

فروشگاه و نمایشگاه کتاب  
دانشگاه علوم پزشکی مشهد  
(شعبه ۲)  
تلفن: ۳۸۸۴۱۶۱۹

ترجمه و تنظیم:

دکتر سعید اسلامی (دانشیار گروه انفورماتیک پزشکی/داروسازی بالینی)

مهندس شهنازی نیا (کارشناس ارشد انفورماتیک پزشکی)

دکتر ژیلا طاهرزاده (استادیار گروه فارماکودینامی و سم‌شناسی)

کلیه حقوق مادی و معنوی این اثر متعلق به دانشگاه علوم پزشکی مشهد است.

تابستان ۱۳۹۷

مقدمه	۳
انتخاب شرایط کنترل و مداخله	۳
انتخاب مداخله	۴
انتخاب کنترل	۶
انتخاب پیامدهای قابل اندازه گیری	۷
تعداد متغیرهای پیامد	۸
پیامدهای ترکیبی (مرکب)	۹
عوارض جانبی	۹
انتخاب شرکت کنندگان	۱۱
تعریف معیار ورود	۱۱
معیار خروج	۱۳
طراحی حجم نمونه و تعیین شرکت کنندگان بر اساس آن	۱۴
اندازه گیری متغیرهای پایه	۱۴
توصیف شرکت کنندگان	۱۵
اندازه گیری مقدار پایه از متغیر پیامد	۱۵
صرفه جویی	۱۶
بانک نمونه	۱۶
تصادفی سازی و پنهان سازی	۱۶
تصادفی سازی	۱۶
تکنیک های خاص تصادفی سازی	۱۸
پنهان سازی	۲۰
در صورتیکه پنهان سازی ممکن نباشد چه کاری باید انجام دهیم؟	۲۲
منابع	۲۳

در کارآزمایی بالینی، محقق مداخله‌ای را ایجاد می‌کند و سپس اثر آن را بر روی یک یا چند پیامد (outcome) مشاهده می‌کند. مزیت اصلی مطالعات کارآزمایی نسبت به مطالعات مشاهده‌ای توانایی در نشان دادن رابطه علت و معلولی (causality) می‌باشد. انتساب تصادفی مداخله، سبب کاهش اثر متغیرهای مخدوشگر می‌شود. پنهان کردن (blindness) اجرای آن، احتمال آنکه اثرات آشکار مداخله سبب اثبات یا قضاوت سوگیرانه (biased) در پیامد گردد را کاهش می‌دهد. با این وجود، بطور کلی کارآزمایی بالینی، گران و وقت گیر است، به یک سوال پاسخ می‌دهد و گاهی شرکت کنندگان در معرض آسیبهای بالقوه قرار می‌گیرند. به این دلایل، مطالعات کارآزمایی بهتر است جهت سوالات پژوهشی استفاده شوند، که مطالعات مشاهده‌ای و دیگر شواهد (evidence) نشان می‌دهند که مداخله احتمالاً موثر و بی-خطر است، اما قبل از آنکه بتواند مورد تایید و/یا توصیه قرار گیرد، شاهد قوی‌تر لازم است. کارآزمایی بالینی نمی‌تواند برای هر سوال پژوهشی استفاده شود. برای مثال آیا درمان دارویی کلسترول بالا در بچه‌ها سبب جلوگیری از حمله قلبی در سنین بالا می‌شود؟ بررسی این مطالعه امکان پذیر نمی‌باشد. یا به طور مثال قرار دادن تصادفی افراد در معرض دود واقعی سیگار جهت بررسی سرطان ریه در آنها، به لحاظ اخلاقی، قابل انجام نیست.

در ادامه به بررسی طراحی کارآزمایی پنهان شده تصادفی (randomized blinded trial) می‌پردازیم. انتخاب شرایط کنترل و مداخله، تعریف پیامدها و عوارض جانبی (adverse effects)، انتخاب شرکت کنندگان، اندازه‌گیری متغیرهای پایه و پیامد، ارزیابی رویکردهای تصادفی کردن و پنهان‌سازی را توضیح خواهیم داد.

