

# اتخاذ تصمیم های آگاهانه: تحقیقات اثربخشی (Effectiveness)

ارزیابی نقاط قوت و ضعف طرح های مطالعاتی  
و روش های تحلیلی برای پژوهش مقایسه ای اثربخشی

کارآمدی (Efficacy)

نشان دهنده ی چگونگی

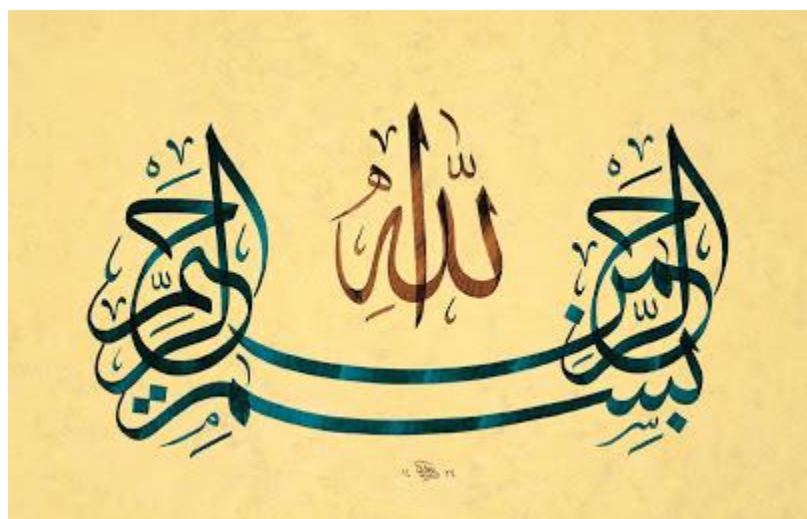
کارکرد مداخلات یا خدمات در

شرایط ایده ال است، در حالیکه

اثربخشی (Effectiveness) به بررسی

چگونگی عملکرد مداخلات با خدمات در

شرایط دنیای واقعی می پردازد.



اتخاذ تصمیم های آگاهانه : تحقیقات اثربخشی (Effectiveness)

نویسندگان : گروهی از نویسندگان

مترجم : فرشید علاء الدینی

ناشر : مولف

تیراژ : 1000 جلد

نوبت چاپ : اول

قیمت : 30000 تومان

سال 1397

۱	مقدمه
۴	طرح‌های مطالعاتی تجربی
۴	کارآزمایی‌های بالینی عملی .....
۷	طرح‌های متقاطع .....
۱۱	کارآزمایی N:1 از کنترل شده تصادفی .....
۱۴	کارآزمایی کنترل شده تصادفی خوشه‌ای .....
۱۸	طرح‌های با شروع تاخیری .....
۲۳	روش‌های تجربی
۲۳	طرح‌های انطباقی و روش‌های بیزین .....
۲۷	طرح‌های مطالعاتی غیرتجربی
۲۷	مطالعات هم‌گروهی (کوهورت) و مورد شاهدی .....
۳۲	روش‌های غیرتجربی
۳۲	طراحی‌های کاربر جدید .....
۳۶	محدودسازی .....
۳۸	تحلیل زیرگروه .....
۴۰	روش‌های نمره‌گرایش (Propensity Score Methods) .....
۴۴	روش‌های متغیرابزاری (Instrumental variable - IV) .....
۴۷	تحلیل حساسیت .....
۴۸	اطلاعات خارجی .....
۵۲	واژه‌نامه

اتخاذ تصمیم‌های آگاهانه : تحقیقات اثربخشی (Effectiveness)

ارزیابی نقاط قوت و ضعف طرح‌های مطالعاتی و روش‌های تحلیلی برای پژوهش  
مقایسه‌ای اثربخشی

کارآمدی (Efficacy) نشان‌دهنده‌ی چگونگی کارکرد مداخلات یا  
خدمات در شرایط ایده-ال است، درحالی‌که اثربخشی (Effectiveness)  
به بررسی چگونگی عملکرد مداخلات یا خدمات در شرایط دنیای واقعی  
می‌پردازد،

## مقدمه

پژوهش مقایسه‌ای اثربخشی<sup>۱</sup> (CER) بدین صورت تعریف می‌شود: «ایجاد و اجرای پژوهشی که مزایا و معایب مداخلات و راهبردهای انجام شده به منظور پیشگیری، تشخیص، درمان و نظارت بر بیماری‌ها در شرایط دنیای واقعی را مقایسه می‌کند.» هدف CER بهبود پیامدهای<sup>۲</sup> بیمار از طریق ارائه‌ی اطلاعات به تصمیم‌گیرندگانی مانند بیماران، ارائه‌دهندگان خدمات سلامت، سیاست‌گذاران، و پرداخت‌کنندگان هزینه‌های سلامت است. این اطلاعات به افراد نشان می‌دهد که کدام یک از مداخلات برای هر یک از انواع خاص بیماران به صرفه‌تر هستند. به دلیل افزایش گزینه‌های درمانی برای بسیاری از بیماری‌ها، تصمیم‌گیرندگان به دنبال اطلاعات مقایسه‌ای هستند که به انتخاب درمانی آگاهانه کمک کنند. با این حال، یا به دلیل عدم اختصاص بودجه، یا به دلیل تمرکز پژوهش‌های بالینی بر نشان دادن کارآمدی<sup>۳</sup> (سودمندی) [مداخلات درمانی]، اطلاعات اثربخشی مقایسه‌ای اغلب در دسترس نیستند. کارآمدی نشان‌دهنده‌ی چگونگی کارکرد مداخلات یا خدمات در شرایط ایده‌آل است، در حالی که اثربخشی به بررسی چگونگی عملکرد مداخلات یا خدمات در شرایط دنیای واقعی می‌پردازد، شرایطی که طی آن بیماران ممکن است با بیماری‌های پیچیده‌تری دست‌وپنجه نرم کنند. بر اساس تخمین موسسه‌ی پزشکی، کمتر از نیمی از تمامی مراقبت‌های پزشکی در ایالات متحده مورد حمایت شواهد اثربخشی کافی هستند.

هدف CER از بین بردن این شکاف شواهد از طریق تولید اطلاعاتی است که تصمیم‌گیرندگان می‌توانند از آن برای اتخاذ تصمیم‌های درمانی و پوششی آگاهانه استفاده کنند. بنابراین، CER باید به منظور پاسخ به نیازهای دنیای واقعی تصمیم‌گیرندگان طراحی شود. این تمرکز عملی CER به الزامات منحصربه‌فردی برای طراحی و اجرای مطالعات نیاز دارد. برای مثال، سبک و سنگین کردن اعتبار، رابطه، امکان پذیر بودن، بجا و مناسب بودن باید با در نظر گرفتن زمینه‌ی تصمیم‌گیرندگان و تصمیمات خاص مطرح شوند. این ملاحظات منحصربه‌فرد باعث رسیدن به سوالاتی می‌شوند که بر اساس آن‌ها طرح‌ها و روش‌های مطالعاتی مناسب برای سوالات CER مشخص می‌شوند.

درک اینکه بهترین رویکرد برای انجام مطالعه‌ی CER و شرایط مناسب آن کدام اند، هنوز مورد مناقشه‌ی بسیار است. عموماً، رویکردهای CER در دو دسته‌ی کلی قرار می‌گیرند: طرح‌ها و روش‌های مطالعاتی کارآزمایی، و طرح‌ها و روش‌های مطالعاتی غیرکارآزمایی. در طرح‌های کارآزمایی، بیماران به صورت تصادفی (بر اساس شانس، نه بنا به تصمیم پزشک) و بر اساس پروتکل مطالعه، به یک درمان خاص اختصاص داده می‌شوند. در طرح‌های غیرکارآزمایی، بیماران و پزشکان تصمیمات درمانی را بر اساس شرایط واقعی انجام داده، و الگوهای مراقبتی و نتایج درمان مشاهده می‌شوند. برخی عنوان می‌کنند که به منظور پاسخ‌دهی به اکثر سوالات CER به طرح‌های مطالعاتی کارآزمایی

1. Comparative effectiveness research

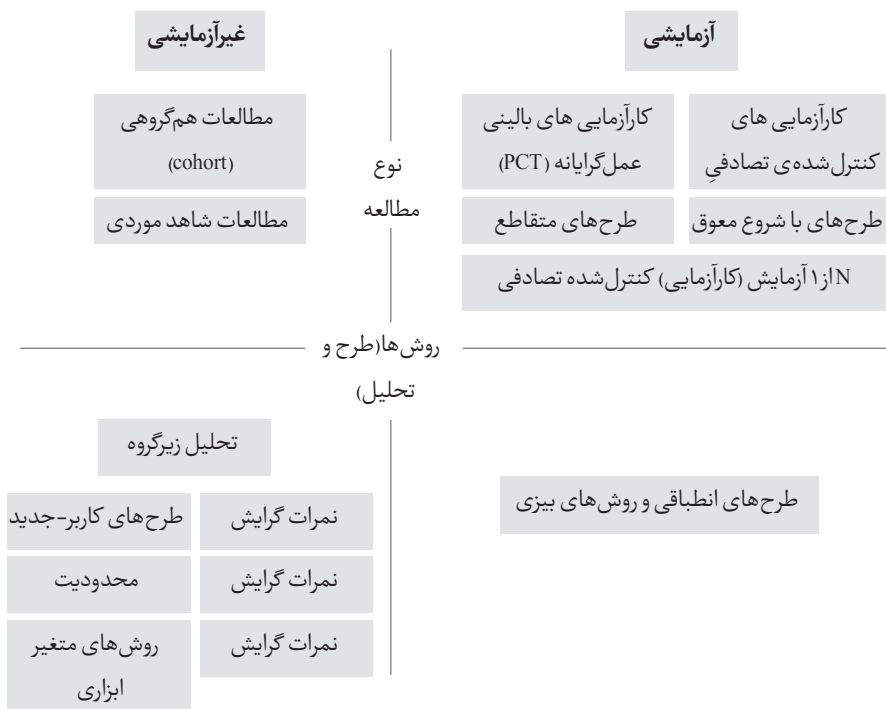
2. Outcomes

3. Efficacy

نیاز است، زیرا تصادفی بودن مطالعه باعث از بین رفتن نگرانی‌های مربوط به سوگیری‌های مسیر (یعنی تمایل متخصص بالینی برای تجویز درمان‌های خاص بر اساس پیش‌آگهی بیمار) شده و با توجه به مخدوشگرهای سنجیده شده و سنجیده نشده (متغیرهای فرعی که ممکن است بر نتیجه‌ی پژوهش تأثیر بگذارند)، به تعادل می‌رسد. دیگران معتقدند که مطالعات غیرکارآزمایی، دربرگیرنده‌ی راه‌هایی هستند که از طریق آن‌ها می‌توان سوگیری مسیر و دیگر مخدوشگرهای موجود در طرح و/یا تحلیل را حذف نمود. از نظر آن‌ها، این نوع مطالعات جایگزین مهمی برای مطالعات تصادفی هستند. در عمل، هر یک از رویکردهای پژوهشی دارای مزایا و معایبی هستند و رویکرد پژوهشی برای سؤال CER باید بر اساس ویژگی‌ها یا خصایص خاص سؤال پژوهش انتخاب شود.

هدف این کتاب ارائه‌ی توصیفی مختصر از طرح‌ها و روش‌های مطالعاتی کارآزمایی و غیرکارآزمایی‌ای است که ممکن است برای پرداختن به سؤالات مطالعه‌ی CER از آن‌ها استفاده شود. هر طرح یا موضوع تحلیل توصیف شده، و نقاط قوت و محدودیت‌های مربوط به آن نیز مطرح شده‌اند. مثال‌هایی به‌منظور نشان دادن استفاده‌ی عملی از روش‌های مذکور در پژوهش‌های قبلی نیز آورده شده‌اند. اگرچه این سند تجویزکننده‌ی روش‌شناسی برای سؤالات خاص پژوهش CER نیست، بخشی از تلاشی بزرگ‌تر برای ایجاد یک رویکرد نظام‌مند جهت تعیین بهترین روش‌ها برای پرداختن به سؤالات CER است. در شکل فعلی، این سند اطلاعات لازم برای پژوهشگران و مصرف‌کنندگان منابع پژوهشی جهت درک نقاط قوت و محدودیت‌های نسبی طرح‌ها و رویکردهای تحلیلی مختلف CER را فراهم کرده و مشخص می‌کنند که چگونه استفاده‌ی آن‌ها از هر یک از این رویکردها و طرح‌ها ممکن است بر نتایج مطالعه و تفسیر آن تأثیر بگذارد.

این کتاب در چهار بخش تنظیم شده است: طرح‌های مطالعاتی کارآزمایی؛ روش‌های تجربی؛ طرح‌های مطالعاتی غیرکارآزمایی؛ و روش‌های غیرتجربی. ساختار این کتاب در شکل ۱ نشان داده شده است. واژه‌نامه‌ی لغات پراستفاده در پایان آورده شده است.



تصویر شماره ۱ روش‌ها و انواع مطالعات کارآزمایی و غیرکارآزمایی.



## طرح‌های مطالعاتی تجربی

### کارآزمایی‌های بالینی عملی

مقدمه:

کارآزمایی‌های بالینی عملی در حقیقت کارآزمایی‌های کنترل‌شده تصادفی‌ای هستند که به منظور تعیین خطرات، فواید، و هزینه‌های یک مداخله در زمان استفاده از آن حین کار بالینی معمول، طراحی می‌شوند. این کارآزمایی‌ها اغلب با کارآزمایی‌های تبیینی (explanatory trials) تضاد دارند. کارآزمایی‌های تبیینی به منظور تعیین اثربخشی یک مداخله‌ی بالینی در شرایط بهینه و ایده-آل (مثلاً، در جمعیتی که به دقت انتخاب شده و بیماران آن به درمان پایبند هستند) انجام می‌شود.

### کارآزمایی‌های عملی در برابر کارآزمایی‌های تبیینی

برخلاف کارآزمایی‌های تبیینی، کارآزمایی‌های بالینی عملی معمولاً شامل دامنه وسیع‌تری از بیماران (از طریق کاهش تعداد معیارهای ورود و خروج در مطالعه)، دامنه وسیع‌تری از مکان‌های مطالعه (با استفاده از سایت‌های مبتنی بر جامعه و غیرآکادمیک)، و نتایجی است که مطابق با نیازهای شواهد تصمیم‌گیرندگان (مثلاً بیماران، متخصصان بالینی، پرداخت‌کنندگان، و سیاست‌گذاران) هستند. شایان ذکر است که کارآزمایی‌های تبیینی و عملی مفاهیم متمایزی نیستند. در عوض، کارآزمایی‌ها ممکن است در بردارنده‌ی مؤلفه‌های تبیینی و عملی به درجات مختلفی باشند. برای مثال، یک کارآزمایی ممکن است دارای معیارهای شایستگی (برای ورود) سفت و سختی باشد. این یعنی تنها شامل بیماران با خطر بالا، مطیع درمان، و پاسخگو به درمان (بخش تبیینی طیف) باشد، اما دارای نظارت حداقلی (یا فقدان نظارت) بر تبعیت پزشک از پروتکل مطالعه بوده و جلسات پیگیری رسمی نیز برای بیمار تعیین نشده باشد (بخش عملی طیف).

اینکه یک کارآزمایی تا چه حد «عملی» یا «تبیینی است» به ارزیابی تمامی مؤلفه‌های کارآزمایی بستگی دارد.

## موارد استفاده‌ی توصیه‌شده

کارآزمایی‌های تبیینی بهترین وسیله برای تعیین کارآمدی و سودمندی هستند، درحالی‌که کارآزمایی‌های بالینی عملی مناسب درک اثربخشی، یا مؤثر بودن یک مداخله‌ی بالینی در دنیای واقعی هستند. به‌منظور ارائه‌ی شواهدی که نشان‌دهنده‌ی فایده‌ی یک مداخله در شرایط دنیای واقعی است، می‌توان در بطن کارآزمایی‌های سودمندی، ویژگی‌هایی عملی را جای داد. ویژگی‌هایی که می‌توانند در مطالعه لحاظ شوند شامل مواردی این‌چنینی هستند: نام‌نویسی از بیمارانی که بیشتر شبیه به بیماران دنیای واقعی هستند و قرار است بعد از تأیید درمان آن را دریافت کنند. دربرداشتن طیف وسیع‌تری از نتایج، با تأکید بیشتر بر وضعیت کارکردی، کیفیت زندگی، و اثرات طولانی‌مدت.

## مسائل و مشکلات بالقوه.

ممکن است اثر درمان در کارآزمایی‌های بالینی عملی در مقایسه با کارآزمایی‌های تبیینی همگنی کمتری (فقدان همسانی) داشته باشد. دلیل این موضوع گسترده‌تر بودن معیارهای صلاحیت و دربرداشتن پزشکان مختلف با درجات متفاوتی از تخصص است. این تنوع موجود در جمعیت بیمار و شرایط اعمال درمان باید هنگام طراحی کارآزمایی و تحلیل موردتوجه قرار بگیرد. همچنین، هدف بسیاری از کارآزمایی‌های بالینی عملی دربرگرفتن کلینیک‌ها و مراکز سلامت عمومی است که به‌طور سنتی در پژوهش‌های بالینی شرکت نکرده‌اند. بنابراین، پژوهشگرانی که دست‌به‌کار آزمایشی بالینی عملی می‌زنند، زمان و منابع زیادی را صرف آموزش افراد موجود در این مراکز می‌کنند تا آن‌ها توانایی لازم برای شرکت در کارآزمایی‌ها را بدست آورند.

## قدرت

• کارآزمایی‌های بالینی عملی اثرات درمان در شرایط دنیای واقعی را ارزیابی می‌کنند. نتیجه‌ی مثبت کارآزمایی بالینی عملی می‌تواند به نفع کار عملی باشد، زیرا شواهدی را ارائه می‌دهد که بر طبق آن درمان/مداخله در شرایط درمانی معمول اثربخش بوده است.

## محدودیت‌ها

- نتیجه‌ی منفی کارآزمایی بالینی عملی نمی‌تواند اطلاعاتی در مورد اثربخش بودن درمان/مداخله در شرایط بهینه ارائه کند.
- شاید لازم باشد که پژوهشگران زمان و منابع بیشتری را صرف کارآزمایی‌های بالینی عملی (در مقایسه با کارآزمایی‌های تبیینی) کنند.
- به دلیل ناهمگنی بالقوه و اثر درمان، کارآزمایی‌های بالینی عملی نسبت به کارآزمایی‌های تصادفی کنترل‌شده چالش‌های طراحی و تحلیل بیشتری دارند.

## مثال‌های انتخابی

• کارآزمایی ملی درمان آمفییزم<sup>۴</sup> (NETT). این کارآزمایی تصادفی کنترل‌شده به مقایسه‌ی جراحی کاهش حجم ریه با مراقبت استاندارد در بیماران مبتلا به آمفییزم پرداخت. این کارآزمایی دارای معیارهای منع (از ورود به مطالعه) حداقلی بوده، در ۱۷ مکان بالینی مختلف صورت گرفته، و معیارهای سنجش آن برای بیماران، متخصصان بالینی، و ارائه‌دهندگان خدمات سلامت آشکار بود: مرگ‌ومیر و حداکثر ظرفیت برای ورزش. به دلیل متغیر بودن تخصص پزشکی و استاندارد مراقبت، از یک طرح عمل‌گراتر استفاده شده است.

• مطالعه‌ی تصویربرداری آینده‌نگر چندمرکزی برای ارزیابی درد سینه<sup>۵</sup> (PROMISE). این کارآزمایی تصادفی کنترل‌شده‌ی بزرگ چندین آزمون غیرتهاجمی مختلف برای بیماری سرخرگ کرونری را مقایسه کرد. به‌طور خاص، این کارآزمایی به بررسی این موضوع پرداخت که آیا برای بیماران شریان کرونری با ریسک پایین-متوسط و درد سینه، یک راهبرد آزمون‌گیری «آناتومیک» اولیه (با استفاده از آنژیوگرافی توموگرافی رایانه‌ای) نسبت به مراقبت معمول یا راهبرد آزمون‌گیری استرس «کارکردی»، از نظر بالینی برتر است یا خیر. این کارآزمایی به ثبت نام از بیماران ادامه می‌دهد و هدف آن رسیدن به ۱۰ هزار بیمار از ۱۵۰ مکان بالینی مختلف است. نقاط پایان کارآزمایی شامل مرگ، آنفارکتوس میوکارد، عوارض شدید پیش از فرایند پزشکی، و بستری شدن به دلیل آنژین ناپایدار است. این کارآزمایی کیفیت زندگی، استفاده از منابع، و مقرون‌به‌صرفه بودن را ارزیابی کرده و امکان ارزیابی اثر فرایند در دنیای واقعی را فراهم می‌کند. به‌منظور اطمینان از عمومیت داشتن نتایج برای طیفی از بیماران، امکان استفاده از دامنه‌ی وسیعی از شرایط درمان، متخصصان پزشکی و انواع آزمون‌های مرتبط با تصمیم‌گیری بالینی داده شد.

4.National Emphysema Treatment Trial

5.Prospective Multicenter Imaging Study for Evaluation of Chest Pain

یک طرح متقاطع به بیماران اجازه می‌دهد تا مانند کنترل خود رفتار کنند و بدین ترتیب، امکان مقایسه درون‌گروهی و بین‌گروهی را فراهم می‌کند. بیماران طی دوره‌های زمانی متوالی، زنجیره‌ای از درمان‌ها را دریافت می‌کنند. در شروع مطالعه، هر بیمار به یک زنجیره اختصاص می‌یابد (مثلاً زنجیره AB در برابر BA). عموماً درمان‌های متوالی با استفاده از یک دوره‌ی خالی (یک دوره‌ی مشخص شده‌ی عدم استفاده [از درمان]، قبل از شروع درمان) از هم جدا می‌شوند. به دلیل اینکه تمامی شرکت‌کنندگان تعداد درمان یکسانی را دریافت می‌کنند (معمولاً با توالی متفاوت)، امکان تعیین نقش ویژگی‌های بیمار در پاسخ به درمان وجود دارد. چندین نوع از طرح‌های متقاطع وجود دارند که اعداد توالی‌ها و دوره‌های زمانی در آن‌ها متفاوت‌اند. هر یک از این طرح‌ها مجموعه الزامات خاص خود را دارند.

