مطالعه حاضر، ۲۶/۶٪ از مطالعات ریزش نداشتند (۱۹).

بررسی Intention to treat analysis (ITT)

هیچ یک از مطالعات بررسی‌شده به استفاده از ITT اشاره‌ای نکرده‌اند، اگرچه پژوهش‌ها نشان داده‌اند که استفاده از عبارت «استفاده از آنالیز ITT» به مفهوم به‌کارگیری حقیقی آن نیست.

تنها ۲۳/۵٪ مطالعات منتشرشده دارای پیامدهایی بودند که از نظر آماری معنی‌دار نبودند و این مسئله را می‌توان یکی از علل گزارش‌دهی ضعیف پیامدها در نیمی از مطالعات قلمداد کرد.

دیگر تورش‌ها

برای وجود تورش‌های متعدد و کیفیت پایین گزارش‌دهی در مقالات پرستاری و مامایی را می‌توان دلایل متعددی، مانند موارد زیر برشمرد:

نخست، بیشتر مطالعات بررسی‌شده در مرکز ثبت کارآزمایی‌ها ثبت نشده بودند. دوم، از آنجا که کارآزمایی‌های حیطه پرستاری و مامایی معمولاً حمایت مالی چندان مناسبی ندارند، ممکن است محققان از حداقل حجم نمونه مورد نیاز، استفاده کنند.

سوم، در بیشتر این کارآزمایی‌ها متخصص آمار تنها در دو مرحله تعیین حجم نمونه و آنالیز داده‌ها مشارکت می‌کند، به‌گونه‌ای که از مشاوره وی در زمینه روش نمونه‌گیری، تصادفی‌سازی و ریزش احتمالی نمونه‌ها کمتر استفاده می‌شود. چهارم، این مسئله که که تحقیقات پرستاری و مامایی کمتر از کار گروهی و رویکرد همکارانه برخوردارند، معمولاً چالش‌های پژوهش با یک رویکرد فردگرایانه حل و فصل می‌شود و مفهوم گروه تحقیقاتی، به معنای واقعی آن، کم‌رنگ است، نیز می‌تواند موجب این ضعف شود. همچنین، این موضوع که، مراکز آموزشی تحقیق را از جمله وظایف ثانویه محققان می‌انگارند، به‌گونه‌ای که معمولاً آموزش از وظایف اولیه تلقی و پژوهش به‌عنوان پلی برای ارتقای افراد در محیط‌های علمی در نظر گرفته می‌شود، سبب می‌شود به کمیت مقالات منتشرشده بیش از کیفیت آن‌ها توجه شود.

افزون بر موارد پیش‌گفته، احتمال دارد آگاهی کمتر افراد از انواع تورش‌ها و چگونگی کنترل آن‌ها در کارآزمایی‌ها از عوامل تأثیرگذار در گزارش‌دهی ضعیف کارآزمایی‌ها باشد. البته، هیچ‌یک از علل گفته شده تداوم ضعف در تحقیقات توجیه نمی‌کنند و انجام‌شدن پژوهش‌های بیشتر درباره بررسی ارتباط عوامل پیش‌گفته با کیفیت مطالعات در حیطه‌های تحقیقات پرستاری و مامایی، ضروری است.

از نقاط قوت این مطالعه، استفاده از ابزار بررسی کیفیت کارآزمایی‌های کاکرین بود که محققان در مرورهای سیستماتیک از آن بسیار استفاده کرده‌اند. این ابزار نه تنها یک ابزار جامع است،

در واقع، یکی از اهداف این دی‌گرام مستندسازی انجام آنالیز ITT و یا انجام‌نشدن آن است (۲۱).

گزارش انتخابی پیامدها

نتایج مطالعه حاضر نشان داد که از ۶۸ مطالعه بررسی‌شده نیمی از مطالعات همه پیامدها را ظاهراً گزارش کرده‌اند. چنان و همکاران (۲۰۰۴) این گونه استنتاج می‌کنند که گزارش انتخابی پیامدها در انتشار مطالعات RCT بسیار رخ می‌دهد و پیامدهای اولیه‌ای که در طرح‌های پژوهشی بیان می‌شوند، با آنچه که در مقالات منتشر می‌شوند، در ۴۰٪ کارآزمایی‌ها متفاوت است. به‌گونه‌ای که تنها ۳۱٪ از پیامدهایی که اثربخشی یک مداخله را بررسی می‌کنند و ۵۹٪ از پیامدهایی که اثرات مضر یک مداخله را بررسی می‌کنند، به‌طور ناقص در مقالات منتشر می‌شوند (۲۲). همچنین، ثبت کارآزمایی در مرکز ثبت کارآزمایی‌ها به محققان کمک می‌کند تا هرگونه گزارش انتخابی پیامدها را بهتر بررسی کنند.