### انتخاب شرایط کنترل و مداخله

طراحی کارآزمایی تصادفی به صورت موازی و بین دو گروه است. گروه مداخله، گروهی است که مداخله را دریافت می‌کند و گروه کنترل که یا هیچ درمان فعالی را دریافت نمی‌کند (استفاده از دارونما) یا آنکه درمان مقایسه‌ای دریافت می‌کند. محقق هر دو گروه را در طول زمان پیگیری می‌کند و پیامد را بین دو گروه مقایسه می‌نماید (تصویر ۱).

مشخص نبود که این نتیجه به دلیل اثر استروژن بوده است یا پروژسترون. به طور کلی، در طراحی کارآزمایی ترجیح داده می‌شود که بین دو گروه مورد مطالعه تنها یک اختلاف اصلی وجود داشته باشد.

محقق باید مواردی نظیر چگونگی ارائه مداخله به گروه دریافت کننده مداخله و چگونگی پنهان‌سازی مداخله را در نظر بگیرد. نکته حائز اهمیت دیگر، چگونگی اجرا مداخله در عمل می‌باشد. مداخلات ساده به مراتب نسبت به مداخلات پیچیده بهتر هستند (مثلا دریافت یک بار در روز یک نوع قرص از دو یا سه بار تزریق زیرجلدی یک دارو توسط بیماران ساده‌تر پذیرفته می‌شود). مداخلات پیچیده با جنبه‌های کیفی، مانند مشاوره‌های چندوجهی برای تغییر رفتار، ممکن است در عمل و در بالین قابل اجرا و بکارگیری نباشند. زیرا تکرار آنها دشوار، وقت گیر و پر هزینه است. چنین مداخلاتی حتی اگر در کارآزمایی اثبات شود که تاثیرگذار هستند، احتمال کمی وجود دارد که در سلامت جامعه تاثیرگذار باشند.

### انتخاب کنترل

بهترین گروه کنترل هیچ درمان فعالی را دریافت نمی‌کند البته به شکلی که بتواند پنهان‌سازی شود. انجام این مهم برای داروها نیازمند یک دارونما است که به صورت غیرقابل تشخیص از درمان فعال (پنهان‌سازی) دریافت شود. این استراتژی برای جبران هر نوع اثر دارونمایی مداخله (نظیر اثرات پیشنهادی یا انتظاری) است و سبب می‌شود که هر اختلاف در پیامد بین گروه‌های مورد مطالعه تنها با اثر خاص مداخله توضیح داده شود.

بهترین مقایسه بین گروه‌های مداخله و کنترل زمانی است که مداخلات همزمان (co-interventions) که می‌توانند خطر بروز پیامد مورد علاقه (outcome of interest) را تغییر دهند نظیر سایر مداخلات دارویی، درمانی یا رفتاری (به غیر از مداخله مورد مطالعه) وجود نداشته باشد. برای مثال در یک کارآزمایی تصادفی که به منظور ارزیابی مداخله یوگا در مقایسه با مراقبت معمولی برای جلوگیری از دیابت انجام می‌شود، مجریان مطالعه ممکن است شرکت‌کنندگان را به ورزش و کاهش وزن تشویق نمایند (مداخله جانبی). بنابراین وجود مداخلات همزمان ممکن است بر کاهش خطر پیشرفت دیابت تاثیرگذار باشد. اگر شرکت‌کنندگان در هر دو گروه از مداخلات همزمان تاثیر بپذیرند، نرخ پیامدها (rate of outcome) و قدرت مطالعه کاهش خواهد یافت و به حجم نمونه بزرگتر یا کارآزمایی بزرگتر نیاز خواهیم داشت. اگر استفاده موثر از مداخلات همزمان در بین دو گروه مداخله و کنترل متفاوت باشد، سبب ایجاد سوگیری (bias) در پیامد خواهد شد. در صورت عدم وجود پنهان‌سازی موثر، پروتکل باید حاوی اقداماتی جهت

بدست آوردن داده برای اصلاح آماری تفاوت بین دو گروه در میزان استفاده از مداخلات همزمان در طول کارآزمایی باشد. با این حال، اندازه‌گیری مداخلات همزمان مشکل است. اصلاح اختلافات پساتصادفی-سازی (postrandomization) بایستی به عنوان یک تحلیل ثانویه یا اکتشافی لحاظ شود زیرا در غیر این صورت قانون intention-to-treat نقض می‌گردد.