### ملاحظات مربوط به طرح‌های متقاطع

هنگام انتخاب بین طرح موازی (سنّتی) و طرح متقاطع، عوامل زیر که تعیین‌کننده‌ی اثربخشی طرح متقاطع هستند باید مدنظر قرار گیرند:

اثرات باقی‌مانده و دوره‌ی زمانی بر نتایج درمان: امکان دارد اثر یک درمان از یک دوره‌ی زمانی به دوره‌ی بعدی انتقال یابد. این اثرات تحت عنوان «اثرات باقی‌مانده» شناخته می‌شوند. همچنین، در حین دوره‌ی بررسی، بیماری ممکن است به صورت طبیعی پیشرفت یا پسرفت کند، یا شدت آن نوسان داشته باشد. این تغییرات تحت عنوان «اثرات دوره‌ی زمانی» شناخته می‌شوند. باید از اندک بودن اثرات باقی‌مانده و دوره‌ی زمانی اطمینان داشته باشیم. در غیر این صورت، یک طرح متقاطع مزایای خود را از دست می‌دهد. به منظور اطمینان از اندک بودن اثرات باقی‌مانده، باید بین دوره‌های درمان فعال، دوره‌های طولانی (به اندازه کافی) بدون درمان وجود داشته باشد.

توالی درمان و گم‌کردن بیمار: در حالت ایده‌آل، توالی تجویز درمان‌ها باید تصادفی باشد. این موضوع باعث حفاظت در برابر سوگیری‌های هشیار و ناهشیار شده و از عدم تفاوت بین بیماران دریافت‌کننده‌ی ترتیب A/B و بیماران دریافت‌کننده‌ی ترتیب B/A اطمینان حاصل می‌کند.

قوانین متقاطع و زمان بندی اندازه‌گیری‌ها: دو نوع از قوانین متقاطع استفاده‌ی بیشتری دارند: در نوع اول، تغییر درمان پس از طول مدت زمان مشخصی (وابسته به زمان) انجام می‌شود. در نوع دوم تغییر درمان بر اساس ویژگی‌های بالینی بیمار (وابسته به حالت بیماری) تعیین می‌شود. به منظور کاهش تأثیر اثرات باقی‌مانده و اثرات دوره‌ی زمانی، این نکات متقاطع باید هم از بی‌ماران و هم از مشاهده‌کنندگان پنهان شوند.

افراد کنار رفته از مطالعه (افرادی که از مطالعه کنار گذاشته شده، یا خود از آن خارج می‌شوند)، داده‌های معیوب و دیگر مشکلات مربوط به داده‌ها: هر بیمار به‌عنوان منبع کنترل خود ایفای نقش

می‌کند و به همین دلیل، یک بیمار به‌تنهایی نسبت بزرگی از اطلاعات کل را ارائه می‌کند. به همین دلیل، نرخ بالای افراد کناررفته یکی از مشکلات مهم این نوع طرح مطالعاتی است. علاوه بر این، نرخ کناررفتنی در طرح‌های متقاطع بیش از طرح‌های موازی است، زیرا به‌منظور ارائه یک نقطه‌ی داده‌ی کامل، بیماران باید حداقل دو درمان را دریافت کنند. اندازه‌ی نمونه‌ی اولیه باید به حد کافی بزرگ باشد تا این اثر را خنثی کند.

تحلیل آماری و اندازه‌ی نمونه: واحد اصلی برای تحلیل آماری، خود بیمار است، نه یک سنجش یگانه. تحلیل‌ها باید بر اساس داده‌های جور (از همان بیمار) انجام شده، و محاسبات اندازه‌ی نمونه باید تغییرپذیری درون-بیمار موجود در خروجی‌ها را مدنظر قرار دهد.

### موارد استفاده‌ی توصیه‌شده

بهترین استفاده‌ی طرح‌های متقاطع، هنگام مطالعه‌ی بیماری‌های مزمن و پایدار است. این طرح‌ها مکرراً به‌منظور مقایسه‌ی درمان‌های جدید و درحال توسعه مورد استفاده قرار می‌گیرند. چنین طرح‌هایی خصوصاً زمانی مفید هستند که بین درمان جدید و درمان استاندارد تفاوت‌های اندکی وجود داشته و اثرات آن‌ها بسیار مشابه است. این طرح‌ها باید برای مطالعاتی مورد استفاده قرار گیرند که در آن‌ها اثرات درمان کوتاه و قابل بازگشت هستند.

### مسائل و مشکلات بالقوه.

برای اینکه طرح‌های متقاطع مؤثر و معتبر باشند، باید تنها در شرایطی بکار گرفته شوند که انتظار می‌رود اثرات باقی‌مانده حداقل و یا صفر، نرخ کناررفتن از مطالعه اندک، و روند بیماری ثابت باشد. استفاده‌ی نامناسب از طرح متقاطع می‌تواند باعث تخمین‌های سوگیرانه‌ی اثر درمان شده و بنابراین، به نتیجه‌گیری‌های نادرست در مورد اثربخشی مقایسه‌ای درمان‌ها ختم شود.

### نقاط قوت

- طرح متقاطع باعث از بین رفتن تغییر بین بیمار می‌شود.
- در مقایسه با یک مطالعه‌ی موازی برای مقایسه‌ی تعداد یکسانی از درمان‌ها، طرح متقاطع نیازمند بیماران کمتری است، زیرا هر واحد کارآزمایی (بیمار) می‌تواند چندین بار مورد استفاده قرار بگیرد. این امر نوعی استفاده‌ی اقتصادی از منابع است.

- بیماران می‌توانند یک درمان را بر دیگری ترجیح دهند، زیرا آن‌ها در یک مطالعه‌ی متقاطع چندین درمان را دریافت می‌کنند.

### محدودیت‌ها

- تضمینی وجود ندارد که دوره‌های بدون درمان، اثرات باقی‌مانده را به‌طور کامل کنترل کنند.
- دوره‌های بدون درمان طولانی ممکن است به‌طور غیرقابل اجتنابی مدت‌زمان کارآزمایی را افزایش دهند.

• نگرانی‌های اخلاقی (اینکه در دوره‌ی عدم درمان، چه مدت می‌توان درمان را از بیمار سلب کرد) و دانش ناکافی (مدت زمان مناسب برای دوره‌ی عدم درمان چقدر است) ممکن است گاهی اوقات باعث دوره‌های عدم درمان ناکافی شوند.

• درون هر واحد یا بیمار، احتمال می‌رود که پاسخ به درمان همبسته باشد (مثلاً، یک پاسخ بیمار به درمان الف با پاسخ همان بیمار به درمان ب همبسته باشد. این یعنی پاسخ‌ها مستقل نیستند). این موضوع باعث بروز پیچیدگی‌های در طرح و تحلیل می‌شود.

### مثال‌های انتخابی

• جوهانسون و همکاران<sup>۶</sup>، ژورنال بررسی درمان سرطان<sup>۷</sup> (۱۹۸۲). پژوهشگران یک کارآزمایی دوسر کور دو دوره‌ای متقاطع انجام دادند تا اثرات درمانی و عوارض دو داروی ضد تهوع را با هم مقایسه کنند. در این کارآزمایی داروی نایلون با داروی متعارف آن دوران، یعنی پروکلروپیرازین مقایسه شد. نمونه‌ی این مطالعه شامل ۲۷ بیمار سرطانی بزرگسال بود که شیمی‌درمانی می‌شدند و علی‌رغم مصرف داروهای ضد تهوع سنتی دچار استفراغ و تهوع کنترل نشده بودند. توالی دوره‌های درمان (A/B یا B/A) به صورت تصادفی انجام شده و قوانین تقاطع بر اساس الگوی متضمن شیمی‌درمانی وضع شدند: بیماران قبل از تقاطع دوز (مقدار) یکسانی از یک داروی شیمی‌درمانی را برای دو دوره‌ی متوالی دریافت می‌کردند. به طور مشابه، از آنجایی که داروهای ضد تهوع مورد کارآزمایی به صورت هم‌زمان با شیمی‌درمانی تجویز می‌شدند، دوره‌ی بدون درمان با زمان توقف بین برنامه‌ی شیمی‌درمانی هم‌زمان شده بود. پژوهشگران یک اثر چشم‌گیر را برای دوره‌ی زمانی<sup>۸</sup> شناسایی کردند (نوسان در شدت یا پیشرفت بیماری) که طی آن نمرات شدیدتر استفراغ در حین دوره‌ی دوم درمان ثبت شده بود. با این حال، تحلیل به شکل واضحی به این نتیجه رسید که اثربخشی نایلون در کاهش تهوع ناشی از شیمی‌درمانی برای بیماران که نسبت به داروهای ضد تهوع متعارف مقاوم بودند، به شکل چشم‌گیری از پروکلروپیرازین بیشتر است. اگرچه پژوهشگران دلیل استفاده از طرح متقاطع را به صورت واضح تشریح نکردند، خود مطالعه حاکی از سودمندی طرح متقاطع است. نویسندگان توانستند با استفاده از تعداد اندکی از بیماران، به یک نتیجه‌گیری معنادار از نظر آماری دست پیدا کنند.

6. Johansson

7. Cancer Treat Rev

8. Period effect

• کوکه و همکاران<sup>۹</sup>، ژورنال درد<sup>۱۰</sup> (۲۰۰۴) پژوهشگران یک کارآزمایی یکسر کور تصادفی متقاطع را انجام دادند تا اثرات ضد درد سه روش متفاوت تحریک الکتریکی عصب از طریق پوست (TENS) را با هم مقایسه کنند. این سه روش از نظر فرکانس و شدت با یکدیگر متفاوت بوده و برای ۱۸۰ بیمار بزرگسال مبتلا به درد مزمن انجام شد که هیچ درمان ضد درد دیگری دریافت نمی‌کردند. درمان‌های TENS دو بار در روز و به مدت ۲ هفته تجویز شده و توسط یک دوره‌ی بدون درمان دوهفته‌ای جدا شدند. نتایج در ابتدا و انتهای هر دوره‌ی دوهفته‌ای ارزیابی شد. شدت درد با استفاده از یک مقیاس آنالوگ بصری معتبر (VAS) سنجیده شد. به دلیل ناهمگنی (عدم تجانس) بین جمعیت بیماران درد مزمن از طرح متقاطع استفاده شد. هر بیمار کنترل خودش بود، موضوعی که ناهمگنی مطالعه را کاهش و قدرت آن را افزایش داد.

• مجوری<sup>۱۱</sup> و همکاران. ژورنال آمریکایی بیماران کلیه<sup>۱۲</sup> (۲۰۰۲). اگرچه بسیاری از گزارش‌ها عنوان می‌کنند که استفاده از دیالیز خنک هنگام درمان همودیالیز دارای اثری حفاظتی بر فشارخون است، کارآزمایی‌های بالینی رسمی برای تأیید این موضوع وجود ندارند. به‌منظور حذف نقش تفاوت‌های درون فردی در نتایج همودیالیز، از یک کارآزمایی باز تصادفی با طرح متقاطع استفاده شد. این کارآزمایی به‌منظور مقایسه‌ی اثر دو فرایند متفاوت برای توازن گرمایی بر فرکانس جلسات همودیالیزی انجام شد که فشارخون پایینی نشانه‌ای آن‌ها را دشوار کرده بود. فرایند تحقیقاتی یا همان دیالیز ایزوترمیک، با فرایند کنترل، یعنی دیالیز ترمونوترال (خنثی از نظر حرارتی) در ۱۱۶ بیمار بزرگسال مقایسه شد. این بیماران حداقل به مدت ۳ ماه تحت درمان همودیالیز بوده و خطر دوره‌های افت فشارخون نشانه‌ای را بروز داده بودند. بیماران برای یک دوره‌ی ورود به مطالعه به مدت یک هفته تحت همودیالیز استاندارد قرار گرفته و سپس به‌صورت تصادفی در یکی از توالی‌های درمانی A/B یا B/A قرار گرفتند. هر دوره‌ی درمان ۴ هفته طول کشید. پژوهشگران فرض را بر این گذاشتند که هیچ اثر باقی‌مانده‌ای از یک دوره‌ی درمانی به دوره‌ی بعد انتقال پیدا نمی‌کند

9. Koke

10. Pain

11. Maggiore

12. Am J Kidney Dis

## کارآزمایی N از ۱ کنترل شده تصادفی

### مقدمه

کارآزمایی N از ۱، طرح مقطعی کنترل شده تصادفی دوسر کوری است که دو یا چند گزینه‌ی درمانی برای یک بیمار را مقایسه می‌کند. در چنین مطالعه‌ای، بیمار تحت چند دوره‌ی متوالی درمان قرار می‌گیرد که طی آن یک مداخله‌ی فعال با یک دارونمای مطابق یا درمان دیگر جور می‌شود. برای هر دوره، ترتیب اعمال درمان فعال یا مقایسه‌کننده به صورت تصادفی (مثلاً با انداختن سکه) انجام می‌شود. در حالت ایده‌آل، هم بیمار و هم متخصص بالینی نسبت به قرار گرفتن در توالی‌ها کور هستند (اطلاعی از آن‌ها ندارند). نتایج مناسب (نتایجی که برای بیمار سودمند بوده و به آسانی توسط او گزارش می‌شوند) اغلب با استفاده از یک پرسشنامه یا دفترچه یادداشت اندازه‌گیری می‌شوند.

### ویژگی‌های کارآزمایی N از ۱

انجام کارآزمایی‌های N از ۱ به شناسایی چندین ویژگی اساسی نیاز دارد:

- بیماری‌ای که مداخله برای آن صورت می‌گیرد باید مزمن و نسبتاً پایدار است.
- نیمه‌عمر مداخله نسبتاً کوتاه بوده یا تغییر انجام گرفته، برگشت پذیر است.
- مداخله دارای شروع و پایانی سریع است.
- اثرات مداخله را می‌توان با استفاده از یک ابراز معتبر خروجی ارزیابی کرد.
- مداخله باعث تغییر در علت بنیادی بیماری نمی‌شود.

متخصصان بالینی در استفاده از این روش برای کار بالینی روزمره عملکرد کندی داشته‌اند. این موضوع می‌تواند دلایل متعددی داشته باشد. محدودیت‌های زمانی، هزینه‌ی طراحی و اجرای طرح، و خطر کلی فقدان آگاهی نسبت به این نوع پژوهش.

### تحلیل و تفسیر داده‌ها

آسان‌ترین روش برای تفسیر داده‌های ناشی از کارآزمایی N از ۱ قرار دادن آن در یک جدول و بررسی آن به صورت دیداری است. اگرچه این روش در معرض سوگیری قرار دارد، می‌تواند در مواردی که تفاوت اثر بین مداخله‌ی فعال و مقایسه‌کننده مدنظر است، متقاعدکننده باشد. روش‌های ساده‌ی آماری نیز می‌توانند برای تفسیر داده‌ها مورد استفاده قرار بگیرند.

همچنین در صورتی که مطالعات به دنبال بررسی مجموعه‌ی یکسانی از مداخلات هستند، می‌توان نتایج چندین کارآزمایی N از ۱ را با هم ترکیب کرد. برخی از مطالعات داده‌های مربوط به تحلیل کارآزمایی N از ۱ را جمع‌آوری کرده و با استفاده از روش‌های مدل‌یابی آماری مختلف آن‌ها را با هم ترکیب کرده‌اند. این نوع تحلیل‌ها به پژوهشگران اجازه می‌دهند تا نتایج آن دسته از کارآزمایی‌های N از ۱ که به صورت تکی مفید هستند را بر یک پژوهش جمعیت محور اعمال کرده و تعمیم‌پذیری آن‌ها را افزایش دهند.



## موارد استفاده‌ی توصیه‌شده

متخصصان بالینی اغلب زمانی از کارآزمایی‌های N از ۱ استفاده می‌کنند که با عدم قطعیت درمانی (مشخص نیست کدام درمان برای بیمار اثربخش است) در یک بیمار خاص مواجه می‌شوند. این رویکرد یک راه جایگزین برای اتخاذ تصمیمات درمانی شخصی‌سازی شده براساس داده‌های عینی، ارزش‌های بیمار، و نتایج بیمار-محور ارائه می‌کند.

## مسائل و مشکلات بالقوه

مسائل و مشکلات این طرح نیز همانند مشکلات مربوط به طرح‌های متقاطع هستند. کارآزمایی‌های N از ۱ تنها باید زمانی استفاده شوند که وضعیت بیماری پایدار بوده، اثرات باقی مانده وجود نداشته و یا بسیار کم باشند. استفاده‌ی نادرست از این طرح ممکن است بردقت نتایج تأثیر منفی بگذارد. علاوه بر این، انجام کارآزمایی‌های N از ۱ هر روز دشوارتر از قبل می‌شود، زیرا زمان لازم برای مشاهده‌ی اثرات درمان افزایش می‌یابد.

## نقاط قوت

- کارآزمایی‌های N از ۱ به پزشکان اجازه می‌دهند تا در کار بالینی، درمان را برای هر بیمار شخصی‌سازی کنند.
- پزشکان و بیماران به صورت مشترک در تصمیم‌گیری شرکت می‌کنند.
- برخی از شواهد موجود نشان می‌دهند که بیماران شرکت‌کننده در کارآزمایی‌های N از ۱ دانش و آگاهی بیشتری از بیماری خود دارند، موضوعی که ممکن است به تعهد بیشتر آن‌ها به درمان و در نتیجه مدیریت بهتر بیماری منجر شود.
- کارآزمایی‌های N از ۱ ممکن است درک پژوهشگر از رابطه‌ی بین مطالعات جمعیت محور، ویژگی‌های فردی، و پاسخ‌های مربوط به درمانشان را ارتقا بخشند.
- هزینه‌ی کارآزمایی‌های N از ۱ به شکل چشم‌گیری از کارآزمایی‌های تصادفی کنترل شده کمترند.

## محدودیت‌ها

- کارآزمایی‌های N از ۱ ممکن است زمان بر بوده و برای یک پزشک (به صورت تکی) هزینه‌بر باشند، خصوصاً در شرایطی که تجربه‌ی مدیریتی وجود نداشته باشد.
- کارآزمایی نیازمند همکاری با همکاران داروساز با تجربه به منظور آماده‌سازی دارونمای مطابق با کارآزمایی است.

## مثال‌های انتخابی

• زاگرو همکاران<sup>۱۳</sup>. ژورنال روماتولوژی<sup>۱۴</sup> (۲۰۰۶). پژوهشگران مطالعه‌ای را ترتیب دادند که طی آن نتایج چندین کارآزمایی N از ۱ با هم ترکیب شدند. هدف این مطالعه مقایسه‌ی درمان‌های مختلف فیبرومیالژی و ارزیابی امکان پذیر بودن و نتایج این نوع طرح مطالعاتی برای پژوهش مؤثر عملی بود. نتیجه‌ی اصلی کارآزمایی‌های N از ۱، نمره‌ی پرسشنامه‌ی اثر فیبرومیالژی (FIQ) بود. این پرسشنامه ابزاری معتبر برای ارزیابی کیفیت زندگی بیماران مبتلا به فیبرومیالژی است. هشت متخصص روماتولوژی ۵۸ بیمار را وارد کارآزمایی‌های N از ۱ کردند. این کارآزمایی‌های به صورت تکی تصادفی، دوسر کور، چند مقطعی انجام شده و هدف آن مقایسه‌ی یک دارو به صورت تکی با درمان ترکیبی بود. از یک داروخانه‌ی مرکزی به منظور آماده‌سازی کیت‌های درمانی مرتب شده به صورت تصادفی، استفاده شد. همان‌گونه که در بالا گفته شد، نمرات FIQ جمع‌آوری شده و با نتایج مشابه از یک کارآزمایی متقاطع سنتی فیبرومیالژی که با استفاده از یک آزمون t انجام شده بود، مقایسه شد.

• لولی<sup>۱۵</sup> و همکاران. ژورنال کلینیکال تراپیوتیکس<sup>۱۶</sup> (۲۰۰۹). پژوهشگران با استفاده از کارآزمایی N از ۱ تصادفی کنترل شده دوسر کور، اثربخشی ترامادول ۵۰ میلی گرم را با دارونما مقایسه کردند. نمونه شامل بیماران مبتلا به روماتیسم مفصلی و بیماری ریه بینایی بود که دچار سرفه‌ی خشک مزمن شده بودند. دارو و دارونما توسط یک پژوهشگر خارجی به صورت کپسول‌هایی کاملاً مشابه آماده شدند. این پژوهشگر هیچ ارتباطی با بیماران نداشته و از نتایج مطالعه نیز آگاه نبود.

---

13. Zucker

14. J Rheumatol

15. Louly

16. Clin Ther

## کارآزمایی کنترل شده تصادفی خوشه ای

### مقدمه

کارآزمایی‌های کنترل شده‌ی تصادفی خوشه ای مطالعاتی هستند طی آن‌ها بیماران بر اساس جغرافیای کاری (مثلاً مراقبان، بیمارستان‌ها، جوامع) دسته‌بندی (خوشه) شده و سپس به صورت تصادفی در گروه مداخله یا کنترل قرار می‌گیرند. کارآزمایی تصادفی کنترل شده‌ی خوشه‌ای مبتنی بر جامعه معمولاً تعداد خوشه‌های اندکی (بیمارستان‌ها) داشته و تعداد بیماران هر خوشه بالا است. این کارآزمایی‌ها به‌عنوان روشی عمل‌گرایانه جهت سنجش اثربخشی یک مداخله در مقیاس بزرگ دیده می‌شوند.