هرچند که نبود گزارش‌دهی و ثبت مطالعات در مرکز ثبت کارآزمایی‌ها قطعه گمشده در بیشتر مطالعات بود، اما ثبت این مطالعات همچنان گزارش انتخابی پیامدها در این گونه کارآزمایی‌ها نیز مرسوم است. در مطالعه‌ای توسط متینو (۲۰۰۹) که به مقایسه ۱۴۷ کارآزمایی ثبت‌شده با مقالات منتشرشده از آن‌ها پرداخت به این نتیجه رسید که بین پیامدهای اولیه موجود در کارآزمایی ثبت شده با پیامدهایی منتشر شده ۳۱٪ تفاوت وجود دارد. متینو (۲۰۰۹) گزارش انتخابی پیامدها را مواردی مانند موارد زیر عنوان می‌کند: حذف پیامدهای اولیه ثبت شده (۱۰/۲٪)، پیامدهایی که در طرح ثبت‌شده، بیان نشده ولی در مقاله گزارش شده‌بودند (۱۵٪)، بررسی پیامد اولیه در زمانی متفاوت از آنچه ثبت شده است (۲/۷٪)، انتشار پیامد ثانویه به‌عنوان پیامد اولیه (۵/۴٪) و انتشار پیامد اولیه به‌عنوان پیامد ثانویه (۶٪) (۲۳).

چنان (۲۰۰۵) شایع‌ترین علل حذف برخی از پیامدها را محدودیت در حجم نوشتاری مقالات، فقدان اهمیت بالینی و عدم معنی‌دارنبودن به‌لحاظ آماری گزارش می‌کند. پیامدهایی که از نظر آماری معنی‌دار می‌شوند، نسبت به پیامدهایی که از نظر آماری معنی‌دار نمی‌شوند، از شانس بیشتری برای انتشار کامل، برخوردار می‌شوند (۲۴). در مطالعه حاضر نیز بیشتر مطالعات (۷۶/۵٪) دارای پیامدهایی بودند که از نظر آماری معنی‌دار بودند و

قوت نتایج مطالعات را تأیید کند. یافته‌های حاصل از این مطالعه نشان داد که بخش زیادی از این کارآزمایی‌ها دارای تورش بالا و یا مبهم در زمینه‌های تولید توالی تصادفی و تخصیص آن، کورسازی، ریزش نمونه‌ها، گزارش انتخابی پیامدها و دیگر تورش‌ها بودند و باید یادآور شد که گزارش‌ندادن موارد مذکور به ابهام و سردرگمی خوانندگان درباره کیفیت مطالعه منجر می‌شود و تورش در این ابعاد به میزانی قابل توجه اعتبار داخلی نتایج مطالعات را تحت تأثیر قرار می‌دهد. بنابراین، تفسیر یافته‌های این گونه مطالعات باید با احتیاط صورت پذیرد.

با استناد به نتایج به‌دست‌آمده از این مطالعه پیشنهاد می‌شود که نخست، مجلات تخصصی پرستاری و مامایی در ایران نیز همچون دیگر مجلات بین‌المللی، از CONSORT پشتیبانی کنند. این مسئله محققان پرستاری را در بهبود طراحی پژوهش‌ها و گزارش‌دهی مطلوب کارآزمایی‌ها کمک خواهد کرد تا نتایج مطالعات آن‌ها تأثیری مستقیم بر کاربرد بالینی آن‌ها داشته باشد. دوم، هرگونه انحراف از طرح پژوهشی تدوین‌شده، باید در مقالات به‌دقت شرح داده شود تا خواننده بتواند تورش‌های احتمالی در مطالعه را بررسی کند. سوم، انتشار مقالات مربوط به مطالعات کارآزمایی بالینی به ثبت پیشین آن در یکی از مراکز بین‌المللی ثبت کارآزمایی‌های بالینی و یا مرکز ثبت کارآزمایی بالینی ایران، (www.irct.ir) برای شفاف‌سازی برای عموم مردم و محققان، مشروط شود. چهارم، برای محدودسازی تورش گزارش انتخابی پیامدها، محققان و سردبیران مجلات باید اطمینان بدهند که همه پیامدهای مورد نظر در طرح پژوهشی، ورای نتایج آن‌ها، به‌طور کامل گزارش شده‌اند. در پایان، پیشنهاد می‌شود طرح‌های پژوهشی‌ها همزمان با مقالات در مجلات تأیید و پذیرش شوند.

تشکر و قدردانی

مقاله حاضر نتیجه طرح تحقیقاتی مصوب دانشگاه علوم پزشکی اصفهان به شماره ۲۹۰۱۹۹ در تاریخ ۲۵/۷/۹۰ است. نویسندگان از معاونت تحقیقات دانشگاه برای تأمین بودجه طرح تشکر می‌کنند.