در اغلب موارد جلوگیری از درمان‌های دیگر به غیر از مداخله اصلی امکان پذیر نمی‌باشد. برای مثال در اجرای یک کارآزمایی جهت بررسی یک داروی جدید برای کاهش خطر ابتلا به سکته قلبی در افراد مبتلا به بیماری عروق کرونر قلب (CHD)، محقق به لحاظ اخلاقی نمی‌تواند شرکت‌کنندگان را از مصرف داروهای درمانی که برای افراد مبتلا به CHD شناخته شده می‌باشند، منع یا دلسرد نماید. یک راه حل آن است که داروهای استاندارد به همه شرکت‌کنندگان در کارآزمایی تجویز شود. البته این رویکرد سبب می‌شود تا نرخ پیامد کاهش یافته و نیاز به حجم نمونه بزرگتر باشد. این گونه آزمایشات بهتر است برای سوالات پژوهشی استفاده شوند که می‌خواهند بررسی کنند آیا مداخله جدید در کنار درمان‌های دیگر سبب بهبود پیامد می‌شود یا خیر؟

در کارآزمایی‌هایی که قرار است مداخله در کنار درمان‌هایی شناخته شده و موثر بکار برود، کارآزمایی بهتر است به صورت non-inferiority یا به صورت هم ارزی (equivalence) طراحی شود. در این حالت درمان جدید با درمانی مقایسه می‌شود که اثر آن اثبات شده می‌باشد.

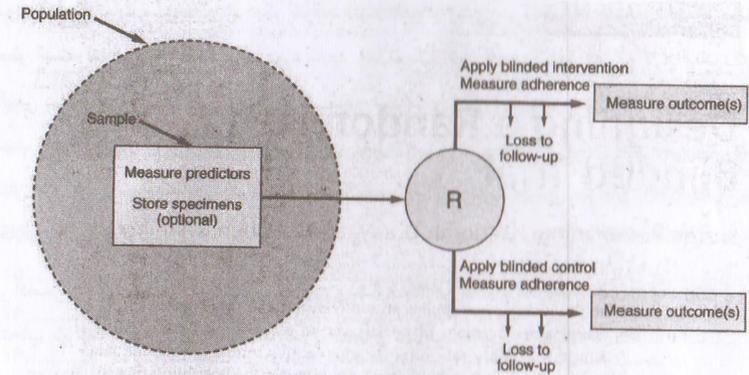
### انتخاب پیامدهای قابل اندازه‌گیری

انتخاب و تعریف پیامدهای خاص کارآزمایی، روی مولفه‌های دیگر مطالعه از جمله هزینه و امکان پذیر بودن آن تاثیر می‌گذارد. کارآزمایی معمولاً باید شامل چند پیامد جهت افزایش غنای نتایج و تجزیه و تحلیل ثانویه باشد. با این حال، تنها یکی از پیامدها به عنوان پیامد اصلی در نظر گرفته می‌شود که پاسخ دهنده سوال اصلی تحقیق می‌باشد. محاسبه حجم نمونه و تعیین اولویتها در اجرای مطالعه بر اساس پیامد اصلی صورت می‌گیرد.

**پیامدهای بالینی**، بهترین شواهد برای چگونه استفاده کردن درمان‌ها یا مداخلات پیشگیرانه می‌باشند. با اینحال برای پیامدهایی که نادر می‌باشند، مانند بروز سرطان، انجام کارآزمایی وسیع، طولانی و گران لازم است. در مقایسه با پیامدهای دوتایی (binary) با پیامدهایی که به صورت متغیرهای پیوسته اندازه‌گیری