### موارد استفاده‌ی توصیه شده

کارآزمایی‌های تصادفی کنترل شده‌ی خوشه‌ای به‌طور رایجی جهت ارزیابی رساندن خدمات سلامت، اثرات مداخلات آموزشی-تربیتی، یا اثرات تغییرات سازمانی به کار گرفته شده‌اند. اخیراً پژوهشگران پیشنهاد کرده‌اند هنگامی که بین الگوهای کاری رابطه‌ای طبیعی وجود دارد، می‌توان از کارآزمایی‌های تصادفی کنترل شده‌ی خوشه‌ای برای تعیین اثربخشی مقایسه‌ای دو یا چند درمان استفاده کرد. مطالعات طولی گاهی اوقات از فنون خوشه‌ای به‌منظور بررسی اثر مداخلات بر یک گروه از بیماران در طول زمان استفاده می‌کنند.

یک دلیل برای استفاده از کارآزمایی‌های تصادفی کنترل شده‌ی خوشه‌ای اجتناب از «آلودگی» بین بیماران دریافت‌کننده‌ی درمان و بیمارانی است که درمان را دریافت نمی‌کنند. چنین آلودگی‌هایی ممکن است تخمین اثر درمان را ضعیف کنند. برای مثال، افرادی که بخشی از یک مطالعه‌ی مربوط به مداخله‌ی رفتاری جهت کاهش مصرف سیگار یا پیشگیری از بیماری شریان کرونری هستند، ممکن است دستورالعمل‌های مشترکی دریافت کرده و اثرات مداخله را تحریف کنند.

کارآزمایی‌های کنترل شده‌ی تصادفی خوشه ای همچنین زمانی توصیه می‌شوند که تصادفی سازی در سطح بیمار عملی نبوده و یا مطلوب نیست. برای مثال، در مطالعه‌ای که هدف آن ارزیابی اثر رهنمودهای بالینی جدید است، «مداخله‌ی» تحقیقاتی در سطح مراقب یا بیمارستان، و نه سطح بیمار، انجام می‌شود.

با این حال، اجرای رهنمودهای جدید می‌تواند بر تصادفی سازی در سطح بیمار تأثیر گذاشته و در این مورد، تصادفی سازی در سطح بیمار مناسب سؤال پژوهش نیست. در عوض، رهنمودهای کاری جدید در کل بیمارستان‌ها و کلینیک‌ها اجرا شده، و هر یک از این مکان‌های مطالعه به‌صورت تصادفی به‌عنوان یک گروه کارآزمایی در نظر گرفته شده‌اند. این مثال نشان می‌دهد که چرا طرح‌های خوشه‌ای اغلب برای مداخلاتی بکار می‌روند که شامل آموزش فعالان حوزه سلامت می‌شوند.

## مسائل و مشکلات بالقوه

طراحی و تحلیل خوشه‌های کارآزمایی‌های تصادفی کنترل شده بسیار پیچیده‌تر از کارآزمایی‌های تصادفی به صورت تکی هستند. نقاط داده‌ی مربوط به بیماران موجود در یک خوشه معمولاً همبسته بوده و این همبستگی باید هم در طرح مطالعه و هم در تحلیل آن مدنظر قرار گیرد. همچنین فرایند تصادفی سازی در خوشه‌ی کارآزمایی تصادفی کنترل شده متفاوت از کارآزمایی تصادفی کنترل شده سنتی است. از آنجایی که پس از تصادفی سازی در خوشه‌ی کارآزمایی‌های تصادفی کنترل شده توافق افراد کسب می‌شود، اگر افرادی که رضایت می‌دهند با افرادی که رضایت نمی‌دهند تفاوت داشته باشند، احتمال سوگیری انتخاب وجود دارد.

## نقاط قوت

- کارآزمایی تصادفی کنترل شده خوشه‌ای اثربخشی یک مداخله در دنیای واقعی، و نه سودمندی آن را ارزیابی می‌کند. این نوع طرح می‌تواند برای پژوهش مقایسه‌ی اثربخشی مفید باشد، زیرا در این پژوهش‌ها بردرک اثرات یک مداخله در شرایط عملی دنیای واقعی تمرکز می‌شود.
- این روش همچنین یک جایگزین برای ارزیابی اثربخشی درمان‌ها در شرایطی است که تصادفی سازی در سطح فردی نامناسب یا غیرممکن است.

## محدودیت‌ها

- طرح پیچیده‌ی کارآزمایی تصادفی کنترل شده‌ی خوشه‌ای نیازمند درک و اجرای مناسب و گزارش دهی درست توسط پژوهشگر است. به منظور اطمینان از تفسیر دقیق کارآزمایی تصادفی کنترل شده‌ی خوشه‌ای توسط خوانندگان، در سال ۲۰۰۴ استانداردهای مشترک گزارش دهی کارآزمایی-ها (CONSORT) برای گزارش دادن کارآزمایی‌هایی که به صورت تکی تصادفی سازی شده بودند، گسترش یافتند تا رهنمودهای مربوط به گزارش دهی کارآزمایی تصادفی کنترل شده‌ی خوشه‌ای را پوشش دهند.

- به دلیل همبستگی درون خوشه‌ای، طرح کارآزمایی تصادفی کنترل شده‌ی خوشه‌ای در مقایسه با کارآزمایی تصادفی کنترل شده تکی نیازمند بیماران بیشتری است. به منظور رسیدن به قدرت آماری مشابه با طرح‌های تصادفی سازی تکی، باید از رویکردهای اندازه‌ی نمونه‌ی غیراستاندارد به منظور اجتناب از مطالعات کم قدرت آماری استفاده کرد (مثلاً تورم با استفاده از یک ضریب همبستگی درون خوشه‌ای محاسبه شده).

- در طرح‌های خوشه‌ای، تصمیم‌گیری مستقل بیمار ممکن است از بین برود، زیرا بیماران احتمالاً به درمان‌ها یا فرایندهایی که تنها در دیگر مکان‌ها وجود دارند، دسترسی ندارند. بنابراین، رضایت آگاهانه یک مسئله‌ی پیچیده‌ی خاص در کارآزمایی‌های تصادفی کنترل شده‌ی خوشه‌ای به حساب می‌آید. هنگامی که می‌توان اطلاعات بدست آمده از طرح خوشه‌ای را با استفاده از تصادفی سازی تکی به دست آورد، نباید از طرح خوشه‌ای استفاده کرد.

## مثال‌های انتخابی

• باس<sup>۱۷</sup> و همکاران، ژورنال انجمن پزشکی کانادا<sup>۱۸</sup> (۱۹۸۶). به منظور ارزیابی یک مداخله جهت ارتقا اثربخشی مدیریت و غربالگری فشارخون بالا در کار بالینی، پژوهشگران ۳۴ دفتر طبابت خصوصی پزشک را به صورت تصادفی به هر یک از دو گروه مداخله اختصاص دادند. هدف مطالعه تعیین اثربخشی یک سیستم مراقبتی جدید با سیستم قدیمی بود. در سیستم جدید یک دستیار پزشکی بر غربالگری نظارت کرده و کارآموزش بیماران با فشارخون بالا، همراهی آن‌ها با درمان، و جلسات را پیگیری می‌کرد. در سیستم قدیمی، بیماران صرفاً معاینه شده، داروی مناسب برای آن‌ها تجویز شده و هیچ‌گونه مدیریت یا مراقبت بعدی وجود نداشت. ۱۷ مطب پزشک که شامل ۱۵۶۵۹ بیمار بود با یک گروه کنترل متشکل از ۱۷ مطب پزشک که شامل ۱۶۴۶۵ بیمار بود، مطابقت داده شد. با توجه به ذات مداخله‌ی مورد مطالعه (یک سیستم مراقبتی جدید)، تصادفی سازی در سطح بیمار غیرعملی بوده و بنابراین از یک طرح خوشه‌ای استفاده شد.

• گروه همکاری اروپایی سازمان بهداشت جهانی<sup>۱۹</sup>. لنست<sup>۲۰</sup> (۱۹۸۶). این گروه همکاری به منظور ارزیابی اثرات یک راهبرد چندعاملی پیش‌گیری از بیماری شریان کرونری (CHD)، یک کارآزمایی طراحی کردند. این مطالعه بر روی ۸۰ کارخانه با نمونه‌ای برابر با ۶۰۸۸۱ مرد انجام شد. واحد تصادفی سازی کارخانه بود. نتایج ارزیابی شده شامل کاهش تعداد کل CHD کشنده، آنفارتوس میوکارد غیرکشنده، و تعداد کلی مرگ و میر بود. به دلیل مؤلفه‌ی آموزشی مداخله، از یک کارآزمایی تصادفی کنترل شده‌ی خوشه‌ای استفاده شد. طرح از آلودگی اجتناب کرد (مثلاً اگر تصادفی سازی در سطح بیمار انجام شده بود، بیماران شرکت‌کننده در کارآزمایی احتمالاً مواد آموزشی مشترکی با افراد غیر شرکت‌کننده داشته و باعث تحریف اثرات می‌شدند).

• زاکر<sup>۲۱</sup> و همکاران. کارآزمایی‌های بالینی کنترل شده<sup>۲۲</sup> (۱۹۹۵). پژوهشگران یک کارآزمایی سلامت با مقیاس بزرگ در جامعه برای پیشگیری از بیماری‌های قلبی عروقی در بین کودکان و بزرگسالان طراحی کردند. مداخلات در سطح مدارس انجام شد تا دانش‌آموزان یک مدرسه مداخله‌ی مشابهی را دریافت کنند. سؤال پژوهش این بود که آیا یک برنامه‌ی آموزش بهداشت قلبی عروقی مبتنی بر رفتار می‌تواند باعث ایجاد نتایج مثبتی در کودکان شود یا خیر. این کارآزمایی شامل ۹۶ مدرسه‌ی ابتدایی می‌شد که ۴۰ مورد از آن‌ها بخشی از شاخه‌ی کنترل بودند. ۲۸ مورد از آن‌ها تنها یک برنامه‌ی مداخله‌ی مدرسه محور داشته و ۲۸ مورد باقی‌مانده هم برنامه مدرسه محور و هم خانواده محور را اجرا کردند. به دلیل ذات مداخله از کارآزمایی تصادفی کنترل شده‌ی خوشه‌ای استفاده شد.

17. Bass

18. Can Med Assoc J

19. World Health Organisation European Collaborative Group

20. Lancet

21. Zucker

22. Control Clin Trials

• پلت<sup>۳۳</sup> و همکاران. ژورنال مراقبت پزشکی<sup>۳۴</sup> (۲۰۱۰). پژوهشگران یک کارآزمایی تصادفی کنترل شده‌ی خوشه‌ای را گزارش کردند که اثربخشی سه اقدام پیشگیرانه و غربالگری به منظور کاهش عفونت استافیلوکوک اورئوس مقاوم به متی‌سیلین (MRSA) در آی سی یو را مقایسه می‌کردند. واحد تصادفی سازی انتخاب شده در این کارآزمایی، بیمارستان بود. این کارروشی عملی برای کنترل آلودگی کارآزمایی در بین خوشه‌های درمان بود، زیرا اقدامات پیشگیرانه و غربالگری در سطح بیمارستانی انجام می‌شدند.

---

23. PlatMed Caret

24. Med Care

## طرح‌های با شروع تاخیری

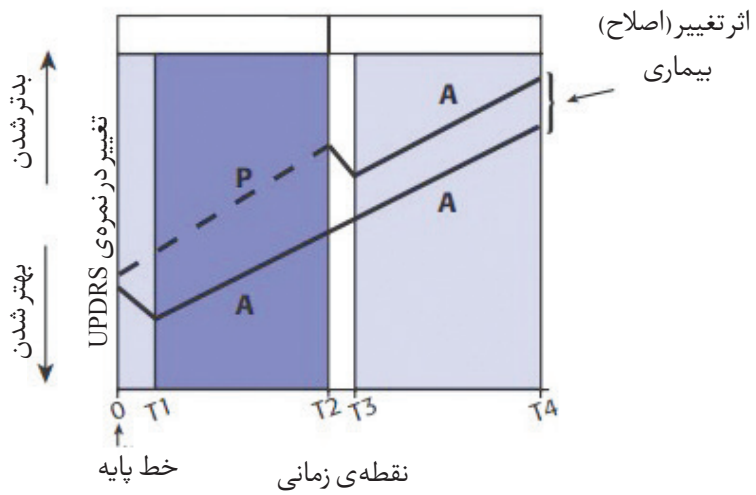
### مقدمه

استفاده از طرح‌های با شروع تاخیری این امکان را به ما می‌دهد تا مشخص کنیم اثرگذاری مداخله از طریق کاهش نشانه‌ها یا تغییر در بیماری است بخصوص در بیماری‌هایی مانند پارکینسون که به آرامی پیشروی می‌کنند.

### طرح مطالعه

یک مطالعه با شروع تاخیری در دو فاز انجام می‌شود. در فاز اول، بیماران برای دریافت درمان فعال یا یک درمان کنترل، تصادفی سازی می‌شوند. مدت زمان فاز اول باید آن قدر طولانی باشد که امکان مشاهده اثرات درمان بر نشانه‌های بیماری را به طور کامل فراهم کند. در پایان فاز اول، هرگونه تفاوت معنادار بین دو گروه (در صورت وجود) نشان دهنده اثرات نشانه‌ای، اثرات تغییر بیماری، یا هر دو است. در فاز دوم، بیماران گروه کنترل، درمان فعال را به صورت «شروع با تأخیر» دریافت می‌کنند. اکنون تمامی بیماران موجود در مطالعه درمان فعال را دریافت کرده‌اند. سپس وضعیت بیماران برای مدت زمانی که از پیش تعیین شده است، پیگیری می‌شود. هرگونه اثر نشانه‌ای مثبت باید در بین دو گروه برابر باشد. هرگونه تفاوت باقی مانده بین اثرات باید به اثر درمان در اصلاح بیماری نسبت داده شود.

شکل ۱ نمای هندسی یک طرح پژوهشی با شروع تاخیری را نشان می‌دهد. بیماران در فاز اول، یعنی ۰ تا  $T_2$  به صورت تصادفی درمان فعال یا کنترل دریافت می‌کنند. داده‌ها از بازه‌ی زمانی  $T_1$  تا  $T_2$  جمع‌آوری شده و به منظور تخمین اثر عامل بر نشانه‌ها بکار گرفته می‌شوند. همچنین، داده‌ها برای وجود اثر احتمالی عامل در اصلاح بیماری بررسی می‌شوند. تمامی بیماران در فاز دوم (از  $T_2$  تا  $T_4$ ) درمان فعال را دریافت می‌کنند. داده‌های جمع‌آوری شده از بازه‌ی زمانی  $T_3$  تا  $T_4$  به منظور تعیین اثر عامل در اصلاح بیماری استفاده می‌شوند. داده‌های مربوط به دوره‌های زمانی ۰ تا  $T_1$  و  $T_2$  تا  $T_3$  ممکن است نشان دهنده‌ی پاسخ‌های انتقالی باشند و معمولاً در تحلیل استفاده نمی‌شوند (UPDRS مقیاس درجه بندی یکپارچه‌ی بیماری پارکینسون است؛ A نشان دهنده‌ی درمان فعال و P نشان دهنده‌ی دارونما است).



شکل ۱: طرح با شروع تاخیری



## موارد استفاده‌ی توصیه‌شده

طرح‌های مطالعاتی با شروع تاخیری زمانی مفید هستند که به یک نقطه‌ی پایان بالینی برای سنجش موثق پیشروی بیماری نیاز است، نقطه‌ای که تحت تأثیر اثرات مداخله بر نشانه‌ها قرار نگرفته باشد. مزیت اصلی طرح مطالعاتی با شروع تاخیری این است که چنین طرحی به جداسازی اثرات اصلاح بیماری یک درمان از اثرات کوتاه‌مدت و مفید آن بر روی نشانه‌ها کمک می‌کند. به‌طور خلاصه، این طرح مطالعاتی از طریق اثر نشانه‌ای مداخله‌ی صورت گرفته در پژوهش، اختلاط را کنترل می‌کند. طرح‌های با شروع تاخیری همچنین زمانی مفید هستند که زمان مطالعه طولانی باشد. برای مثال، نمی‌توان انتظار داشت اثرات محافظت عصبی یک درمان در بیماری‌هایی با روند آهسته مانند پارکینسون یا آلزایمر سریع مشخص شوند. می‌توان از این طراحی برای مطالعه‌ی درمان‌های مربوط به بیماری‌هایی که روند آهسته‌ای دارند، اما نهایتاً ضعیف‌کننده و فرسایشی هستند، استفاده کرد. بیماری‌هایی مانند آلزایمر، پارکینسون، روماتیسم مفصلی، و بیماری مزمن انسدادی ریه (COPD) از جمله‌ی این بیماری‌ها هستند. همچنین، این طرح‌ها راهبرد مفیدی برای ارزیابی اجرای ضربدری (اولیه‌ی) برخی از تغییرات سیاست خاص هستند. سیاست‌هایی مانند نتایج مربوط به محدودیت‌های پوشش درمان از جمله‌ی مواردی هستند که می‌توانند با استفاده از این طرح ارزیابی شوند.

## مسائل و مشکلات بالقوه

برنامه‌ریزی و طراحی یک مطالعه با طرح شروع تاخیری نیازمند پس‌زمینه‌ی آماری قدرتمند و دانش کافی در مورد مکانیسم بیماری است. مدت‌زمان فاز اول، مدت‌زمان «حیطه‌هایی که داده‌های آن‌ها استفاده نمی‌شوند» (در شکل ۱ به صورت بازه‌های زمانی انتقال از  $T_1$  و  $T_2$  تا  $T_3$  مشخص شده‌اند)، تعداد اندازه‌گیری‌های مکرر در حین فازها، روش تحلیلی مناسب، و آزمون‌های مناسب آماری، همگی باید به‌دقت تعیین شوند.

داده‌های از دست‌رفته نیز می‌توانند یک مشکل بزرگ باشند. به دلیل عدم وجود اثر درمان، احتمال کناره‌گیری شرکت‌کنندگان گروه کنترل بیشتر است. این موضوع می‌تواند به نرخ کناره‌گیری متفاوتی بین گروه کنترل و گروه درمان منجر شود. پروتکل مربوطه باید روش‌های آماری مناسب جهت کنترل این مشکل را مشخص کند. همچنین می‌توان از تحلیل کوواریانس و نمرات گرایش به منظور بررسی عدم توازن بین گروه کنترل و درمان استفاده کرد. در تمامی موارد باید تحلیل حساسیت انجام شود.

## نقاط قوت

- طرح‌های با شروع تاخیری اثرات اصلاح بیماری را از اثرات کوتاه‌مدت آن بر نشانه‌های بیماری جدا می‌کنند.
- طرح مطالعه همچنین مسائل اخلاقی مربوط به کارآزمایی‌های تصادفی کنترل شده را نیز مدنظر

قرار می‌دهد. در مقایسه با کارآزمایی‌های سنتی، در این روش بیماران بیشتری مداخله‌ی فعال را دریافت می‌کنند. تمامی شرکت‌کنندگان در نهایت مداخله‌ای که احتمالاً از نظر پزشکی مفید و مؤثر است را دریافت می‌کنند، درحالی‌که گروه کنترل در فاز اول نگاه داشته می‌شود.

### محدودیت‌ها

- طرح‌های با شروع تاخیری نیازمند درک کافی از طرح مطالعه و پیشرفت بالینی بیمار هستند. درک این موارد به منظور تعیین دقیق مدت زمان فاز اول و دوم، و روش آماری لازم جهت تحلیل ضروری می‌باشند.
- تنها نیمه‌ی اول مطالعه دوسر کور است. نیمه‌ی دوم برچسب آشکار است. به دلیل عدم کور بودن، این محدودیت ممکن است باعث سوگیری مطالعه شود.
- طرح‌های با شروع تاخیری ممکن است با مشکلات نام‌نویسی از شرکت‌کنندگان مواجه شود. در این طرح باید بیمارانی ثبت نام شوند که در صورت قرارگیری (تصادفی) در شاخه‌ی کنترل، مایل‌اند در نیمه‌ی اول مطالعه از درمان نشانه‌ای صرف‌نظر کنند. در مورد بیماری پارکینسون، اکثر بیماران نیازمند درمان نشانه‌ای در زمان دریافت تشخیص بیماری بوده و بنابراین، گزینه‌ی مناسبی برای شرکت در مطالعه‌ی با شروع تاخیری نیستند.
- تنها بیمارانی که بیماری آن‌ها در ابتدای روند قرار داشته، خفیف بوده و سرعت پیشروی کمی دارد، ممکن است مناسب این نوع مطالعه باشند. یافته‌های مطالعه ممکن است قابل تعمیم به بیمارانی که بیماری‌شان در مراحل پیشرفته‌تر قرار دارد، نباشد.
- در فاز اول گروه دارونما، احتمال کناره‌گیری و عدم ادامه‌ی مطالعه از سوی بیماران بالا است، زیرا ممکن است بیماران هیچ‌گونه اثر درمانی را مشاهده نکنند. ویژگی‌های سطح پایه‌ی متفاوت بین بیماران فاز دوم و بیمارانی که از مطالعه کناره‌گیری کرده‌اند ممکن است باعث اختلاط آماری شده و بر نتایج تأثیر منفی بگذارند.

### مثال‌های انتخابی

• مطالعه‌ی گروهی پارکینسون<sup>۲۵</sup>، ژورنال آرکایو آو نورولوژی<sup>۲۶</sup> (۲۰۰۴). داروی TVP-۱۰۱۲ در تک درمانی زود هنگام بیماران مبتلا به پارکینسون غیربستری. در این مطالعه اثربخشی شروع زود هنگام درمان با راساگلین با شروع دیر هنگام آن مقایسه شده و تأثیر این دو بر پیشرفت بیماری و ازکارافتادگی بیمار بررسی شد. یک کارآزمایی دوسر کور گروه موازی، تصادفی شده و با شروع تاخیری اجرا شد. در این کارآزمایی ۴۰۴ بیمار که در مراحل اولیه‌ی بیماری پارکینسون قرار داشته و نیازمند درمان دوپامینرژیک نبودند، ثبت نام شدند. این مطالعه در ۳۲ مرکز درمانی در ایالات متحده و کانادا انجام شد. بیماران به صورت تصادفی در دو گروه قرار گرفتند. در گروه اول، درمان با راساگلین ۱ یا ۲ میلی گرم

25. Parkinson Study Group

26. Arch Neuro

روزانه به مدت یک سال ادامه می‌یافت. در گروه دوم، بیمار به مدت شش ماه دارونما را مصرف کرده و در شش ماه دوم ۲ میلی گرم راساگلین به صورت روزانه دریافت می‌کرد.