بلکه راهنمای کامل آن باعث‌شده تا اختلاف آرا در میان بررسی‌کنندگان مطالعات کاهش یابد.

چندین محدودیت نیز در پژوهش حاضر وجود داشت. نخست، کیفیت روش‌شناسی کارآزمایی‌ها در این مطالعه به‌طور مستقیم بررسی نشد، زیرا تنها مقالات منتشرشده از مطالعه بررسی شدند و اطلاعات به‌دست‌آمده در تماس با نویسندگان و یا اطلاعات موجود در طرح‌های پژوهشی آن‌ها بررسی نشد و از آنجا که ممکن است جزئیات روش‌شناختی مهمی در مقالات منتشرشده حذف شده‌باشد، کیفیت گزارش‌دهی تنها جانشین کیفیت واقعی مطالعه در نظر گرفته شد. با این وجود به دلیل آن که مقاله منتشرشده معمولاً تنها منبع علمی برای کارکنان بهداشت و درمان و دیگر محققان برای داوری در مورد اعتبار و تعمیم‌پذیری نتایج در نظر گرفته می‌شود (۱۷)، به‌رغم وجود این محدودیت، کیفیت گزارش‌دهی مقالات دارای ارزیابی قابل توجه است.

دوم، این که فقط مطالعات تخصصی پرستاری و مامایی وارد شده‌اند، ممکن است تعمیم‌پذیری نتایج مطالعه ما را تحت تأثیر قرار دهد. همچنین، تنها RCT‌های حیطه پرستاری و مامایی که در مجلات تخصصی پرستاری و مامایی منتشرشده بودند، بررسی شد، درحالی که دیگر RCT‌های مرتبط با مطالعات پرستاری و مامایی که در دیگر مجلات علمی و پژوهشی ایرانی منتشر شده‌بودند، در مطالعه وارد نشدند. سوم، برخی از مطالعات منتشرشده در مجله زنان و زایمان صرفاً در حیطه مداخلات پرستاری و مامایی نبودند، اگرچه اغلب نویسندگان این مطالعات محققان مامایی بودند که با متخصصان زنان و مامایی همکاری پژوهشی داشتند. چهارم، همه مجلات بررسی‌شده ایرانی بودند و ممکن است تعمیم‌پذیری نتایج مطالعه را به مجلات فراملی تحت تأثیر قرار دهد. هرچند از ۸ مجله بررسی‌شده یک مجله به‌صورت کاملاً انگلیسی منتشر می‌شد و ۷ مجله به زبان فارسی منتشر می‌شدند و دارای چکیده انگلیسی نیز بودند، اما همه مجلات بررسی‌شده در بانک‌های اطلاعاتی مهم بین‌المللی نمایه می‌شدند و از این رو نباید این گونه تصور شود که بازخورد کیفیتی این مطالعات صرفاً تأثیرات ملی دارد، بلکه بهبود کیفیت گزارش‌دهی این مقالات می‌تواند تأثیرات فراملی نیز داشته باشد و از نتایج آن‌ها در تصمیم‌گیری‌های بالینی و راهنماها و کتاب‌های پرستاری و مامایی استفاده کرد.

نتیجه‌گیری

کارآزمایی‌های کنترل‌دار تصادفی در تحقیقات پرستاری و مامایی بسیار استفاده می‌شوند و گزارش درست مطالعات می‌تواند