• هاشر<sup>۲۷</sup> و همکاران. ژورنال اختلالات حرکتی<sup>۲۸</sup> (۲۰۰۹). مطالعه‌ی گسترش برچسب باز TEMPO: در این مطالعه اثرات بالینی طولانی‌مدت درمان زود هنگام و دیر هنگام با راساگلین بر مراحل اولیه‌ی پارکینسون مقایسه شد. این مطالعه در واقع ادامه و گسترش مطالعه‌ی TVP-۱۰۱۲ (مطالعه‌ای که در بالا توصیف شد) محسوب می‌شد. بیماران به صورت تصادفی در گروه درمان زود هنگام با راساگلین (گروه شروع زود هنگام) قرار گرفته و یا ابتدا به مدت شش ماه دارونما و سپس راساگلین را دریافت کردند (گروه با شروع تاخیری). مطالعه‌ی گسترش TEMPO نتایج طولانی‌مدت گسترش برچسب باز مطالعه‌ی TEMPO را توصیف کرد، پژوهشی که طی آن بیماران به مدت ۶٫۵ سال تحت درمان با راساگلین قرار گرفتند.

• اولانو<sup>۲۹</sup> و همکاران. [چندین ناشر مختلف] (۲۰۰۸-۲۰۱۱) مطالعه‌ی کاهش سرعت پیشروی بیماری با استفاده از آزبلیکت (راساگلین) به صورت یکبار در روز (ADAGIO): این مطالعه اثرات بالقوه‌ی راساگلین در اصلاح بیماری پارکینسون را بررسی کرد. این مطالعه یک کارآزمایی دوسر کور بود که طی آن ۱۱۷۶ بیمار مبتلا به پارکینسون درمان نشده به صورت تصادفی در دو گروه قرار گرفتند. گروه اول به مدت ۷۲ هفته راساگلین (با دوز ۱ یا ۲ میلی گرم در روز) را دریافت می‌کرد (گروه با شروع زود هنگام). گروه دوم ابتدا به مدت ۳۶ هفته دارونما را مصرف کرده و سپس در ۳۶ هفته‌ی باقی‌مانده راساگلین (با دوز ۱ یا ۲ میلی گرم) مصرف نمود (گروه با شروع تاخیری). به منظور تعیین نتیجه‌ی مثبت با هر یک از دوره‌ای دارو، گروه درمان زود هنگام باید به هر یک از نقاط پایانی سلسله مراتبی در تحلیل اصلی می‌رسید. این نقاط پایانی بر اساس مقیاس یکپارچه‌ی درجه بندی پارکینسون (UPDRS) قرار دارند: برتری درمان زود هنگام نسبت به دارونما در نرخ تغییر نمرات UPDRS بین هفته‌ی ۱۲ و ۳۶. برتری درمان با شروع تاخیری در نرخ تغییر نمرات UPDRS بین خط پایه و هفته‌ی ۷۲، و عدم برتری درمان با شروع تاخیری در نرخ تغییر نمرات UPDRS بین هفته‌ی ۴۸ و ۷۲.

27. Hauser

28. Movement Disorders

29. Olanow

## روش‌های تجربی

### طرح‌های انطباقی و روش‌های بیزین

مقدمه

یک طرح انطباقی «از داده‌های جمع‌آوری شده به منظور تصمیم‌گیری در مورد چگونگی تغییر جنبه‌های کارآزمایی، هنگام ادامه‌ی آن استفاده می‌کند، بدون آنکه اعتبار و یکپارچگی مطالعه را خدشه‌دار کند.» شیوه‌های تغییر احتمالی کارآزمایی یا فرایندهای آماری عموماً از قبل تعریف می‌شوند، اما می‌توانند در حین انجام کارآزمایی نیز تغییر کنند (بدون آنکه از قبل برنامه‌ریزی شده باشند). طرح‌های انطباقی شامل نقاط میانی از پیش تعیین شده‌ای (یا همان فازها) در پروتکل هستند که پژوهشگران می‌توانند فرضیه یا فرایندهای آماری برنامه‌ریزی شده را مجدداً ارزیابی کنند. در هر یک از فازهای برنامه‌ریزی شده، تحلیل داده‌ها انجام شده و اطلاعات به‌روزرسانی شده به منظور انجام انطباق‌های خاصی بکار گرفته می‌شوند.

### دسته‌بندی‌های انطباق

انطباق‌ها در سه دسته‌ی کلی قرار می‌گیرند:

**انطباق‌های مربوط به آینده:** این انطباق‌ها قبل از شروع کارآزمایی در پروتکل مشخص شده‌اند. برخی از مثال‌های این نوع انطباق شامل تصادفی سازی انطباقی؛ متوقف ساختن پیش از موعد کارآزمایی به دلیل مسائل ایمنی، بیهودگی، یا سودمندی در تحلیل میانی؛ یا کنار گذاشتن گروه‌های درمانی ضعیف‌تر می‌شوند.

**انطباق‌های هم‌زمان:** این انطباق‌ها تغییراتی هستند که هنگام ادامه‌ی کارآزمایی صورت می‌گیرند. انطباق‌های هم‌زمان، از قبل به شکل خاص برنامه‌ریزی نشده‌اند، بلکه زمانی وارد طرح می‌شوند که به آن‌ها نیاز باشد. تغییر معیارهای ورود/خروج از مطالعه، معیارهای سنجش پذیری (تعریف کردن معنی عبارت «درمان شدن» بیمار)، رژیم دارویی (میزان دوز دارو)، یا مدت زمان دوره‌ی درمان؛ یا تغییر فرضیه‌ها یا نقاط پایان مطالعه از جمله‌ی این انطباق‌ها هستند.

**انطباق‌های مربوط به گذشته:** این انطباق‌ها عموماً در برنامه‌های تحلیل آماری انجام می‌شوند. انطباق‌های مربوط به گذشته می‌توانند پس از جمع‌آوری داده‌ها انجام شوند، اما این کار باید پیش از غیرکوری شدن عوامل درمانی انجام شود.

راهنمای سازمان غذا و داروی آمریکا در مورد طراحی کارآزمایی انطباقی توصیه می‌کند که تمامی انطباق‌ها در فاز طراحی کارآزمایی و پیش از اجرایی شدن آن مشخص شوند. این کار باعث می‌شود بتوانیم نرخ خطای نوع اول (احتمال رد شدن فرضیه‌ی صفر، هنگامی که فرضیه درست است) را به درستی محاسبه کنیم. یک برنامه‌ی قبلی این نگرانی را برطرف می‌کند که استفاده از طرح‌های انطباقی ممکن است آن قدر کارآزمایی را تغییر دهند تا از سؤال اولیه‌ی پژوهش کاملاً دور شود.

باین حال، هنگامی که انطباق‌های موردی و خاص رخ می‌دهند، بر اساس متمم‌های پروتکل پیش خواهند رفت.

### روش‌های بیزین در طرح انطباقی

اگرچه مطالعات انطباقی می‌توانند از روش‌های تحلیلی بیزین یا غیربیزین استفاده کنند، آمار بیزین به‌طور خاص مناسب مطالعات انطباقی هستند. طرح‌های انطباقی استفاده‌کننده از رویکردهای بیزین از تمامی اطلاعات موجود برای پاسخ به سؤال مدنظر استفاده می‌کنند. برعکس رویکردهای سنتی، رویکرد بیزین هم از شواهدی انباشت شده در طول زمان (مثلاً داده‌های کارآزمایی بالینی میانی) و هم اطلاعات قبلی (مثلاً منابع یا مطالعات مرتبط قبلی) استفاده می‌کند. اطلاعات قبلی با داده‌های جدید ترکیب می‌شوند تا یک «توزیع پسین» (این عبارت به معنای احتمال رویدادی است که وقوع آن به اطلاعات قبلی و داده‌های فعلی بستگی دارد) حاصل شود. سپس این توزیع تحلیل شده و به‌منظور رسیدن به نتیجه‌گیری استفاده می‌شود.

### موارد استفاده‌ی توصیه‌شده

این روش می‌تواند به‌طور خاص مفید باشد، زیرا می‌تواند شواهد با کیفیتی که جدیداً در دسترس قرار گرفته‌اند را وارد مطالعه کند. این موضوع باعث کاهش اندازه‌ی نمونه، زمان صرف شده، و هزینه‌ی لازم جهت کسب اطلاعات مرتبط برای تصمیم‌گیرندگان است. از آنجایی که با کسب اطلاعات جدید در حین کارآزمایی، این طرح امکان بروز تغییر در طرح را فراهم می‌کند، برای مطالعه‌ی بیماری‌های نادر یا پزشکی کودکان که دانش حیطه‌ی آن‌ها محدود است، مفید خواهد بود. برای مثال، هنگامی که سودمندی چندین گروه درمانی مشخص نیست، طرح‌های انطباقی (مثلاً طرح‌ها کردن بازنده) می‌توانند از داده‌های میانی استفاده کرده و به ما بگویند که چه زمانی یک شاخه‌ی درمانی ضعیف و بی‌اثر را «رها کنیم». زمان رها کردن در پروتکل از پیش تعیین شده است. سپس می‌توان بیماران را به‌صورت تصادفی در گروه‌های «برتر» باقی‌مانده قرار داد. مورد پذیرش‌ترین طرح‌های انطباقی شامل تصادفی‌سازی انطباقی پاسخ در فاز دوم کارآزمایی محصولات، تخمین مجدد اندازه‌ی نمونه به‌صورت کور، و توقف زود هنگام کارآزمایی به دلیل ایمنی، سودمندی، یا بی‌په‌ودگی مطالعه هستند.

### مسائل و مشکلات بالقوه

نباید از چندین انطباق در یک کارآزمایی استفاده کرد، زیرا این کار باعث پیچیدگی طرح شده و تحلیل تفسیر نتایج را دشوار می‌سازد. هنگام طراحی یک کارآزمایی این‌چنینی، باید از انتشار پیش از موعد اطلاعات میانی برای عموم اجتناب کرد، زیرا چنین کاری یکپارچگی کارآزمایی را به خطر خواهد انداخت. علاوه بر این، استفاده از روش‌های بیزین پیچیده بوده و نیازمند تلاش بیشتر هم در زمینه‌ی

طراحی و هم اجرای طرح بوده و به منظور جلوگیری از تفسیر نادرست نتایج به تبیین واضحی نیاز دارند.

### نقاط قوت

• طرح مطالعه‌ی انطباقی باعث انعطاف در طراحی مجدد کارآزمایی‌های بالینی در مراحل میانی می‌شوند. این انعطاف‌پذیری می‌تواند باعث کاهش قرارگیری بیمار در معرض آسیب شده و هزینه و زمان کارآزمایی را کاهش دهد.

• پژوهشگران می‌توانند فرضیات نامناسب در مورد برخی از پارامترهای برنامه‌ریزی خاص را شناسایی و اصلاح کنند. پارامترهایی مانند بزرگی اثر درمان که در مرحله‌ی طراحی کارآزمایی به خوبی درک نشده بودند.

### محدودیت‌ها

• انطباق‌ها ممکن است باعث بروز سوگیری شده و تحلیل و تفسیر را چالش‌برانگیز کنند. در صورتی که کور بودن مطالعه در حین تحلیل‌های میانی از بین برود، طرح ممکن است سوگیرانه شود. به منظور اجتناب از سوگیری، پژوهشگران باید اختصاص یافتن درمان به هرآزمودنی را کور نگه‌داشته یا از یک فرد مستقل بخواهند تا تحلیل را انجام دهد.

• بررسی میانی چندباره‌ی داده‌ها در حین کارآزمایی احتمال خطای نوع اول، یعنی مثبت کاذب، را افزایش می‌دهد. افزایش این خطا باعث کاهش قدرت آماری مقایسه‌ها می‌شود. روش‌های مربوط به کنترل نرخ خطای نوع اول باید از پیش تعیین شده و کنترل شوند.

• اقدامات آماری به منظور سنجش اطمینان نتایج مطالعه (مثلاً، PV و بازه‌های اطمینان مربوط به آن‌ها) برای اثرات درمان ممکن است قابل اعتماد نباشند. دلیل این موضوع عدم توانایی در حفظ و کنترل نرخ خطای نوع اول است.

• تدارکات مربوط به کارآزمایی‌های انطباقی پیچیده‌تر از تدارکات مربوط به کارآزمایی‌های استاندارد هستند.

### مثال‌های انتخابی

• نلسون<sup>۳۰</sup>. ژورنال موسسه‌ی ملی سرطان<sup>۳۱</sup> (۲۰۱۰). در یک کارآزمایی انطباقی فاز دوم مداخلات برای سرطان ریه‌ی غیرسلول کوچک، پژوهشگران پاسخ‌های متفاوت به عوامل ضد تومور را نسبت به مجموعه‌ای از زیرگونه‌های نشانگر زیستی تومور بررسی کردند. در فاز نخست مطالعه، پژوهشگران با استفاده از سوزن بزرگ، مثال‌های انتخابی تازه‌ای از تومور را بافت‌برداری کرده و آن‌ها را به منظور وجود چهار زیرگونه‌ی نشانگر زیستی بررسی نمودند. آن‌ها به صورت تصادفی ۴۰ درصد از جمعیت

30. Nelson

31. J Natl Cancer Inst

مطالعه (N=255) را در چهار گروه درمانی مختلفی قرار دادند. هر یک از این گروه‌ها دارویی را دریافت می‌کردند که مکانیسم اثر آن با یکی از نشانگرهای زیستی مطرح‌شده مرتبط بود. (برای مثال یکی از نشانگرهای زیستی، نشانگر ژنتیک گیرنده‌ی فاکتور رشد اپیدرمی (EGFR) بود که KRAS/BRAF نام داشت. این در حالی است که یکی دیگر از درمان‌های مورد استفاده در مطالعه ارلوتینیب بود که خود یک بازدارنده‌ی EGFR محسوب می‌شود.) در فاز دوم که بخش انطباقی کارآزمایی محسوب می‌شد، پاسخ‌های بیماران به درمان با وضعیت نشانگر زیستی تومورشان همبسته بود. این اطلاعات برای تصمیم‌گیری جهت گماشتن ۶۰ درصد باقی‌مانده‌ی بیماران در گروه‌های درمان استفاده شد. این تصمیمات درمانی در فاز سوم کارآزمایی عملی شدند. به‌طور کل، اختصاص درمان نشانگر زیستی به بیماران بر اساس نتایج حاصل از فاز انطباقی مطالعه باعث شد تا بیماری ۴۶ درصد از بیماران پس از هشت هفته به حالت پایدار برسد. تنها ۳۰ درصد از بیماران مبتلا به سرطان ریه‌ی سلول غیر کوچک که شیمی‌درمانی معمول را دریافت می‌کردند پس از هشت هفته به وضعیت پایدار رسیدند.

## طرح‌های مطالعاتی غیرتجربی

### مطالعات هم‌گروهی (کوهورت) و مورد شاهدی

#### مقدمه

مطالعات غیرتجربی طرح‌هایی هستند که در آن‌ها بیماران به صورت تصادفی یک درمان یا مداخله را دریافت نمی‌کنند. در این نوع مطالعه، بیماران و پزشکان تصمیمات درمانی را بر اساس شرایط دنیای واقعی اتخاذ کرده و سپس به مشاهده‌ی الگوهای مراقبت و نتایج می‌پردازند. یک طرح مطالعاتی غیرتجربی ممکن است برای انواع مختلفی از مطالعات مانند توصیفی، اثربخشی، مقایسه‌ی اثربخشی، یا ایمنی استفاده شود.

#### انواع مطالعات غیرتجربی

مطالعات هم‌گروهی و موردشاهدی دو نوع از طرح‌های مطالعاتی غیرتجربی هستند که می‌توان از آن‌ها برای پژوهش مقایسه‌ی اثربخشی استفاده کرد.

#### مطالعات هم‌گروهی

مطالعات هم‌گروهی یک گروه از بیماران مبتلا به یک بیماری (های) مشابه را دنبال می‌کنند (یا بیمارانی که در معرض یک درمان یا عامل مشابه قرار گرفته‌اند). مطالعات هم‌گروهی ممکن است تنها شامل بیمارانی شوند که مبتلا به بیماری مدنظر پژوهشگر (یا بیمارانی که در معرض عامل مدنظری قرار گرفته‌اند) هستند. همچنین، این مطالعات ممکن است شامل یک گروه مقایسه‌ی بیمارانی باشند که مبتلا به بیماری مدنظر نبوده و یا در برابر عامل مدنظر قرار نگرفته‌اند.

در یک مطالعه‌ی هم‌گروهی معمول جهت مقایسه‌ی اثربخشی، بیمارانی که در معرض مداخله‌ی مدنظر قرار می‌گیرند و واجد معیارهای ورود و خروج هستند، ثبت‌نام می‌شوند. بیماران برای مدتی مشخص دنبال شده، و داده‌های مربوط به نتایج مدنظر جمع‌آوری می‌شوند. برای مثال، یک مطالعه‌ی هم‌گروهی جهت مقایسه‌ی اثربخشی ممکن است بیمارانی را وارد مطالعه کند که از دو نوع داروی کاهش کلسترول استفاده می‌کنند. سپس به مدت یک سال آن‌ها را به‌منظور ارزیابی کاهش کلسترول و دیگر نتایج مدنظر مانند آنفارکتوس میوکارد حاد (AMI) دنبال می‌کنند.

#### مطالعات مورد شاهدی

مطالعات مورد شاهدی بیمارانی را وارد مطالعه می‌کنند که الف، نتیجه‌ی خاص و مدنظر پژوهشگر را تجربه کرده باشند («موارد») و ب، بیمارانی که نتیجه‌ی مدنظر پژوهشگر را تجربه نکرده‌اند، اما همانند «موارد» نماینده‌ی همان گروه جمعیتی («کنترل‌ها») هستند. برای هر یک از موارد ثبت‌نام شده در مطالعه، یک یا چند مورد کنترل نیز شناسایی و ثبت‌نام می‌شوند. برای مثال، یک مطالعه مورد شاهدی ممکن است رابطه‌ی بین داروهای ضدبارداری خوراکی و آنفارکتوس میوکارد حاد را بررسی کند. زنانی که در سن باروری هستند و دچار آنفارکتوس میوکارد حاد شده‌اند به‌عنوان موارد، و زنانی



که در سن باروری بوده و سابقه‌ی آنفارکتوس میوکارد حاد را نداشتند، به‌عنوان کنترل‌ها وارد مطالعه می‌شدند. سپس رابطه‌ی بین استفاده از قرص‌های ضدبارداری و آنفارکتوس میوکارد حاد بررسی شد. برای پژوهش اثربخشی مقایسه، می‌توان از طرح‌های مورد شاهدهی برای مقایسه‌ی فراوانی نتایج، خصوصاً نتایج نادر، در بین بیماران دریافت‌کنندهی درمان‌های مختلف استفاده کرد.

#### موارد استفاده‌ی توصیه‌شده

مطالعات غیرتجربی به‌طور خاص در شرایطی سودمند هستند که بنابه دلایل اخلاقی یا لجستیک امکان تصادفی سازی وجود ندارد. برای مثال، ممکن است در جمعیت‌های حساس (مانند زنان باردار)، بنابه دلایل اخلاقی، امکان تصادفی سازی وجود نداشته باشد. بنابه دلایل لجستیک، در مطالعاتی که نیازمند جمعیت ثبت‌نام‌کنندهی بزرگ یا داده‌های پیگیری طولانی‌مدت هستند، ممکن است امکان تصادفی سازی وجود نداشته باشد. در مقایسه با کارآزمایی‌های تصادفی، مطالعات غیرتجربی عموماً معیارهای شمول گسترده‌تری دارند و بنابراین، ممکن است بتوانند یک جمعیت بیمار نماینده‌تر (جمعیتی که بیشتر نمایندهی بیماران هستند) را وارد مطالعه کند (برای مثال بیماران مبتلا به چندین بیماری، یا بیماران مسن). مطالعات غیرتجربی همچنین می‌توانند اطلاعاتی در مورد الگوی‌های استفاده ارائه کنند. مطالعات مورد شاهدهی به‌طور خاص برای مطالعه‌ی نتایج نادر مفید هستند. در این حالت، ثبت‌نام از تعداد نمونه‌ی کافی در یک مطالعه‌ی هم‌گروهی دشوار است. در مثال مربوط به قرص‌های ضدبارداری مطرح‌شده در بالا، نتیجه‌ی مدنظر (آنفارکتوس میوکارد حاد) در جمعیت مدنظر (زنانی که در سن باروری هستند) نسبتاً نادر است. بنابراین، یک مطالعه هم‌گروهی باید نسبتاً بزرگ باشد تا بتواند بیماران لازم برای بررسی این سؤال را داشته باشد. این محدودیت باعث می‌شود تا مطالعه‌ی مورد شاهدهی یک گزینه‌ی جذاب به نظر برسد.

داده‌های لازم برای یک مطالعه‌ی غیرتجربی ممکن است به‌صورت گذشته‌نگر یا آینده‌نگر جمع‌آوری شوند. در مطالعات که داده‌ها به‌صورت آینده‌نگر جمع‌آوری می‌شوند، مواجهه‌ها (مثلاً مواجهه با درمان مدنظر) قبل از بروز نتایج مدنظر سنجیده می‌شوند. برای مثال، در مطالعه‌ی کاهش کلسترول مطرح‌شده در بالا، بیمارانی که داروی مدنظر پژوهش را مصرف می‌کنند وارد مطالعه شده و تحت بررسی قرار می‌گیرند تا مشخص شود آیا نتیجه‌ی مدنظر (مثلاً آنفارکتوس میوکارد حاد) در آن‌ها بروز می‌کند یا خیر. در مطالعه‌هایی که جمع‌آوری داده‌ها به‌صورت گذشته‌نگر است، داده‌ها از منابع داده‌ی موجود (مثلاً پایگاه‌های داده‌ی مدیریتی، سوابق پزشکی و...) گردآوری می‌شوند. در نتیجه، مواجهه‌ها پس از بروز نتیجه‌ی مدنظر اندازه‌گیری می‌شوند. در مطالعه‌ی کاهش کلسترول مطرح‌شده در بالا، تمامی بیماران مصرف‌کنندهی داروی مدنظر در یک پایگاه داده‌ی از پیش موجود شناسایی می‌شوند. سپس داده‌های مربوط به این بیماران به‌منظور بررسی وجود هم‌زمان نتیجه‌ی مدنظر بررسی می‌شوند.

## مسائل و مشکلات بالقوه

مشکل اصلی مطالعات غیرتجربی، پتانسیل آن‌ها برای مخدوش شدن است. به دلیل عدم قرار گرفتن بیماران در گروه‌های درمانی به صورت تصادفی، امکان عدم توزیع یکسان برخی از ویژگی‌های بیماران بین دو گروه وجود دارد. مخدوش شدن زمانی رخ می‌دهد که این ویژگی‌ها با نتیجه‌ی مدنظر پژوهش رابطه داشته باشند. در چنین مواردی، این ویژگی‌ها (ویژگی‌های بیماران)، متغیرهای مخدوش کننده نامیده می‌شوند. تحلیل پژوهش باید این متغیرها را نیز شامل شود. برای مثال، مطالعه‌ای که به مقایسه‌ی اثربخشی دو داروی کاهش کلسترول می‌پردازد، باید اطلاعات مربوط به رژیم غذایی بیماران را نیز جمع‌آوری کرده و مدنظر قرار دهد. امکان بروز مخدوش شدن از طریق انتخاب درمان در طرح‌های غیرتجربی وجود دارد. در این نوع مخدوش شدن، یکی از ویژگی‌های بیمار که با نتیجه‌ی مدنظر رابطه دارد، بر انتخاب درمان تأثیر می‌گذارد. برای مثال، یک مطالعه ممکن است دو محصول (داروی) آسم را با هم مقایسه کند. یک محصول جدید و یک محصول قدیمی. بیماران مبتلا به آسم شدیدتر که به داروی قدیمی پاسخ مناسبی نمی‌دهند، ممکن است برای دریافت داروی جدید در اولویت باشند. این موضوع باعث به وجود آمدن گروه‌های درمانی نابرابر می‌شود. علاوه بر این، در طرح‌های غیرتجربی، هنگام وجود یک عامل ناشناس و اندازه‌گیری نشده در مطالعه که بر نتایج اثر می‌گذارد، احتمال بروز مخدوش شدن اندازه‌گیری نشده وجود دارد. برای مثال، در مطالعه‌ی مربوط به داروهای کاهش کلسترول، ممکن است آنفارتوس میوکارد حاد یکی از نتایج مدنظر پژوهش باشد. دیابت یکی از عوامل خطر برای آنفارتوس میوکارد حاد است که با استفاده از داروهای کاهش کلسترول ارتباط دارد. در این شرایط، اگر مطالعه اطلاعات مربوط به دیابت، به عنوان یک عامل خطر، را جمع‌آوری نکند، احتمال وجود سوگیری در نتایج مطالعه وجود دارد. می‌توان از چندین روش تحلیل آماری برای رفع مشکل مخدوش شدن بالقوه استفاده کرد.

مطالعات غیرتجربی آینده‌نگر می‌توانند داده‌های خاص مدنظر سؤالات پژوهش را تعریف و جمع‌آوری کنند. با این حال، مطالعات آینده‌نگر نیازمند زمان برای جمع‌آوری داده‌ها بوده و نسبت به مطالعات گذشته‌نگر هزینه‌ی بیشتری دارند. در طرف مقابل، مطالعات گذشته‌نگر نسبتاً سریع به پایان می‌رسند و معمولاً مقرون به صرفه هستند. با این وجود، مطالعات گذشته‌نگر محدود به داده‌های موجود هستند. عناصر داده‌ای که در منابع داده‌ی قبلی وجود نداشته باشند برای تحلیل موجود نبوده و اطلاعات لازم در مورد چگونگی تعریف عناصر داده نیز ممکن است وجود نداشته باشد.

## نقاط قوت

- در شرایطی که استفاده از تصادفی سازی امکان‌پذیر یا اخلاقی نیست، می‌توان از طرح‌های هم‌گروهی و مورد شاهدی استفاده کرد.
- معیارهای شمول (ورود) گسترده و معیارهای عدم شمول (خروج) کم باعث ایجاد جمعیتی

بالینی می‌شوند که بیشتر نماینده‌ی جمعیت هدف بوده و بنابراین نتایج مطالعه تعمیم‌پذیری بیشتری دارد.

• مطالعات گذشته‌نگر در مقایسه با دیگر انواع مطالعات نسبتاً سریع‌تر انجام شده و مقرون به صرفه هستند.

• مطالعات مورد شاهدهی برای بررسی نتایج نادر مفید هستند.

### محدودیت‌ها

• در این نوع مطالعه، مخدوش شدن مشکل بزرگی محسوب می‌شود. می‌توان با استفاده از روشهای آماری مخدوش شدن را کنترل کرد، اما مخدوش کننده اندازه‌گیری نشده همیشه یکی از نگرانی‌های مطالعات هم‌گروهی محسوب می‌شود.

• در مطالعات گذشته‌نگر، برخی از عناصر داده‌ی مدنظر ممکن است در دسترس نباشند.

### مثال‌های انتخابی

• کرلی کاوسکی<sup>۳۲</sup> و همکاران. سالنامه‌ی پزشکی داخلی<sup>۳۳</sup> (۲۰۱۱). پژوهشگران به منظور مقایسه‌ی اثربخشی ماموگرافی دیجیتال با آنالوگ در جامعه‌ی آمریکا (درمانگاه‌ها و مطب‌های شخصی)، از یک مطالعه هم‌گروهی گذشته‌نگر استفاده کردند. آن‌ها ۳۲۹,۲۶۱ زن بین ۴۰ تا ۷۹ سال که تحت ماموگرافی دیجیتال یا آنالوگ قرار گرفته بودند را وارد مطالعه کردند. نتایج مدنظر اصلی پژوهشگران شامل حساسیت ماموگرافی، اختصاصی بودن آن، نرخ تشخیص سرطان، و نتایج مربوط به تومور بود. از یک مطالعه هم‌گروهی برای ارزیابی و مقایسه‌ی اثربخشی چندین نتیجه‌ی مختلف استفاده شد، کاری که با استفاده از مطالعات مورد شاهدهی قابل انجام نیست.

• مساروی<sup>۳۴</sup> و همکاران، سالنامه‌ی جراحی آنکولوژی<sup>۳۵</sup> (۲۰۱۱) پژوهشگران برای بررسی ایمنی و اثربخشی حذف با بسامد رادیویی (RFA) در درمان کارسینومای هیپاتوسلولار (سرطان کبد)، از یک مطالعه هم‌گروهی گذشته‌نگر استفاده کردند. این مطالعه از پایگاه داده‌ی همه‌گیرشناسی، پایش، و نتایج نهایی (به اختصار SEER) که به مدیکر دیتا (Medicare Data) متصل بود، به منظور مقایسه‌ی بیماران دریافت‌کننده‌ی درمان حذف با بسامد رادیویی با بیماران جراحی شده، یا بیماران بدون درمان استفاده کرد. نتایج مدنظر مطالعه شامل نرخ مرگ‌ومیر و مراجعه‌ی مجدد به بیمارستان بود، مواردی که احتمالاً نیازمند یک دوره‌ی پیگیری طولانی‌مدت، خارج از یک مطالعه‌ی گذشته‌نگر بودند.

• مطالعه‌ی گروهی سازمان جهانی بهداشت. لنست (۱۹۹۷). به منظور بررسی رابطه‌ی بین

32. Kerlikowske

33. Ann Intern Med

34. Massarweh

35. Ann Surg Oncol

قرص‌های ضدبارداری خوراکی و آنفارکتوس میوکارد حاد (که رویداد نادری است)، از یک مطالعه‌ی موردشاهدی استفاده شد. زنان بین ۲۰ تا ۴۴ ساله که به دلیل آنفارکتوس میوکارد حاد حتمی یا احتمالی بستری شده بودند، به‌عنوان مورد وارد مطالعه شدند. هر مورد به حداکثر سه مورد کنترل جور شد. این کنترل‌ها بنا به دلیلی جز آنفارکتوس میوکارد حاد بستری شده بودند. به‌منظور مقایسه‌ی خطر آنفارکتوس میوکارد حاد در بین مصرف‌کنندگان فعلی قرص‌های ضدبارداری خوراکی و زنانی که این قرص‌ها را مصرف نکرده بودند، از شرکت‌کنندگان هنگام حضور در بیمارستان مصاحبه به عمل آمد و نسبت‌های شانس (Odds Ratios) نیز محاسبه شد.

## روش‌های غیرتجربی

### طراحی‌های کاربر جدید

#### مقدمه

طرح کاربر جدید نوعی مطالعه‌ی غیرتجربی است که تمامی بیماران موجود در یک جمعیت تعریف‌شده را هنگام شروع درمان شناسایی می‌کند. این طرح از طریق خارج کردن کاربران متداول، سوگیری‌های خاصی را از بین برده، و تنها بیمارانی را در مطالعه نگه می‌دارد که قبل از درمان حداقل استفاده از درمان را داشته، یا فاقد سابقه‌ی درمان باشند.

### طرح مطالعه و روش‌های آن

#### کاربران جدید:

با توجه به دشواری‌های مربوط به شناسایی کاربران جدید، خصوصاً در مطالعات همه‌گیرشناسی دارویی (فارماکوپیدمیولوژی)، در صورتی که بیماران قبل از شروع درمان برای یک دوره‌ی مشخص (دوره‌ی عدم استفاده از درمان) از دارو یا درمان استفاده نکرده باشند، به‌عنوان کاربران «جدید» شناسایی و وارد مطالعه می‌شوند. با افزایش مدت زمان دوره‌ی بدون درمان قبل از شروع پیگیری، احتمال اینکه بیمار وارد شده به مطالعه‌ی کاربر واقعاً جدید باشد، بیشتر خواهد شد.

#### گروه‌های مقایسه:

انتخاب یک گروه مقایسه می‌تواند موضوعی پیچیده و ذهنی باشد. مقایسه‌ی ایده‌ال در شرایطی صورت می‌گیرد که توزیع عوامل خطر اندازه‌گیری شده و نشده برای نتیجه‌ی مطالعه، در بین بیماران یکسان باشد. طرح‌های کاربر-جدید را می‌توان به‌صورت یک مطالعه‌ی هم‌گروهی یا مطالعه‌ی مورد شاهدی لانه‌گزینی شده (یک مطالعه‌ی مورد شاهده‌ی که در داخل یک هم‌گروهی موجود انجام می‌شود) انجام داد. در این طرح‌ها، گروه درمان به‌صورت ایده‌ال متشکل از کاربرانی است که یک درمان جایگزین را دریافت می‌کنند، یا هیچ درمانی را دریافت نمی‌کنند.

#### کاربران درمان جایگزین:

بیمارانی که از یک درمان جایگزین استفاده می‌کنند و تشخیص پزشکی یکسانی دارند (نشانه‌های بیماری‌شان یکسان است).

#### کاربرانی که درمانی استفاده نمی‌کنند:

بیمارانی که از یک درمان جایگزین استفاده نمی‌کنند. می‌توان آن‌ها را شناسایی کرد و یک تاریخ تصادفی برای شروع پیگیری به آن‌ها اختصاص داد.

### موارد استفاده‌ی توصیه‌شده

زمانی باید از طرح‌های کاربر جدید استفاده کرد که نرخ بروز نتایج (تغییرات) مرتبط با درمان، براساس

زمان گذشته از شروع درمان متفاوت است. اغلب، دوره‌ی استفاده از دارو در ابتدای درمان با افزایش یا کاهش خطر بروز نتیجه رابطه دارد. این یعنی در صورت وارد کردن کاربران رایج دارو به مطالعه، مطالعه سودار شده و خطر بروز نتیجه با گذشت زمان از شروع درمان تغییر می‌کند. مطالعات انجام شده در دهه‌ی ۹۰ میلادی را در نظر بگیرید که گزارش داده بودند خطر ترومبوآمبولیسم سیاهرگی در کاربران قرص‌های ضدبارداری نسل سوم در مقایسه با کاربران نسل‌های قبلی این داروها افزایش می‌یابند. از آنجایی که خطر ترومبوآمبولیسم رگی در ابتدای درمان بالاتر بود و کاربران قرص‌های ضدبارداری نسل سوم در مقایسه با کاربران نسل‌های قدیمی‌تر این داروها، به‌تازگی شروع به مصرف آن کرده بودند، احتمال دارد که افزایش خطر گزارش شده در مورد این داروها (در مقایسه با داروهای نسل قبل) را بتوان به این تفاوت در مدت درمان بین گروه‌های مقایسه نسبت داد.

اگر کاربران داروهای قدیمی‌تر از زمان شروع درمان تحت نظارت قرار می‌گرفتند، احتمال داشت نتایج تغییر کند. طرح‌های کاربر جدید این‌گونه سوگیری‌های عدم اطمینان را برطرف می‌کنند، زیرا برای هر بیمار، تحلیل با شروع دوره‌ی فعلی درمان آغاز می‌شود.

همچنین زمانی که مداخله باعث تغییر احتمالی عوامل خطر بیماری می‌شود، می‌توان از یک طرح کاربر جدید استفاده کرد. اگر مطالعه شامل کاربران متداول [یک دارو یا درمان] باشد، کنترل آن دست از عوامل بیماری که احتمالاً توسط مداخله تغییر یافته‌اند، چالش برانگیز خواهد بود. در یک طرح کاربر جدید می‌توان به‌منظور حذف اثر عوامل خطر بیماری بر درمان، قبل از شروع درمان آن‌ها را اندازه‌گیری کرد. این کار امکان تنظیم متغیرهای مخدوش کننده قبل از شروع درمان را فراهم کرده و یک توالی زمانی مشخص برای ارزیابی و تعدیل متغیرهای مخدوش کننده را قبل از شروع درمان تعیین می‌کند.

### مسائل و مشکلات بالقوه

دشواری لجستیک (تدارکاتی) شناسایی زمان شروع درمان و جمع‌آوری اطلاعات در مورد متغیرهای مخدوش کننده بالقوه قبل از شروع درمان، یکی از محدودیت‌های اصلی طرح‌های کاربر جدید است. استفاده از پایگاه‌های داده‌ی خودکار بزرگ و پیوند سوابق امکان شناسایی مؤثر کاربران جدید را فراهم کرده و اطلاعات دقیقی را در اختیار پژوهشگر قرار می‌دهد. می‌توان از این اطلاعات برای تعریف متغیرهای مخدوش کننده بالقوه و استفاده از دارو استفاده کرد.

خارج کردن کاربران متداول از یک طرح مطالعه، اندازه‌ی نمونه و بالطبع آن قدرت مطالعه را کاهش می‌دهد. در صورتی که پایگاه داده‌ی خودکار دارای داده‌های طولی مناسب در مورد تاریخچه‌ی بیماران باشد، قدرت مطالعه کاهش نخواهد یافت، زیرا می‌توان زمان اولین استفاده‌ی هر کاربر [از دارو یا درمان] را شناسایی کرد. در صورتی که تاریخچه‌ی مناسب بیماران در دسترس نباشد، پژوهشگران می‌توانند بزرگی سوگیری‌های بالقوه‌ی مربوط به ورود کاربران متداول به مطالعه را ارزیابی کنند. اگر نشانه‌ای مبتنی بر سوگیری یافت نشود، می‌توان داده‌های مربوط به کاربران متداول را وارد تحلیل کرد.

## نقاط قوت

- طرح مطالعاتی کاربر جدید از طریق خارج کردن کاربران متداول از مطالعه، عدم اطمینان از رخدادهای اولیه (داروها یا درمان‌هایی که بیماران قبلاً مصرف کرده‌اند) را از بین می‌برد.
- این‌گونه طرح‌ها عوامل خطر بیماری را به‌عنوان متغیرهای مخدوش‌کننده در نظر می‌گیرند.

## محدودیت‌ها

- تعیین زمان شروع درمان ممکن است دشوار باشد.
- طرح کاربر جدید به دلیل خارج کردن کاربران متداول از مطالعه، اغلب اندازه‌ی نمونه را کاهش می‌دهد.

## نمونه‌های انتخابی

• اشنی و ایس<sup>۳۶</sup> و همکاران، ژورنال مراقبت پزشکی<sup>۳۷</sup> (۲۰۰۷). در این پژوهش که بر روی افراد تحت پوشش بیمه‌های مدیکرو مدیکید انجام شد، پژوهشگران نرخ مرگ‌ومیر در بین مصرف‌کنندگان داروی استاتین در طی یک سال را بررسی کردند. میانگین مدت مصرف دارو ممکن بود اثرات مربوط به شروع درمان را کم‌تر، و اثرات مربوط به استفاده‌ی طولانی‌مدت از درمان را بیشتر نشان دهد. به همین دلیل، یکی از اولین محدودیت‌های مطالعه‌ی هم‌گروهی این بود که فقط از مصرف‌کنندگان جدید یا تصادفی استاتین در مطالعه استفاده شود.

• ری، دلبیوای<sup>۳۸</sup> و همکاران. لنست (۲۰۰۲). این مطالعه از داده‌های گذشته‌نگر چندین مکان استفاده کرد. داده‌ها از برنامه‌ی بیمه‌ی مدیکید تنسی در کانادا و بریتانیا جمع‌آوری شدند. هدف این مطالعه بررسی خطر آنفراکتوس میوکارد و بیماری کرونر قلبی کشنده در بیمارانی بود که به‌صورت عمومی برایشان داروهای ضدالتهاب غیراستروئیدی تجویز شده بود. جامعه‌ی مدنظر، بیمارانی بودند که اخیراً به دلیل بیماری کرونر قلبی بستری شده بودند. ایمنی قلبی عروقی داروهای ضدالتهاب غیراستروئیدی بسیار بحث‌برانگیز بوده و خطر بروز عواقب قلبی عروقی در آینده با سابقه‌ی ضعیف همراه بوده و احتمالاً باعث بروز مخدوش‌کنندگی در نتایج می‌شود. یک کوهورت (هم‌گروهی) از کاربران جدید داروهای ضدالتهاب غیراستروئیدی براساس سن، جنسیت، و تاریخ شروع استفاده از دارو با گروهی دیگر از افرادی که این داروها را مصرف نمی‌کردند، جور شدند. نقطه‌ی پایان مطالعه بستری شدن در بیمارستان به دلیل آنفراکتوس میوکارد حاد یا مرگ در اثر بیماری کرونر قلبی بود.

• ری، دلبیوای، و همکاران، ژورنال پیشگیری از آسیب<sup>۳۹</sup> (۲۰۰۲). این مطالعه از داده‌های برنامه‌ی بیمه‌ی مدیکد تنسی استفاده کرده و به‌صورت گذشته‌نگر خطر شکستگی سر استخوان ران در بین مصرف‌کنندگان استاتین را با خطر شکستگی در بین مصرف‌کنندگان دیگر داروهای کاهش کلسترول

36. Schneeweiss

37. Medical Care

38. Ray WA

39. Inj Prev

مقایسه کرد. مصرف‌کنندگان جدید تمامی داروهای کاهش چربی و افراد غیر مصرف‌کننده‌ی گروه کنترل، همگی حداقل ۵۰ سال سن داشته، مبتلا به بیماری کشنده نبوده، در آسایشگاه سکونت نداشته، یا دچار زوال عقل یا پوکی استخوان نبودند. معیار خروجی اصلی مطالعه، شکستگی سر استخوان ران بود. شکستگی‌های پاتولوژیک یا شکستگی‌های ناشی از ضربه‌ی شدید در نظر گرفته نمی‌شدند. پژوهشگران تنها از مصرف‌کنندگان جدید داروهای کاهش چربی استفاده می‌کرد تا مطمئن شوند تمامی شکستگی‌های کشف‌شده در طول مطالعه پس از مصرف داروی کاهش چربی رخ داده‌اند. اگر از یک مطالعه‌ی هم‌گروهی مرسوم استفاده می‌شد، رخدادهای (شکستگی‌ها) زود هنگامی که احتمالاً پس از شروع درمان به وقوع می‌پیوستند، احتمالاً نادیده گرفته می‌شدند.



محدود سازی روشی است که به منظور کاهش سوگیری بکار می‌رود. این روش از طریق همانندسازی گروه‌های مقایسه تا حد ممکن، این کار را انجام می‌دهد. هدف این است که گروه‌ها از نظر وجود عوامل مخدوش کننده بالقوه و اثرات درمان شبیه به هم شوند. بیماران به سطوح تعریف شده از ویژگی‌های فردی قابل اندازه‌گیری (مثلاً جنسیت، گروه سنی، خطر بیماری، توصیه‌ی درمان) «محدود» می‌شوند. اگر از یک ویژگی به عنوان یک متغیر محدود سازی استفاده شود، آن ویژگی دیگر نمی‌تواند به عنوان متغیر مخدوش کننده ایفای نقش کند.

### موارد استفاده‌ی توصیه شده

در دنیای واقعی، درمان‌ها به صورت تصادفی استفاده نمی‌شوند، بلکه بر اساس نشانه‌های خاص و برای جمعیت‌های خاص بکار گرفته می‌شوند. وضعیت بیمار تعیین‌کننده‌ی شروع درمان، و نوع درمان انتخابی است. هنگامی که هم‌گروه‌های بیمار نشانه‌های یکسانی نداشته باشند، درمان می‌تواند با شدت و مدت بیماری همبسته باشد، موضوعی که باعث مخدوش شدن از طریق انتخاب درمان می‌شود. مخدوش شدن از طریق انتخاب درمان، با افزایش، کاهش، یا معکوس کردن مصنوعی رابطه‌ی تخمینی، اعتبار درونی را تضعیف می‌کند. برای از بین بردن مخدوش شدن از طریق انتخاب درمان، می‌توان از محدود سازی استفاده کرد.