منابع

- Sohrabi M. Blinding and selection bias in randomized controlled trial: An article critique. *Pejouhesh* 2009; 33: 64-5.
- Boutron I, Estellat C, Ravaud P. A review of blinding in randomized controlled trials found results inconsistent and questionable. *J Clin Epidemiol* 2005; 58: 1220-6.
- Schulz Kf, Chalmers I, Grimes Da, Altman Dg. Assessing the quality of randomization from reports of controlled trials published in obstetrics and gynecology journals. *JAMA* 1994; 272: 125-8.
- Crocetti Mt, Amin Dd, Scherer R. Assessment of risk of bias among pediatric randomized controlled trials. *Pediatrics* 2010; 126: 298-305.
- Attia AM. Bias in RCTs: confounders, selection bias and allocation concealment. *Middle East fertility society journal* 2005; 10: 258.
- Beller Em, GebSKI V, Keech Ac. Randomisation in clinical trials. *Med J Aust* 2002; 177: 565-7.
- Schulz Kf, Grimes DA. Generation of allocation sequences in randomised trials: chance, not choice. *The Lancet* 2002; 359: 515-19.
- Moher D, Pham B, Jones A, et al. Does quality of reports of randomised trials affect estimates of intervention efficacy reported in meta-analyses? *The Lancet* 1998; 352: 609-13.
- Schulz Kf, Grimes DA. Allocation concealment in randomised trials: defending against deciphering. *The Lancet* 2002; 359: 614-18.
- Higgins Jpt, Green S, Collaboration C. *Cochrane handbook for systematic reviews of interventions*. 5 ed. Wiley online library, 2008.
- Schulz Kf, Grimes DA. Blinding in randomised trials: hiding who got what. *The Lancet* 2002; 359:696-700.
- Pildal J, Hro'bjartsson A, JøRGENSEN KJ, HILDEN J, ALTMAN DG, GøTZSCHE PC. Impact of allocation concealment on conclusions drawn from meta-analyses of randomized trials. *Int J Epidemiol* 2007; 36: 847-57.
- Devereaux Pj, Manns Bj, Ghali Wa, Quan H, Guyatt GH. The reporting of methodological factors in randomized controlled trials and the association with a journal policy to promote adherence to the Consolidated Standards of Reporting Trials (CONSORT) checklist. *Control Clin Trials* 2002; 23: 380-88.
- Hutton JI, Williamson PR. Bias in metaanalysis due to outcome variable selection within studies. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 2000; 49: 359-70.
- Schulz Kf, Grimes DA. Unequal group sizes in randomised trials: guarding against guessing. *The Lancet* 2002; 359: 966-70.
- Adetugbo K, Williams H. How well are randomized controlled trials reported in the dermatology literature? *Arch Dermatol* 2000; 136: 381.
- Rios Lp, Oduyungbo A, Moitri Mo, Rahman Mo, Thabane L. Quality of reporting of randomized controlled trials in general endocrinology literature. *J Clin Endocrinol Metab* 2008; 93:3810-3816.
- Devereaux Pj, Choi Ptl, El-Dika S et al. An observational study found that authors of randomized controlled trials frequently use concealment of randomization and blinding, despite the failure to report these methods. *J Clin Epidemiol* 2004; 57: 1232-6.
- Polit Df, Gillespie Bm. The use of the intention-to-treat principle in nursing clinical trials. *Nurs Res* 2009; 58: 391.
- Kruse RI, Alper Bs, Reust C, Stevermer Jj, Shannon S, Williams RH. Intention-to-treat analysis: Who is in? Who is out? *J Fam Pract* 2002; 51: 969-71.
- Schulz K, Altman D, Moher D. CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. *BMC Med* 2010; 8:18.
- Chan Aw, Kroleza-Jeric K, Schmid I, Altman DG. Outcome reporting bias in randomized trials funded by the Canadian Institutes of Health Research. *CMAJ* 2004; 171: 735-40.
- Mathieu S, Boutron I, Moher D, Altman Dg, Ravaud P. Comparison of registered and published primary outcomes in randomized controlled trials. *JAMA* 2009; 302: 977-84.
- Chan Aw, Altman DG. Identifying outcome reporting bias in randomised trials on PubMed: review of publications and survey of authors. *BMJ* 2005; 330: 753.

Risk of Bias in Randomized Controlled Trials Published in Iranian Nursing and Midwifery Journals in 2010

Mohammady M¹, Toghian Chaharsougi N², Adoli S³

1- MSc in nursing, lecturer, Nursing & midwifery faculty, Isfahan University of Medical Sciences, Isfahan, Iran

2- MSc in pediatric nursing, Isfahan, Iran

3- Ph.D. in nursing, assistant professor, Nursing & midwifery care research center, Isfahan University of Medical Sciences, Iran

Corresponding author: Toghian Chaharsougi N., toghiann@gmail.com

Background & Objectives: Randomized controlled trials are the most reliable type of study to be able to compare different interventions in scientific research. The introduction of bias into the design and conduct of randomized controlled trials can seriously affect the accuracy of the results and led to the results be invalid. The aim of this study was to assess the bias in randomized controlled trials published in Iranian nursing and midwifery journals.

Methods: We searched all published articles in eight professional nursing and midwifery Iranian journals in 2010. Risk of bias in RCTs was assessed using of Cochrane group risk of bias standard tool by two reviewers independently.

Results: After reviewing the eight journals, 68 randomized controlled trials were retrieved. The majority of studies had unclear or high risk of bias. Random sequence generation in 22%, concealment allocation in 4.5%, blinding in 22%, incomplete outcome data in the 35.3%, selective outcome reporting in 51.5%, and other bias in 36.7% of the studies were in low risk of bias.

Conclusion: A large proportion of the studies had high or unclear risk of bias. It was considerable that none reporting bias led to uncertainty reader from study quality. Bias might significantly affect the internal validity of the studies results; therefore in interpretation of the findings of these studies caution should be taken.

Keywords: Randomized controlled trial, Risk of bias, Intention to treat analysis, Quality appraisal