مثالی را در نظر بگیرید که دو دسته‌ی داروهای کاهنده فشارخون، یعنی بتا بلاکرها و ادرارآورها را از نظر تأثیر بر آنفارکتوس میوکارد با هم مقایسه می‌کند. از آنجایی که بتا بلاکرها برای درمان آنژین صدری نیز بکار گرفته می‌شوند (آنژین صدری یکی از نشانه‌ها برای آنفارکتوس میوکارد به شمار می‌رود)، این مقایسه در معرض سوگیری مخدوش شدن از طریق انتخاب درمان قرار دارد. محدود سازی تحلیل در این مثال به بیمارانی که دارای سابقه‌ی بالینی بیماری قلبی عروقی نیستند باعث هم‌گروه‌هایی از بیماران می‌شود که پیش‌آگهی‌شان تحت تأثیر علائم بالینی بیماری قلبی عروقی قرار ندارد.

محدود سازی هم‌گروه‌های مطالعه به بیمارانی که از نظر نشانه‌های لازم برای مطالعه همگن هستند، باعث بهبود توازن پیش‌بین‌های نتیجه‌ی بیماران در گروه مواجهه (درمان) شده و بنابراین، مخدوش شدن را کاهش می‌دهد. علاوه بر این، محدود سازی هم‌گروه‌های مطالعه احتمال پاسخ مشابه تمامی بیماران مطالعه به درمان را افزایش داده و بنابراین، احتمال تغییر اثر را کاهش می‌دهد.

### مسائل و مشکلات بالقوه

تعمیم‌پذیری نتایج مطالعه می‌تواند بر اساس نوع محدودیت صورت گرفته در تحلیل متفاوت باشد. برای مثال، اگر فقط مصرف‌کنندگان جدید و غیر مصرف‌کنندگان را وارد هم‌گروهی کنیم، نقش اثرات درمانی صورت گرفته پس از مدت کوتاهی از شروع درمان دست‌کم گرفته نشده، و در نتیجه

تعمیم‌پذیری محدود نمی‌شود.

با این حال، زمانی که مطالعه زیرگروه‌های پرخطر یا کم‌خطر را بر اساس شدت بیماری، یا بیماری‌های همپوش محدود می‌کند، تعمیم‌پذیری نتایج آن خدشه می‌شود، زیرا جمعیت بیمارانی که پزشک می‌تواند نتایج را برای آن‌ها اعمال کند، محدود خواهد شد. همچنین، پژوهشگران باید متغیرهای انتخاب شده برای محدود سازی جمعیت مطالعه را به دقت بررسی کنند تا اطمینان یابند که محدودیت‌های وضع شده سهواً باعث سوگیری نمی‌شوند.

### نقاط قوت

- محدود سازی باعث رفع مخدوش شدن از طریق انتخاب درمان می‌شود.
- می‌توان از این روش برای متوازن کردن ویژگی‌های بیماران استفاده کرد.

### محدودیت‌ها

- بر اساس معیارهای انتخاب شده برای محدود سازی جمعیت بیماران، تعمیم‌پذیری نتایج ممکن است محدود شود.

### مثال‌های انتخابی

• اشنی وایس و همکاران، ژورنال مراقبت پزشکی (۲۰۰۷). در این پژوهش که بر روی افراد تحت پوشش بیمه‌های مدیکرومدیکید انجام شد، پژوهشگران نرخ مرگ‌ومیر در بین مصرف‌کنندگان داروی استاتین در طی یک سال را بررسی کردند. پژوهشگران به منظور کاهش سوگیری مطالعه، سطوح فزاینده‌ای از محدودیت را اعمال کردند: ۱. فقط مصرف‌کنندگان تازه را وارد مطالعه کردند (افرادی که در ۱۲ ماه قبل از تاریخ شاخص استاتین مصرف نکرده بودند)؛ ۲. انتخاب یک گروه مقایسه که بیش از همه به گروه مداخله شباهت داشت (با محدود سازی گروه مقایسه به مصرف‌کنندگان تازه یک درمان پیشگیر دیگر)؛ ۳ حذف کردن بیمارانی که دچار موارد منع ورود به مطالعه هستند (از طریق تخمین زدن مدل نمره گرایش و اصلاح نخستین صدک توزیع نمرات گرایش)؛ ۴. حذف بیمارانی که پایبندی پایینی به درمان دارند (بیمارانی که طی ۱۸۰ روز پس از شروع درمان، داروهای تجویز شده برای بار دوم و سوم را مصرف نمی‌کنند)؛ و ۵. سنجش زیرگروه‌ها برای وجود خطر یافت شده در کارآزمایی‌ها (از طریق محدود سازی ارزیابی به بیمارانی که صلاحیت شرکت در کارآزمایی PROSPER را داشته‌اند). نرخ خطر تعدیل شده‌ی اولیه‌ی جمعیت برای استفاده از استاتین برابر با ۰٫۶۲٪ بود، اما پس از اعمال چهار محدودیت به ۰٫۸۰٪ رسید. محدودیت سطح ۵ که بیماران را محدود به افرادی که می‌کرد که

شایسته‌ی شرکت در کارآزمایی PROSPER بودند، نرخ خطر تعدیل شده‌ای برابر با ۰٫۷۹ داشت. این کار پژوهشگران را قادر ساخت تا ارزش محدودیت برای تعدیل اختلاط در مطالعات غیرکارآزمایی را نشان دهند. محدود سازی در چهار سطح اول نرخ خطری همانند نرخ خطر مشاهده شده در جمعیت کارآزمایی بالینی داشت.

## تحلیل زیرگروه

### مقدمه

مطالعات غیرکارآزمایی شامل دامنه‌ی گسترده‌ای از بیماران می‌شوند که طی کار بالینی معمول درمان را دریافت می‌کنند. به دلیل تفاوت احتمالی بیماران یک جمعیت با دیگری (یا به عبارتی ناهمگن بودن آن‌ها)، پاسخ آن‌ها به درمان نیز احتمالاً متفاوت خواهد بود. تحلیل زیرگروه روشی رایج برای ارزیابی این موضوع است که آیا اثرات درمان بین زیرگروه‌های تعریف شده از بیماران در یک مطالعه‌ی غیرکارآزمایی متفاوت است یا خیر.

### موارد استفاده‌ی توصیه شده

تحلیل زیرگروه باید در مواردی انجام شود که احتمال می‌رود اثرات درمان بین زیرمجموعه‌های بیماران یک مطالعه متفاوت باشد. به منظور تعیین متفاوت بودن اثر درمان بر زیرگروه‌ها، ابتدا یک آزمون آماری انجام می‌شود (آزمونی برای «برهم‌کنش»). این آزمون تعیین می‌کند که آیا اثرات درمان در سطوح مختلف یک ویژگی بیمار (مثلاً سن، گروه‌ها) از نظر آماری معنادار است یا خیر. در صورت وجود تفاوتی که از نظر آماری معنادار باشد، آزمون‌های بیشتری انجام می‌شود تا مشخص گردد کدام سطوح پاسخ (مثلاً کدام گروه‌های سنی) اثرات درمانی متفاوتی را تجربه می‌کنند.

### مسائل و مشکلات بالقوه

بسیاری از مطالعات غیرکارآزمایی به منظور ارزیابی تفاوت در اثربخشی درمان بین دو گروه درمانی اصلی طراحی می‌شوند، به نحوی که در ابتدا و هنگام طراحی، تحلیل زیرگروه مدنظر نیست. در نتیجه، بسیاری از مطالعات تنها دارای قدرت آماری لازم برای کشف تفاوت‌های کلی اثر اصلی در بین تمامی گروه‌های مطالعه هستند. تحلیل‌های زیرگروه ممکن است قادر نباشند تفاوت‌های معنادار آماری را در یک یا چند زیرگروه کشف کنند، زمانی که چنین تفاوت‌هایی واقعاً وجود دارند. اگر اثر زیرگروه واقعاً وجود داشته باشد، صرفاً به دلیل اینکه مطالعه به اندازه‌ی کافی بزرگ نیست، ممکن است کشف نشود. با توجه به تعداد بالای متغیرهای خط پایه، بسیاری از تحلیل‌های زیرگروه در مطالعات غیرکارآزمایی، به جای انجام شدن بر اساس یک برنامه‌ی تحلیل زیرگروه از قبل تعریف شده، ناشی از اکتشاف داده‌های Post hoc (تعمیقی) هستند. این تحلیل‌ها باید اکتشافی بوده و تنها در شرایط استثنایی بر نتیجه‌گیری‌های حاصل از مطالعه تأثیر بگذارند.

## قدرت

• تحلیل زیرگروه به ارزیابی این موضوع می‌پردازد که آیا زیرگروه‌های بیماران سود یا زیان متفاوتی از یک درمان را تجربه می‌کنند یا خیر. آگاهی از این موضوع به تصمیم‌گیری بالینی کمک می‌کند.

### محدودیت‌ها

• تحلیل زیرگروه چندگانه به‌طور رایجی انجام می‌شود. هنگام انجام تحلیل زیرگروه چندگانه (استفاده از چندین آزمون برهم‌کنش)، احتمال مشاهده‌ی مثبت کاذب (یافتن یک برهم‌کنش معنادار، درحالی‌که واقعاً وجود ندارد) افزایش می‌یابد. این موضوع به نتیجه‌گیری نادرستی ختم می‌شود که بنابراین اثر درمان در بین زیرگروه‌ها متفاوت است، درحالی‌که واقعاً چنین نیست.

### مثال‌های انتخابی

• ماراسکینی<sup>۴۰</sup> و همکاران، ژورنال تعاملی جراحی قلبی عروقی و صدی<sup>۴۱</sup> (۲۰۱۰). داده‌های قبلاً گزارش شده حاکی از این بودند که هم جنسیت و هم سن ممکن است در مرگ‌ومیر ۳۰ روز بستری در بیمارستان نقش داشته باشند. با این حال، بسیاری از مطالعات قدرت آماری لازم را نداشته، یا داده‌های متناقضی را گزارش کرده بودند. این مطالعه‌ی غیرکارآزمایی اثر جنس و سن بر مرگ‌ومیر طی ۳۰ روز پس از جراحی شریان کرونری در ۷۴۵۷۷ بیمار را بررسی کرد. این بیماران بر اساس جنس و سن طبقه‌بندی شدند تا اثر تغییر هر دو متغیر مورد آزمون قرار گیرد. نتایج مطالعه حاکی از این بود که بیماران مؤنث و مسن نرخ مرگ‌ومیر بیشتری در ۳۰ روز بستری در بیمارستان را دارند.

40. Maraschini

41. Interact Cardiovasc Thorac Surg

از نمرات گرایی به منظور حل مشکل مخدوش شدن در مطالعات غیرکارآزمایی استفاده می‌شود. یک نمره‌ی گرایی، متغیر خلاصه‌ای است که از طریق ترکیب چندین متغیر اندازه‌گیری شده (مثلاً سن، جنسیت، نژاد و غیره) و تبدیل آن به یک متغیر، یعنی نمره‌ی گرایی محاسبه می‌شود. نمره‌ی گرایی را می‌توان به‌عنوان احتمال پیش‌بینی شده‌ی درمان تفسیر کرد (گرایی درمان). این نمره بر اساس مقادیر متغیرهای مشاهده شده‌ای است که به‌منظور محاسبه‌ی نمره‌ی گرایی از آن‌ها استفاده می‌شود.

اگرچه می‌توان از نمره‌ی گرایی به تعدیل چندین متغیر مخدوش کننده اندازه‌گیری شده استفاده کرد، نمی‌توان از آن برای پرداختن به متغیرهای اندازه‌گیری نشده یا ناشناخته‌ای که ممکن است بر نتیجه‌ی مطالعه اثر بگذارند (مثلاً سطوح مختلف سواد بهداشتی در بیمارانی که چند دارو مصرف می‌کنند)، استفاده کرد.

### موارد استفاده‌ی توصیه‌شده

به‌طور کل، فایده‌ی نمره‌ی گرایی، توانایی آن برای توازن متغیرهای همراه بین گروه‌های درمان است (گروه‌های درمانی را از نظرویزگی‌های مشاهده‌شده، تا حد امکان شبیه به هم می‌کند). گاهی اوقات تعداد متغیرهای یک مطالعه، نسبت به تعداد نتایج آن زیاد است. در چنین شرایطی، استفاده از روش نمره‌ی گرایی می‌تواند مفید باشد. پس‌ازآنکه نمره‌ی گرایی ایجاد شد، می‌توان از آن برای هم‌تاسازی (جور کردن) یا محدود سازی طرح مطالعه، طبقه‌بندی، مدل‌سازی، یا تحلیل مطالعه استفاده کرد. موارد استفاده‌ی مختلف نمرات گرایی به‌صورت خلاصه در زیر توصیف شده‌اند.

#### محدودیت نمره‌ی گرایی:

جمعیت به نحوی محدود می‌شود که بیماران موجود در گروه‌های مقایسه مقادیر نمره‌ی گرایی همپوشی داشته باشند. این کار باعث شباهت بیشتر گروه‌ها می‌شود.

#### جور کردن نمره‌ی گرایی:

از نمره‌ی گرایی برای جور کردن بیماران هر یک از گروه‌های مقایسه با دیگری استفاده شده، و اثرات درمان با استفاده از مقایسه‌های جور شده تخمین زده می‌شود. برای مثال، می‌توان بیماران در گروه درمانی الف با نمره‌ی گرایی بین ۰٫۶-۰٫۷ را با بیماران گروه ب که نمره‌ی گرایی بین ۰٫۶-۰٫۷ دارند، جور کرد. این فرایند جور سازی برای کل جمعیت مطالعه و تمامی دامنه‌های نمره‌ی گرایی انجام می‌شود. بنابراین، مقایسه‌ها تنها بین آن دسته از بیماران گروه الف و ب انجام می‌شود که بسیار شبیه به هم هستند (دارای نمره‌ی گرایی یکسانی هستند)، موضوعی که باعث افزایش اعتبار تخمین اثر درمان می‌شود.

### طبقه‌بندی نمره‌ی گرایش:

نمره‌ی گرایش طبقه‌بندی شده و داده‌ها در داخل این طبقه‌ها بررسی می‌شوند. برای مثال، یک پژوهشگر ممکن است تصمیم بگیرد جمعیت مطالعه را به سه دسته تقسیم کند، افرادی که نمرات گرایش پایین، متوسط و بالایی دارند. بررسی اثرات درمان در داخل هر یک از طبقه‌های نمره‌ی گرایش، مخدوش‌کنندگی ناشی از متغیرهایی که نمره‌ی گرایش توسط آن‌ها ایجاد شده است (مثلاً سن، جنس، نژاد و غیره) را کاهش می‌دهد.

### مدل‌سازی نمره‌ی گرایش:

در یک مدل آماری که هدف آن ارزیابی رابطه‌ی بین درمان و نتیجه‌ی مدنظر است، از نمره‌ی گرایش به‌عنوان یک متغیر همراه استفاده می‌شود. پژوهشگران در واقع نمره‌ی گرایش را همانند سایر متغیرهای مستقل موجود در مدل در نظر می‌گیرند. این کار باعث می‌شود تا هرگونه مخدوش شدن بالقوه‌ی ایجاد شده توسط متغیرهای نمره‌ی گرایش (متغیرهایی که از آن‌ها برای ایجاد نمره‌ی گرایش استفاده شده است) رفع شود.

### وزن دادن توسط نمره‌ی گرایش:

از نمره‌ی گرایش جهت توزین مجدد گروه مواجهه، گروه عدم مواجهه، یا هر دو گروه استفاده می‌شود، به نحوی که توزیع نمره‌ی گرایش در هر دو گروه یکسان باشد. این کار اغلب تخمین توزین شده‌ی احتمال معکوس (inverse probability weighted estimation-IPW) نامیده می‌شود و هدف آن اطمینان از مشابه بودن گروه‌های درمان است.

### مسائل و مشکلات بالقوه

استفاده‌ی چندباره از نمره‌ی گرایش (مثلاً محدود سازی، جور کردن) جمعیت مطالعه را محدود می‌کند. محدود سازی، بیمارانی که نمرات گرایش هم‌پوشی ندارند را حفظ نمی‌کند. جور کردن نیز باعث حفظ بیمارانی که توانایی جور شدن ندارند، نمی‌شود. حذف بیماران بدین شکل باعث کاهش اندازه‌ی مطالعه شده و در نتیجه قدرت آماری آن را کاهش می‌دهد.

علاوه بر این، اگرچه برخی کاربردهای نمره‌ی گرایش امکان مقایسه‌ی ساده و فهرست‌وار بیماران موجود در هر یک از گروه‌های مقایسه، پس از ایجاد نمره‌ی گرایش را می‌دهند، در روش‌شناسی‌های پیچیده‌تر (مثلاً وزن دادن، مدل‌سازی) نمی‌توان به‌منظور نشان دادن توازن متغیرهای همراه بین گروه‌های مقایسه، به‌سادگی داده‌ها را وارد جدول و فهرست کرد. به عبارت ساده‌تر، این روش‌های پیچیده‌تر باعث می‌شوند تا پس از استفاده از نمره‌ی گرایش، نشان دادن شباهت گروه‌های مقایسه به‌صورت تصویری (با استفاده از نمودار یا جدول) دشوار باشد.

## قدرت

• استفاده از نمرات گرایش امکان تعدیل چندین متغیر مخدوش کننده مشاهده شده، با استفاده از یک سنجی خلاصه را می دهد.

## محدودیت ها

- اگر تمامی مشاهدات در تحلیل بکار گرفته نشوند، استفاده از نمره ی گرایش گاهی اوقات می تواند باعث کاهش قدرت آماری شود.
- برخی کاربردهای نمره ی گرایش امکان بررسی دیداری ساده ی توازن متغیرهای همراه (پس از ایجاد نمره ی گرایش) را نمی دهند.

## نمونه های انتخابی

### محدودیت

• اشنی وایس و همکاران، ژورنال آرشیو روان پزشکی عمومی (۲۰۱۴). پژوهشگران رابطه ی بین استفاده از داروهای ضد افسردگی مختلف و خطر اقدام به خودکشی را در یک پایگاه داده بررسی کردند. برای تحلیل جمعیت های قابل مقایسه در بین انواع مختلف اختلالات روانی و کلاس های دارویی، پژوهشگران یک نمره ی گرایش ایجاد کردند. این نمره با استفاده از تمامی متغیرهای مخدوش کننده بالقوه (مثلاً سن، جنس، بیماری های هم پوش، داروها، و...) ایجاد شد و جمعیت مطالعه را بر اساس هم پوشی در نمرات گرایش محدود کرد. این کار باعث شد گروه های درمانی شبیه به هم باشند.

### جور کردن

• جوهانس<sup>۴۲</sup> و همکاران، ژورنال فارماکواپیدمیولوژی وایمنی دارو<sup>۴۳</sup> (۲۰۰۷). پژوهشگران این موضوع را بررسی کردند که آیا خطر بیماری شریان کرونری (CHD) بین بیماران دیابتی درمان شده با تiazولیدیندیون و بیماران دیابتی درمان شده با متفورمین خوراکی و سولفونیل اوره متفاوت است یا خیر. نمره ی گرایش بر اساس متغیرهای همراه از پیش تعیین شده (عوامل خطر بیماری قلبی، نشانگرهای شدت بیماری دیابت بیمار و...) محاسبه شد. سپس بیماران موجود در هر یک از گروه های مقایسه بر اساس نمره ی گرایش خود جور شدند. جورشدگی بیماران مخدوش شدن ناشی از متغیرهایی که برای محاسبه ی نمره ی گرایش بکار رفته بودند را کاهش داد.

42. Johannes

43. Pharmacoepidemiol Drug Saf

## طبقه‌بندی

• جار<sup>۴۴</sup> و همکاران، سالنامه‌ی پزشکی داخلی<sup>۴۵</sup> (۲۰۰۵). پژوهشگران دیالیز صفاقی را با همودیالیز از نظر خطر مرگ مقایسه کردند. آن‌ها متوجه شدند بیمارانی که دیالیز صفاقی انجام داده بودند نسبت به بیمارانی که همودیالیز کرده بودند، پروفایل موردی-آماری بهتری در خط پایه دارند. با استفاده از متغیرهای خط پایه‌ای که انتظار می‌رفت باعث مخدوش شدن شوند (مثلاً سن، جنس، نژاد، وضعیت اشتغال)، نمره‌ی گرایش محاسبه شد. در ادامه، برای کاهش مخدوش شدن ناشی از متغیرهایی که برای محاسبه‌ی نمره‌ی گرایش از آن‌ها استفاده شده بود، جمعیت به سه گروه تقسیم شدند.

## مدل‌سازی

• اپل‌گیت<sup>۴۶</sup> و همکاران. ژورنال کاترگذار و مداخلات قلبی عروقی<sup>۴۷</sup> (۲۰۰۶). پژوهشگران ابزارهای بستن عروق (VCDs) و فشردگی دستی را از نظر خطر بروز عوارض عروقی با هم مقایسه کردند. ابتدا نمره‌ی گرایش بر اساس متغیرهای مخدوش شدن بالقوه (مثلاً ویژگی‌های جمعیت شناختی، تظاهرات بالینی، ناحیه سطح بدن و...) محاسبه شد. سپس در یک مدل رگرسیون از نمره‌ی گرایش به‌عنوان یک متغیر استفاده شد تا مخدوش شدن ناشی از متغیرهای مورد استفاده برای محاسبه‌ی نمره‌ی گرایش کاهش یابد.

## وزن کردن

• اسکات<sup>۴۸</sup> و همکاران، سالنامه‌ی نوآوری و پژوهش در جراحی<sup>۴۹</sup> (۲۰۱۰). پژوهشگران جراحی لوبکتومی قفسه‌ی سینه با کمک ویدئو را از نظر خطر عواقب حین عمل (مثلاً نرخ مرگ‌ومیر، نرخ عوارض، و مدت‌زمان ماندن بیمار در اتاق عمل) با لوبکتومی باز مقایسه کردند. نمره‌ی گرایش بر اساس متغیرهای مخدوش‌کنندگی اندازه‌گیری شده (سن، جنس، میزان FEV<sub>1</sub> در حین عمل و...) از نمره‌ی گرایش برای انجام IPW و کاهش مخدوش شدن ناشی از متغیرها استفاده شد (متغیرهایی که از آن‌ها به‌منظور محاسبه‌ی نمره‌ی گرایش استفاده شده بود).

44. Jaar

45. Ann Intern Med

46. Applegate

47. Catheter Cardiovasc Interv

48. Scott

49. Ann Surg Innov Res



روش‌های متغیر ابزاری به منظور تخمین دقیق اثرات درمان در حضور متغیرهای مخدوش کننده اندازه‌گیری نشده (متغیرهای مخدوش کننده یا ناشناخته هستند و یا در داده‌های مورد تحلیل نیستند) طراحی شده‌اند. در این شیوه، متغیر مرتبط با تغییر ناشی از درمان شناسایی شده و از آن برای به توازن رساندن متغیرهای مخدوش کننده اندازه‌گیری نشده در گروه‌های درمان استفاده می‌شود. یک متغیر ابزاری، متغیری است که ۱. به درمان مرتبط است و ۲. نه به صورت مستقیم و نه غیرمستقیم به نتیجه‌ی مدنظر ارتباطی ندارد. برای مثال، مطالعه‌ای که اثرات درمان (intensive) بیماران آنفارکتوس میوکارد حاد و خطر مرگ را بررسی می‌کند، می‌تواند از نزدیکی به مرکز مراقبت ثالث به عنوان متغیر ابزاری استفاده کند. نزدیکی به مرکز مراقبت ثالث رابطه‌ی محکمی با دریافت یا عدم دریافت درمان فشرده برای آنفارکتوس میوکارد حاد دارد و می‌توانیم به صورت منطقی فرض کنیم نزدیکی به مرکز مراقبت ثالث به شکل مستقیم یا غیرمستقیم با مرگ و میر مرتبط نیست.

### موارد استفاده‌ی توصیه شده

روش‌های متغیر ابزاری در شرایطی مفید هستند که احتمال وجود متغیرهای مخدوش کننده چشم‌گیر بالا بوده، و متغیر ابزاری دارای رابطه‌ی محکمی با درمان در دسترس باشد.

### مسائل و مشکلات بالقوه

برای طرح‌های غیرکارآزمایی، شناسایی و ارزیابی متغیرهای ابزاری بالقوه می‌تواند بسیار دشوار باشد. پژوهشگران باید از دانش تخصصی برای توجیه استفاده از متغیر ابزاری استفاده کنند. هیچ راه تجربی‌ای وجود ندارد که بتوان از طریق آن مشخص کرد یک متغیر ابزاری به صورت مستقیم یا غیرمستقیم با نتیجه‌ی مطالعه ارتباط دارد یا خیر. در صورتی که متغیر ابزاری رابطه‌ی محکمی با درمان نداشته باشد، یا با نتیجه‌ی آن رابطه داشته باشد، نتایج ممکن است گمراه‌کننده باشند.

علاوه بر این، روش‌های متغیر ابزاری ناکارآمد بوده و به منظور تخمین دقیق اثر درمان، نیازمند نمونه‌ی بزرگی هستند و هنگامی که پژوهشگران احتمال وجود مخدوش کننده اندازه‌گیری نشده‌ی زیادی را می‌دهند، از این روش‌ها به عنوان یک روش تحلیلی اصلی استفاده می‌کنند. در دیگر موارد، روش‌های متغیر ابزاری به عنوان تحلیل ثانویه به کار گرفته می‌شوند.

### قدرت

• روش‌های ابزاری پتانسیل لازم برای کاهش چشم‌گیر سوگیری به دلیل مخدوش کننده اندازه‌گیری نشده را دارند.

## محدودیت‌ها

- در دسترس بودن متغیرهای ابزاری معتبر باعث محدودیت این روش می‌شود (آن دسته از متغیرهای ابزاری که با درمان رابطه‌ی محکمی داشته و با نتیجه‌ی مطالعه به صورت مستقیم یا غیرمستقیم رابطه نداشته باشند).
- در صورتی که متغیر ابزاری به صورت مستقیم یا غیرمستقیم با نتیجه رابطه داشته باشد، تخمین اثر درمان دچار سوگیری خواهد شد (و این رابطه را نمی‌توان به صورت تجربی تأیید یا رد کرد).
- در صورتی که رابطه‌ی بین متغیر ابزاری و درمان ضعیف باشد، سوگیری‌ها تقویت خواهند شد.
- تحلیل‌های سنتی متغیر ابزاری بیمارانی که ممکن است در طول زمان از درمان خارج شوند یا تغییر گروه دهند را مدنظر قرار نمی‌دهند. به همین دلیل، به روش‌های تحلیل پیچیده‌تری نیاز است.
- علاوه بر این، در صورت ناهمگن بودن اثر درمان، به فرضیات غیرقابل تأییدی نیاز خواهد بود.
- روش متغیر ابزاری نسبت به دیگر روش‌های تعدیل کارآمدی کمتری دارد (بازهی اطمینان بزرگ‌تر).

## نمونه‌های انتخابی

- بروکس و همکاران<sup>۵۰</sup>، ژورنال پژوهش خدمات انسانی<sup>۵۱</sup> (۲۰۰۳). این مطالعه جراحی حفظ پستان همراه با رادیوتراپی (BCSI) را از نظر میانگین زنده ماندن با ماستکتومی مقایسه کرد. در این مطالعه از دو متغیر ابزاری استفاده شد: درصد BCSI در ناحیه‌ی بیمار (به صورت دو گروه بیشتر و کمتر. در صورتی که کمتر ۲۰ درصد از جراحی‌های سرطان سینه‌ی مراحل اولیه در ۵۰ مایلی اطراف محل سکونت بیمار انجام می‌شد، بیمار در گروه کمتر قرار می‌گرفت) و فاصله تا یک مرکز درمان رادیوتراپی (به صورت دو گروه نزدیک و دور. اگر فاصله تا مرکز درمان رادیوتراپی در سال تشخیص سرطان کمتر از ۱۹ مایل بود، بیمار در گروه نزدیک قرار می‌گرفت). این متغیرهای ابزاری به این دلیل انتخاب شدند که نزدیکی به مراکز درمانی مدنظر رابطه‌ی محکمی با انجام دادن درمان دارد. علاوه بر این، به طور منطقی فرض بر این گذاشته شد که نزدیک بودن به این مراکز به صورت مستقیم یا غیرمستقیم ربطی به میانگین زنده ماندن ندارد.

50. Brooks

51. Health Serv Res

• اشنی وایس و همکاران، ژورنال پزشکی نیوانگلند<sup>۵۲</sup> (۲۰۰۸). این مطالعه به بررسی نقش آپروتینین (یک پروتئین تزریقی که هنگام جراحی‌های پیچیده از آن استفاده می‌شود) و استفاده از آن حین جراحی کنارگذر سرخرگ تاجی (CABG) و خطر مرگ پرداخت. ترجیح جراح برای استفاده از آپروتینین، متغیر ابزاری محسوب می‌شد. اگر جراح این دارو را برای بیش از ۹۰ درصد از بیماران تجویز می‌کرد، آپروتینین داروی ترجیحی وی محسوب می‌شد. در صورتی که جراح داروی را برای کمتر از ۱۰ درصد از بیماران خود تجویز می‌کرد، آپروتینین داروی ترجیحی وی محسوب نمی‌شد. پژوهشگران همچنین به تفاوت بین ۱۰۰ درصد در مقابل ۰ درصد توجه کردند. (اگر یک جراح دارو را برای ۱۰۰ درصد از بیماران خود تجویز می‌کرد، آپروتینین را ترجیح می‌داد، و اگر برای هیچ‌یک از بیماران خود [۰ درصد] تجویز نمی‌کرد، دارو را ترجیح نمی‌داد) از آنجایی که ترجیح جراح برای استفاده از این دارو رابطه‌ی محکمی با درمان انتخابی داشت، به عنوان متغیر ابزاری انتخاب شد. (از آنجایی که این مورد یک مطالعه‌ی غیرکارآزمایی بود، درمان انتخابی بنا به صلاحدید جراح انتخاب می‌شد.) همچنین، به طور منطقی فرض بر این گذاشته شد که ترجیح جراح برای استفاده یا عدم استفاده از آپروتینین رابطه‌ی مستقیم یا غیرمستقیمی با خطر مرگ نداشت.

تحلیل حساسیت اثر بالقوه‌ی متغیرهای مخدوش کننده اندازه‌گیری نشده بر نتایج مطالعه را می‌سنجد. پس از آنکه همه‌ی تلاش‌ها جهت رفع، کاهش، یا کنترل سوگیری در طراحی مطالعه، جمع‌آوری داده‌ها، و تحلیل آن‌ها انجام شده باشد، تحلیل حساسیت آخرین خط دفاعی در برابر سوگیری است.

به منظور بررسی وجود مخدوش کننده اندازه‌گیری نشده، فرضیات آگاهانه‌ای ایجاد می‌شوند. هدف این فرضیات کمی ساختن اثر مخدوش‌گرهای اندازه‌گیری نشده بر اثر درمان است.

### انواع تحلیل حساسیت

چندین رویکرد برای انجام تحلیل حساسیت وجود دارد. هر یک از این رویکردها محدودیت‌ها و کاربردهای خاص خود برای مطالعات غیرتجربی دارند. مهم است که فرضیات هر مدل را بشناسیم، محاسبات را انجام داده و به تفسیر نتایج توجه کنیم. صفحات گسترده و پکیج‌های آماری اغلب تحلیل را ساده‌تر کرده و تصاویری گرافیکی ایجاد می‌کنند که برای درک نتایج مفید هستند. دو رویکرد رایج برای انجام تحلیل حساسیت شامل رویکرد آرایه (array approach) و رویکرد کنار گذاشتن (rule-out approach) است، از رویکرد آرایه برای درک چگونگی اثرگذاری قدرت یک متغیر مخدوش کننده اندازه‌گیری نشده (که بین گروه‌های درمانی متوازن نیست) بر اثر درمانی مشاهده شده استفاده می‌شود.

نتیجه‌ی نهایی این رویکرد یک آرایه است که تخمین‌های خطر متفاوت را در طول دامنه‌ی وسیعی از مقادیر پارامتر نشان می‌دهد. برای مثال، اگر شیوع سیگار کشیدن در یک جمعیت مشخص نباشد، جمعیت مشاهده‌ی کننده‌ی تغییرات مرتبط در تخمین اثر ممکن است از ۲۰ تا ۹۰ درصد متفاوت باشد.

رویکرد کنار گذاشتن برای مشخص شدن این موضوع بکار می‌رود که برای تبیین اثر درمانی مشاهده شده، به چه میزان مخدوش کننده ناشی از یک متغیر نیاز است. می‌توان آن دسته از متغیرهای مخدوش کننده که قدرتشان برای حذف اثر درمانی مشاهده شده کافی نیست را کنار گذاشت (نفی کرد)

### موارد استفاده‌ی توصیه شده

از تحلیل حساسیت باید در مواردی استفاده کرد که احتمال وجود متغیرهای مخدوش کننده اندازه‌گیری نشده می‌رود. تحلیل حساسیت میزان سوگیری را مشخص می‌کند. اگرچه در این خلاصه‌ی ارائه شده، به طور کامل شرح داده نشده است، می‌توان از تحلیل حساسیت برای سنجش حساسیت یافته‌های مطالعه نسبت به تغییرات مواجهه، تعاریف خروجی، و دیگر فرضیات ایجاد شده در حین

انجام مطالعه استفاده کرد.

## مسائل و مشکلات بالقوه

اگرچه چندین رویکرد کمی برای سنجش حساسیت نتایج مطالعه نسبت به متغیرهای مخدوش کننده بالقوه و اندازه‌گیری نشده وجود دارد، سنجش این موضوع که آیا تحلیل در واقع نسبت به مخدوش کننده اندازه‌گیری نشده غیر حساس است یا خیر، هنوز توسط قضاوت انسانی انجام می‌شود.

## نقاط قوت

• تحلیل حساسیت می‌تواند اثرات بالقوه‌ی متغیرهای مخدوش کننده اندازه‌گیری نشده‌ی مطالعه را بسنجد.

## محدودیت‌ها

• هر یک از رویکردها محدودیت‌های خاص خود را دارند. رویکرد کنار گذاشتن محدود به یک متغیر مخدوش کننده دوتایی بوده و بزرگی اختلاط را اندازه‌گیری نمی‌کند. این کار به چندین رویکرد دیگر که نیازمند درک فنی گسترده و مهارت‌های برنامه‌نویسی هستند نیاز دارد. پژوهشگران باید با محدودیت‌های هر رویکرد آشنا بوده و تحلیل مناسب را انتخاب کنند.

## مثال‌های انتخابی

• گرون ولد<sup>۵۳</sup> و همکاران، ژورنال بین‌المللی همه‌گیرشناسی<sup>۵۴</sup> (۲۰۱۰). پژوهشگران روش‌های مختلف تحلیل حساسیت را مورد بحث قرار داده و آن‌ها را بر روی یک نمونه‌ی بالینی واکسیناسیون آنفلوانزا و نرخ مرگ‌ومیر اعمال می‌کنند. در یک مثال، پژوهشگران هنگام تغییر چندین رابطه، اتفاق رخ داده برای اثر درمان تخمینی را شبیه‌سازی می‌کند. ۱. قدرت رابطه‌ی بین متغیر مخدوش کننده اندازه‌گیری نشده و وضعیت واکسیناسیون؛ ۲. قدرت رابطه‌ی بین متغیر مخدوش کننده اندازه‌گیری نشده و خطر مرگ‌ومیر؛ و ۳. شیوع متغیر مخدوش کننده. پژوهشگران نتیجه گرفتند که برای حذف اثرات درمان (رسیدن به نتیجه‌ی هیچ‌گونه اثر درمانی)، شیوع متغیر اختلاطی باید بین ۲۰-۴۰ بوده، و رابطه‌ی بین متغیر مخدوش کننده اندازه‌گیری نشده و واکسیناسیون و مرگ‌ومیر باید بسیار محکم باشد (نسبت شانس برابر با ۳/۰).

## اطلاعات خارجی

### مقدمه

در مطالعات غیرکارآزمایی، اطلاعات مهم مانند ویژگی‌های بالینی بیمار در مطالعات claims-based و داده‌های آزمایشگاهی گاهی در مطالعات مبتنی بر پرسشنامه ممکن است در دسترس نباشند. برای

53. Groenwold

54. Intern J of Epidemiol

مثال در مطالعاتی که به بررسی پیامدهای ترمبوآمبولیسم می‌پردازند، وضعیت مصرف دخانیات و چاقی که هم جز عوامل خطر محسوب شده و هم متغیرهای مخدوش کننده بالقوه هستند، ممکن است در منبع داده ثبت نشده باشند. اگرچه تحلیل حساسیت ساده می‌تواند این متغیرهای مخدوش کننده بالقوه را ارزیابی کند، اما هر یک از متغیرها (مصرف دخانیات یا چاقی) را به عنوان یک متغیر مستقل در نظر می‌گیرد و اثر ترکیبی متغیرهای اندازه‌گیری نشده (هم مصرف دخانیات و هم چاقی) مدنظر قرار نمی‌گیرد. هنگامی که گمان می‌رود هر یک از متغیرهای همراه به تنهایی باعث ایجاد نتایج متفاوتی شوند (در مقایسه با اثر مشترک متغیرها)، این موضوع مایه‌ی نگرانی است. می‌توان از اطلاعات خارجی برای تعدیل چندین متغیر مخدوش کننده اندازه‌گیری نشده و اثر مشترک آن‌ها استفاده کرد. دو نوع اطلاعات خارجی وجود دارد که می‌توان از آن‌ها استفاده کرد: داده‌های مربوط به افراد در داخل جمعیت اصلی مطالعه (مطالعات اعتبارسنجی درونی) یا داده‌های مربوط به افراد در خارج از جمعیت اصلی مطالعه (مطالعات اعتبارسنجی بیرونی).

### مطالعات اعتبارسنجی درونی

مطالعات اعتبارسنجی درونی از داده‌های اضافی جمع‌آوری شده از زیرمجموعه‌ای از شرکت‌کنندگان جمعیت اصلی مطالعه استفاده می‌کنند. این نمونه‌ی اضافی می‌تواند یک زیر نمونه‌ی تصادفی از کوهورت (زیرگروه)، یا نمونه‌ای غیرتصادفی باشد که بر اساس اطلاعات مربوط به مواجهه، پیامد و یا هر دو انتخاب شده است. نمونه‌گیری غیرتصادفی «طرح دومرحله‌ای» نامیده می‌شود و نیازمند تحلیل‌هایی است که مکانیسم نمونه‌گیری استفاده شده را مدنظر قرار می‌دهند.

### مطالعات اعتبارسنجی بیرونی

مطالعات اعتبارسنجی خارجی از داده‌های اضافی جمع‌آوری شده از خارج از جمعیت اصلی مطالعه استفاده می‌کنند. داده‌های خارجی معمولاً داده‌های زمینه‌یابی مقطعی بوده و مختص یک فرضیه خاص نیستند. جمع‌آوری این داده‌ها اغلب ارزان است.

### موارد استفاده‌ی توصیه شده

مطالعات اعتبارسنجی بیرونی و درونی زمانی مفید هستند که پایگاه داده‌ی موجود حاوی اطلاعات مهم در مورد عوامل خطر و متغیرهای مخدوش کننده بالقوه‌ی مدنظر نیست. استفاده از اطلاعات خارجی به پژوهشگر اجازه می‌دهد علاوه بر تعدیل چندین متغیر مخدوش کننده بالقوه، مخدوش کنندگی مشترک بالقوه‌ی ناشی از متغیرهای همراه اندازه‌ی گیری نشده را بسنجد (مثلاً اثر مشترک وضعیت مصرف دخانیات و چاقی در مثال بالا).

### مسائل و مشکلات بالقوه

بر اساس در دسترس بودن اطلاعات، نوع مطالعه‌ی اعتبارسنجی (درونی یا بیرونی)، و روش نمونه‌گیری

(تصادفی یا غیرتصادفی)، به روش‌های آماری متفاوتی نیاز خواهد بود. ممکن است از روش‌هایی مانند propensity score calibration یا multiple imputation، maximum likelihood، estimating equations استفاده شود. علاوه بر این، به منظور استفاده‌ی درست از روش‌های آماری، به درک کاملی از این روش‌ها و فرضیه‌های هر یک نیاز است.

### نقاط قوت

- می‌توان به منظور تعدیل چندین متغیر مخدوش کننده اندازه‌گیری نشده اطلاعات خارجی را تحلیل کرد.
- مطالعات اعتبارسنجی خارجی ممکن است برای چندین مطالعه‌ی اصلی بکار گرفته شوند، زیرا مختص هیچ‌یک از سوالات خاص پژوهش نیستند.

### محدودیت‌ها

- تنها زمانی می‌توان از مطالعات اعتبارسنجی درونی استفاده کرد که امکان جمع‌آوری داده‌های اضافی بیماران از منبع داده وجود داشته باشد.
- مطالعات اعتبارسنجی خارجی اغلب دارای ابزارهای سنجش دقیق نبوده، یا به‌طور کامل نماینده جمعیت اصلی مطالعه نیستند.

### نمونه‌های انتخابی

#### مطالعات اعتبارسنجی درونی

• انگ<sup>۵۵</sup> و همکاران، ژورنال فارماکوپایدمیولوژی و ایمنی دارو (۲۰۰۸). پژوهشگران با استفاده از یک پایگاه داده‌ی اجرایی claims مطالعه‌ای را ترتیب دادند که استفاده از داروهای ضدبارداری خوراکی و خطر ترمبوآمبولیسم را بررسی می‌کرد. از آنجایی که برخی از عوامل خطر خاص ترمبوآمبولیسم (مثلاً چاقی و مصرف سیگار) که جزو متغیرهای مخدوش کننده بالقوه نیز هستند در پایگاه داده موجود نبودند، از یک طرح هم‌گروهی موردی برای ارزیابی مخدوش کننده باقی‌مانده‌ی ناشی از متغیرهای اندازه‌گیری نشده در هم‌گروهی مطالعه‌ی اصلی استفاده شد. برای ایجاد هم‌گروه فرعی، بیماران به‌صورت تصادفی از هم‌گروه مطالعه‌ی اصلی انتخاب شدند. اطلاعات اضافی از هم‌گروه فرعی جمع‌آوری شده و از آن برای انتساب مقادیر از دست‌رفته در هم‌گروهی اصلی استفاده شد.

• واکر<sup>۵۶</sup> و همکاران لنست (۱۹۸۱) پژوهشگران مطالعه‌ای ترتیب داده که طی آن وازکتومی و خطر آنفارکتوس میوکارد غیر کشنده در یک جمعیت مربوط به نظام ارجاع خدمات پزشکی بررسی شد. در نمونه‌ی اولیه و پس از کنترل سال تولد و مدت زمان مشاهده‌ی بیمار، میزان شیوع آنفارکتوس میوکارد غیر کشنده در مردان بدون وازکتومی کمی کمتر بود به‌منظور تعدیل آن دسته از عوامل خطر قلبی که ممکن بود باعث مخدوش کنندگی در نتایج شوند، سوابق پزشکی اعضای منتخب بیماران بررسی

55. Eng

56. Walker

شد. با استفاده از نمونه‌گیری غیرتصادفی، زیرمجموعه‌ای از جمعیت که دارای اطلاعات پزشکی بودند، شناسایی شدند.

#### مطالعات اعتبارسنجی خارجی

• استرمر<sup>۵۷</sup> و همکاران ژورنال همه‌گیرشناسی آمریکا (۲۰۰۵) پژوهشگران برای بررسی استفاده از داروهای ضدالتهاب غیراستروئیدی و مرگ‌ومیر یک‌ساله در جمعیت مدیکید نیوجرسی، مطالعه‌ای را اجرا کردند. تخمین اولیه با استفاده از نمره‌ی گرایش برای تعدیل متغیرهای مخدوش‌کننده اندازه‌گیری شده (سن، نژاد، جنس، دیگر بیماری‌ها، داروها و...) حاکی از این بود که داروهای ضدالتهاب غیراستروئیدی با کاهش کوتاه‌مدت مرگ‌ومیر در بیماران سالمندی بستری همراه است. با این حال، پژوهشگران گمان می‌کردند متغیرهای مخدوش‌کننده دیگری وجود دارند که باعث سوگیری تخمین اثر درمان شده‌اند. در این مطالعه از داده‌های زمینه‌یابی ذی‌نفع بیمه‌ی مدیکر (MCBS) که یک زمینه‌یابی سیار است، استفاده شد. داده‌های مربوط به متغیرهای مخدوش‌کننده اندازه‌گیری نشده (مثلاً وضعیت مصرف دخانیات، شاخص توده بدنی، تحصیلات، درآمد و...) از جمعیتی کسب شد که توزیع سن و جنس آن‌ها همانند مطالعه‌ی اصلی بود. سپس اثر تخمینی مطالعه‌ی اصلی با استفاده از فنون آماری تعدیل‌شده و دیگر رابطه‌ای بین داروهای ضدالتهاب غیراستروئیدی و مرگ‌ومیر یک‌ساله مشاهده نشد.



## واژه نامه

تعریف	ترجمه انتخابی	واژه
به «اعتبار بیرونی - external validity» مراجعه کنید.	کاربردپذیری	Applicability
رابطه‌ی بین دو متغیر، به نحوی که اگر یک متغیر تغییر کند، دیگری نیز به شیوه‌ای قابل پیش‌بینی تغییر می‌کند. رابطه‌ی مثبت زمانی رخ می‌دهد که افزایش یک متغیر باعث افزایش دیگری شود. رابطه‌ی منفی زمانی رخ می‌دهد که افزایش یک متغیر باعث کاهش دیگری شود. وجود رابطه به معنای رابطه‌ی علت و معلولی نیست. همبستگی نیز نامیده می‌شود.	رابطه	Association
یک خطای نظام‌مند در طرح مطالعه که باعث ارزیابی نادرست از اثر مداخله بر پیامدهای اندازه‌گیری شده می‌شود. در کارآزمایی‌های بالینی، انواع اصلی سوگیری ناشی از تفاوت نظام‌مند بین گروه‌های مورد مقایسه در مطالعه (سوگیری انتخاب)، مواجهه با عواملی غیر از مداخله‌ی مدنظر (سوگیری عملکرد)، کناره‌گیری یا حذف شرکت‌کننده (سوگیری کناره‌گیری)، یا ارزیابی پیامدها (سوگیری تشخیص) می‌شوند. علاوه بر این، بررسی مطالعات قبلی ممکن است به‌طور خاص در معرض سوگیری گزارش دهی باشد. در این نوع سوگیری، از بین همه‌ی داده‌های مرتبط در دسترس، زیرمجموعه‌ای که سوگیری دارد انتخاب می‌شود.	سوگیری	Bias
یک کارآزمایی بالینی زمانی «کور» است که شرکت‌کننده نداند در کدام شاخه (گروه) قرار دارد. دوسر کور یعنی هم شرکت‌کننده و هم پژوهشگر نمی‌دانند شرکت‌کننده کدام درمان را دریافت خواهد کرد.	کورسازی	Blinding
یک مطالعه‌ی غیرکارآزمایی که افراد مبتلا به یک بیماری خاص یا پیامد مدنظر (موارد) را با افرادی از همان جمعیت مقایسه می‌کند که فاقد آن بیماری یا پیامد هستند (کنترل‌ها). هدف این مطالعه یافتن رابطه بین پیامد با مواجهه‌ی قبلی و عوامل خطر خاصی است. این طرح به‌طور خاص زمانی سودمند است که پیامد مدنظر نادر بوده و امکان اندازه‌گیری قابل اعتماد مواجهه‌ی قبلی وجود دارد. مطالعات مورد شاهده‌ی می‌توانند گذشته‌نگر یا آینده‌نگر باشند.	مطالعه مورد شاهده‌ی	Case-control study
رابطه‌ی بین دو ویژگی که دلیل آن علت و معلولی است (تغییر در یک ویژگی باعث بروز تغییر در دیگری می‌شود).	علیت	Causality
به «علیت» مراجعه کنید.	علت و معلول	Causation

واژه	ترجمه انتخابی	تعریف
Cohort study	مطالعه هم‌گروهی	یک مطالعه‌ی غیرکارآزمایی با گروهی تعریف شده از شرکت‌کنندگان (هم‌گروه) که در طول زمان دنبال می‌شوند. نتایج بین زیرمجموعه‌های این هم‌گروهی که در معرض یک مداخله (یا دیگر عوامل مدنظر) قرار گرفته یا قرار نگرفته بودند (یا سطوح مواجهه متفاوتی داشتند)، مقایسه می‌شوند. مطالعات هم‌گروهی می‌توانند گذشته‌نگر یا آینده‌نگر باشند.
Comorbidity	هم‌پوشی	هنگامی که دو یا چند بیماری به صورت هم‌زمان در بیمار وجود دارند.
Comparative Effectiveness (Research) (CER)	پژوهش اثربخشی مقایسه‌ای	تولید و ترکیب اطلاعات به منظور مقایسه‌ی مزایا و منافع روش‌های دیگر جهت پیشگیری، تشخیص، درمان و نظارت بر یک بیماری بالینی یا بهبود مراقبت. هدف پژوهش اثربخشی مقایسه‌ای کمک به مصرف‌کنندگان، متخصصان بالینی، خریداران و سیاست‌گذاران برای اتخاذ تصمیم‌های آگاهانه‌ای است که باعث بهبود مراقبت‌های بهداشتی درمانی در سطح فردی و عمومی می‌شوند.
Comparison group	گروه‌های مقایسه	به بخش «گروه کنترل» نگاه کنید.
Confidence interval	بازه‌ی اطمینان	دامنه‌ای از مقادیر برای یک متغیر مدنظر (مثلاً یک نرخ). این دامنه به نحوی ساخته شده است که احتمال شمول مقدار واقعی متغیر در آن مشخص است. احتمال مشخص شده سطح اطمینان نام داشته و نقاط پایانی بازه‌ی اطمینان حدهای اطمینان نامیده می‌شوند.
Confounder	مخدوش کننده	به بخش «متغیر مخدوش کننده - Confounding variable» مراجعه کنید.
Confounding by indication	مخدوش شدن از طریق انتخاب درمان	زمانی رخ می‌دهد که تشخیص یا دیگر ویژگی‌های بالینی که باعث استفاده از درمان خاصی شده‌اند، به پیامد بیمار مرتبط باشند.
Confounding variable	متغیر مخدوش کننده	یک متغیر (یا ویژگی) که احتمال حضور آن در یک گروه از شرکت‌کنندگان نسبت به گروه دیگر بیشتر بوده، به پیامد مدنظر مرتبط بوده و پتانسیل مغشوش کردن (مخدوش کردن) نتایج را دارد. برای مثال، اگر افراد گروه کارآزمایی یک کارآزمایی کنترل شده جوان‌تر از افراد گروه کنترل باشند، نمی‌توان به آسانی مشخص کرد که خطر پایین‌تر مرگ در یک گروه به دلیل مداخله است یا تفاوت سنی دلیل خطر پایین‌تر است (سن همان متغیر مخدوش کننده است). مخدوش کنندگی یکی از نگرانی‌های اصلی در مطالعات غیرتصادفی به شمار می‌رود. مخدوش کننده نیز نامیده می‌شود.
Control group	گروه کنترل	شرکت‌کنندگان موجود در شاخه‌ی کنترل مطالعه.
Correlation	همبستگی	به «رابطه - association» مراجعه کنید.
Covariate	متغیر همراه	یک متغیر مستقل که مطالعه آن را دست‌کاری نکرده و بر پاسخ تأثیر می‌گذارد.

واژه	ترجمه انتخابی	تعریف
Cox proportional hazard model	مدل مخاطره متناسب کاکس	یک مدل آماری که به منظور تحلیل داده های survival از آن استفاده می شود.
Effect modification	تغییر دادن اثر	شرایطی که در آن سنجش اثر در مقادیر مختلف یک متغیر دیگر متفاوت است (مثلاً، سنجش اثر در مقادیر مختلف نژاد، سن، و... تفاوت دارد).
Efficacy	سودمندی، کارآمدی	(یک دارو یا درمان) حداکثر توانایی یک دارو یا درمان برای تولید یک نتیجه، فارغ از میزان (دوز) تجویز شده. یک دارو در صورتی کارآزمایی های سودمندی را پشت سر می گذارد که در دوز آزمون شده و علیه بیماری مدنظر (بیماری که داروی برای آن تجویز شده است) موثر باشد.
Endpoint	نقطه ی پایان	به «پیامد - outcome» مراجعه کنید.
Experimental study	مطالعه ی تجربی	مطالعه ای که در آن پژوهشگران به منظور آزمودن یک فرضیه به صورت فعالانه مداخله می کنند. هنگامی که از شرکت کنندگان انسانی به عنوان آزمودنی استفاده شود، کارآزمایی یا کارآزمایی بالینی نامیده می شود.
Explanatory trials	کارآزمایی تبیینی (توضیحی)	یک کارآزمایی کنترل شده که هدف آن سنجش مزایای یک مداخله در شرایط ایده ال (سودمندی) است. این کاراز طریق آزمودن فرضیه ها در یک پژوهش علی انجام می شود. کارآزمایی های مربوط به مداخلات بهداشتی-درمانی اغلب یا تبیینی و یا عملی هستند. به «کارآزمایی عمل گرایانه - pragmatic trial» نیز مراجعه کنید.
External validity	اعتبار بیرونی	اینکه تا چه حد نتوان نتایج مطالعه را به دیگر جمعیت ها یا موقعیت ها تعمیم داد. همچنین تعمیم پذیری، کاربردپذیری نامیده می شود. به «کاربردپذیری - applicability» نیز مراجعه کنید.
Generalizability	تعمیم پذیری	به «اعتبار بیرونی - External validity» مراجعه کنید.
Homogeneous	همگن	مشابه بودن شرکت کنندگان، مداخلات، و سنجش پیامدها در مجموعه ای از مطالعات
Hypothesis testing	آزمون فرضیه	یک چارچوب عینی و عملی برای تعیین احتمال درست بودن یک فرضیه
Incidence	بروز	تعداد موارد جدید یک رویداد (بیماری) که طی یک بازه ی زمانی مشخص و در یک جمعیت تعریف شده ی در معرض خطر بروز می کنند. بروز به صورت تناسب بیان می شود.
Incidence rate	نرخ بروز	یک مقیاس فروانی که بر اساس آن یک رویداد (مانند موارد جدید از یک بیماری) طی یک دوره ی زمانی در یک جمعیت رخ می دهد. در این نسبت، جمعیت در معرض ریسک در مخرج، و تعداد موارد ابتلا در طی یک دوره ی زمانی مشخص در صورت کسر قرار می گیرد.

واژه	ترجمه انتخابی	تعریف
Internal validity	اعتبار درونی	میزان صحت و دقت طرح و اجرای مطالعه و جلوگیری از بروز سوگیری. مطالعاتی که با دقت بیشتری طراحی شده باشند، به احتمال بیشتر نتایج درست‌تری را تولید خواهند کرد. به «سوگیری - bias» مراجعه کنید.
Intra-class correlation (coefficient) (ICC)	ضریب همبستگی درون طبقه‌ای	از این ضریب برای ارزیابی همبستگی بین گروه‌های داده (مثلاً پایایی بین ارزیابان) استفاده می‌شود.
Logistic regression	رگرسیون لجستیک	یک فن آماری که احتمال یک متغیر وابسته‌ی دو حالت (مانند مرگ و زندگی) را پیش‌بینی می‌کند. در این فن معمولاً از ترکیبی از متغیرهای مستقل کمی و کیفی برای پیش‌بینی استفاده می‌شود.
Masked	پوشیده	به «کورسازی - blinding» مراجعه کنید.
Matching	جور کردن	زمانی که موارد بیمار به منظور کاهش اثر عوامل مخدوش‌کنندگی بر رابطه‌ی مدنظر با موارد کنترل دارای عوامل مخدوش‌کننده، مشابه (سن، جنسیت، شاخص توده‌ی بدنی و...) «جور» می‌شوند.
Multiple imputation	انتساب چندگانه	روشی جهت پیش‌بینی و پُر کردن داده‌های گم‌شده یک مطالعه بر اساس داده‌های مشاهده‌شده و الگوی داده‌های گم‌شده.
Multivariate analysis	تحلیل چند متغیری	در این روش یک مدل ریاضیاتی ساخته می‌شود که رابطه‌ی بین مواجهه، بیماری، و متغیرهای مخدوش‌کننده را توصیف می‌کند.
Nested case-control study	مطالعه مورد شاهدی لانه‌گزیده	مطالعه‌ای که در آن بیماران و کنترل از نمونه‌های یک مطالعه‌ی هم‌گروهی انتخاب می‌شود (یک مطالعه مورد شاهدی در داخل یک مطالعه‌ی هم‌گروهی لانه می‌گزیند).
New-user design	طرح کاربرجدید	نوعی مطالعه که تحلیل را محدود به افراد مورد مشاهده در شروع دوره‌ی درمان فعلی می‌کند.
Nonexperimental study design	طرح مطالعاتی غیرکارآزمایی	نوعی مطالعه که طی آن پژوهشگران خط سیر رخدادها را بررسی کرده و بر روی شرکت‌کنندگان مداخله‌ای اعمال نمی‌کنند. همچنین مطالعه مشاهده‌ای - observational نامیده می‌شود.
Observational study	مطالعه مشاهده‌ای	به «طرح مطالعاتی غیرکارآزمایی - Nonexperimental study design» مراجعه کنید.
Outcome	(پیامد) نتیجه	نتیجه‌ی یک مطالعه تجربی (کارآزمایی) که از آن برای ارزیابی سودمندی یک مطالعه استفاده می‌شود. همچنین نقطه‌ی پایان - endpoint نامیده می‌شود.
Patient registry	ریجستری (ثبت) بیمار	یک سیستم سازمان‌یافته که از روش‌های مطالعه‌ی غیرکارآزمایی به منظور کسب اطلاعات همسان (اطلاعات بالینی یا دیگر اطلاعات) استفاده کرده و به ارزیابی پیامدهای خاص برای یک جمعیت مشخص می‌پردازد. در این سیستم جمعیت مشخص بر اساس یک بیماری، شرایط، یا مواجهه‌ی خاص تعریف شده و دارای یک یا چند هدف از پیش تعیین‌شده‌ی علمی، بالینی، یا سیاسی است.

واژه	ترجمه انتخابی	تعریف
Pharmacoepidemiology	فارماکواپیدمیولوژی (همه‌گیرشناسی دارویی)	مطالعه‌ی استفاده، اثرات، و پیامدهای یک درمان دارویی از دیدگاهی اپیدمیولوژیک (جمعیت شناختی)
Placebo	دارونما	یک دارو، درمان، یا فرایند غیرفعال که فاقد ارزش درمانی است. در کارآزمایی‌های بالینی، درمان‌های تجربی اغلب به منظور ارزیابی سودمندی درمان با دارونما مقایسه می‌شوند.
Power	(قدرت) توان	احتمال رد شدن فرضیه‌ی صفر، هنگامی که یک فرضیه‌ی جایگزین خاص صحیح است. توان یک آزمون فرضیه برابر با یک منهای احتمال خطای نوع دوم است. در کارآزمایی‌های بالینی، توان این‌گونه تعریف می‌شود: احتمال کارآزمایی برای شناسایی اثر مداخله‌ای که از نظر آماری چشم‌گیر است. مطالعاتی که دارای تعداد مشخصی از شرکت‌کنندگان هستند، توان بیشتری برای کشف اثرات بزرگ (نسبت به اثرات کوچک) دارند. به طور کل، هنگام محاسبه‌ی اندازه‌ی نمونه، توان به صورت 80 درصد یا بیشتر تعیین می‌شود. توان آماری - statistical power نیز نامیده می‌شود.
Precision	دقت	در آمار، درجه‌ی قطعیت اثر تخمینی برای یک پیامد، دقت نامیده می‌شود. هرچه دقت بالاتر باشد، خطای اندازه‌گیری کمتر است. بازه‌های اطمینان موجود در اطراف تخمین یک اثر، یک راه برای بیان دقت هستند. هرچه دقت بازه‌ی اطمینان کوچک‌تر باشد، دقت بالاتر خواهد بود.
Prevalence	شیوع	نسبتی از جمعیت که در یک نقطه‌ی خاص زمانی مبتلا به یک بیماری یا حالت هستند. اگرچه اغلب به‌طور نادرست از عبارت نرخ شیوع استفاده می‌شود، شیوع در واقع یک نرخ - rate نیست.
Probability	احتمال	به فراوانی نسبی وقوع یک رویداد در بی‌نهایت کارآزمایی، احتمال آن رویداد گفته می‌شود.
Prospective study	مطالعه‌ی آینده‌نگر	مطالعه‌ای که در آن پژوهشگر قبل از رخ دادن پیامدها، مواجهه‌ها را اندازه‌گیری می‌کند.
P-value	P-value	احتمال (دامنه‌ی آن از صفر تا یک) اینکه نتایج مشاهده شده در یک مطالعه (یا نتایج شدیدتر) به دلیل شانس رخ داده باشند.
Quasi-experimental	تحقیقات شبه کارآزمایی	مطالعه‌ای که شبیه به یک طرح کارآزمایی است، اما در آن شرکت‌کنندگان به صورت تصادفی در گروه‌های کنترل و درمان قرار نمی‌گیرند. در شرایطی که پژوهشگر نمی‌تواند تمامی عوامل مؤثر بر نتیجه را کنترل کند، ممکن است برای افزایش رابطه‌ی علی از طرح شبه کارآزمایی استفاده کند. از آنجایی که در این طرح کنترل کارآزمایی کامل وجود ندارد، پژوهشگر باید عوامل تهدیدکننده‌ی اعتبار پژوهش و متغیرهای کنترل نشده‌ای که ممکن است بر نتیجه اثر بگذارند را کاملاً مدنظر قرار داده و برایشان چاره‌ای بیاندیشد.

واژه	ترجمه انتخابی	تعریف
Randomization	تصادفی سازی	فرایند قرار دادن شرکت‌کنندگان در یکی از شاخه‌های کارآزمایی کنترل شده، به صورت تصادفی. این کار باعث می‌شود تا شانس قرارگیری شرکت‌کنندگان در هر یک از شاخه‌های مطالعه برابر و مستقل باشد. تصادفی سازی دو مؤلفه دارد: ایجاد یک توالی تصادفی و اجرای آن. در حالت ایده آل، شرکت‌کنندگان مطالعه نباید از توالی آگاه باشند (اختصاص به شاخه‌ها پنهان باشد).
Randomized controlled trial (RCT)	کارآزمایی کنترل شده تصادفی	یک مطالعه‌ی تجربی (کارآزمایی کنترل شده) که در آن شرکت‌کنندگان به صورت تصادفی در گروه‌های درمان (گروه‌های کارآزمایی و کنترل) قرار می‌گیرند.
Regression analysis	تحلیل رگرسیون	یک روش مدل سازی آماری که برای تخمین یا پیش بینی اثر یک یا چند متغیر مستقل بر یک متغیر وابسته استفاده می‌شود.
Residual confounding	مخدوش کنندگی باقی مانده	مخدوش شدن ناشی از متغیرهای اندازه‌گیری نشده در یک مطالعه
Restriction	محدود سازی	محدود سازی ورود به کوهورت به افرادی که دارای دامنه‌ی خاصی از مقادیر در یک عامل اختلاطی (مثلاً سن، نژاد و...) هستند. هدف این کار کاهش اثرعامل اختلاطی است.
Retrospective study	مطالعه‌ی گذشته‌نگر	مطالعه‌ای که در آن پژوهشگر پس از رخ دادن پیامدها، مواجهه‌ها را اندازه‌گیری می‌کند.
Standard error	خطای استاندارد	انحراف معیار میانگین‌های یک نمونه‌ی توزیع نظری از میانگین حقیقی جمعیت.
Structural equation modeling	مدل سازی معادلات ساختاری	مدل سازی معادلات ساختاری شامل طیف وسیعی از روش‌های تحلیل چند متغیری است که به منظور یافتن روابط متقابل بین متغیرهای مدل خطی بکار می‌روند. یافتن این روابط متقابل از طریق بررسی واریانس و کوواریانس متغیرها انجام می‌شود.
T-test	T آزمون	بررسی آماری دو میانگین جمعیت. آزمون T دونمونه‌ای به بررسی متفاوت بودن میانگین دو نمونه می‌پردازند. این آزمون معمولاً زمانی استفاده می‌شود که واریانس دو توزیع نرمال مشخص نباشد و کارآزمایی از نمونه‌ی کوچکی استفاده کند.
Type I error	خطای نوع اول	این خطای زمانی رخ می‌دهد که فرضیه‌ی صفر در حالیکه درست است، رد شود یا هنگام عدم وجود تفاوت، اعلام شود که تفاوتی وجود داشته است. این خطا همچنین خطای آلفا، مثبت کاذب و اُپیر کاذب نامیده می‌شود.
Type II error	خطای نوع دوم	این خطای زمانی رخ می‌دهد که فرضیه‌ی صفر غلط، رد نشود یا هنگام وجود تفاوت، اعلام شود که تفاوتی وجود نداشته است. این خطا همچنین خطای بتا، فرصت ازدست‌رفته و منفی کاذب نامیده می‌شود.
Variance	واریانس	یک معیار پراکندگی که توسط مجموعه‌ای از مشاهدات تعیین شده، و به صورت مجموع مربع‌های انحراف از معیار تقسیم بر تعداد درجات آزادی موجود در مجموعه‌ی مشاهدات، تعریف می‌شود.