## انتخاب مداخله

اولین مرحله در طراحی کارآزمایی بالینی، انتخاب مداخله می‌باشد. محقق باید چندین مساله را در نظر بگیرد که شامل: میزان استفاده، مدت زمان، و تکرار مداخله می‌باشد، بطوریکه بهترین توازن بین کارایی (efficacy) و ایمنی (safety) را ایجاد نماید. همچنین در نظر گرفتن نکات زیر نیز حائز اهمیت می‌باشد: امکان پنهان‌سازی، درمان با یک یا ترکیبی از مداخله‌ها، قابلیت پذیرش شرکت‌کنندگان و قابلیت تعمیم - پذیری روش درمانی که عملاً استفاده خواهد شد. اگر تصمیم‌گیری در مورد هر یک از موارد فوق نامشخص باشد (مثلاً تعیین آنکه کدام دوز دارو بهترین توازن را بین کارایی و ایمنی ایجاد می‌کند) بهتر است تا زمان کامل شدن مطالعات مقدماتی جهت حل این مسایل، انجام کارآزمایی به تعویق افتد.



تصویر ۱: مراحل طراحی کارآزمایی پنهان شده تصادفی عبارتند از:

- انتخاب نمونه‌ای از جمعیت مناسب برای دریافت مداخله
- اندازه‌گیری متغیرهای پیشگویی و در صورت وجود سطح پایه‌ای از متغیر پیامد.
- در نظر گرفتن چگونگی ذخیره سازی نمونه‌ها برای تحلیلهای بعدی
- انتصاب تصادفی پنهان شده در گروههای کنترل و مداخله
- پیگیری دو گروه در طول زمان، کاهش از دست دادن نمونه در طول پیگیری و ارزیابی منطبق با مداخله و کنترل
- اندازه‌گیری متغیرهای پیامد

بهترین توازن بین کارایی و ایمنی، بستگی به مداخله و شرط مطالعه شده دارد. عموماً در مداخلاتی که جهت درمان بیماری‌های همراه با علائم شدید یا مرگ بکار می‌روند، کارآیی بیشتر مورد توجه قرار می‌گیرد. برای مثال ممکن است بالاترین دوز قابل تحمل برای درمان سرطان متاستاتیک، بهترین انتخاب باشد. از سوی دیگر، ایمنی باید معیار اصلی طراحی مداخلاتی باشد که برای درمان علائمی بکار می‌روند که به ندرت سبب پیشرفت بیماری یا مرگ می‌شوند. برای حفظ سلامت مردم باید آزمونهای دقیق ایمنی در درمان‌های پیشگیرانه انجام شود. اگر درمان موثر هم باشد اما هر فرد تحت درمان در معرض خطر عوارض جانبی درمان نیز می‌باشد. در اینگونه موارد بهتر است دوز دارو به اندازه‌ای انتخاب شود که بیشترین کارآیی همراه با کمترین احتمال بروز عوارض جانبی را داشته باشد. اگر بهترین دوز بر اساس آزمایشات حیوانی و یافته‌های تحقیقاتی انسانی مشخص نیست، ممکن است نیاز به آزمایشات اضافی برای مقایسه اثرات دوزهای متعدد بر روی نشانگرها واسط (intermediate marker) یا پیامدهای بالینی باشد.

گاهی اوقات محقق تصمیم می‌گیرد چندین [گروه مداخله] (چند دوز یا ...) با یک گروه کنترل مقایسه شوند. به عنوان مثال، در یک طراحی کارآزمایی، از آنجا که مشخص نیست کدام دوز رالوکسیفن (۶۰ یا ۱۲۰ میلی گرم) برای جلوگیری از شکستگی مهره موثرتر است همزمان چند نتیجه از ارزیابی داروی رالوکسیفن استنتاج شده است. در این کارآزمایی هر دو دوز برای جلوگیری از شکستگی مهره مورد آزمون قرار می‌گیرد. این روش یک استراتژی مستدل ولی پرهزینه است چراکه این کارآزمایی به علت داشتن چندین فرضیه پیچیده‌تر می‌گردد.

در بعضی درمانها، دوز دارو تا رسیدن به یک اثر بهینه برای هر بیمار مرتباً باز تنظیم می‌شود مانند کاهش بار ویروسی در هپاتیت C. در این موارد ممکن است بهتر باشد که در مطالعه نیز مداخله به گونه‌ای طراحی شود که دوز دارو برای رسیدن به یک نتیجه بالینی مرتباً باز تنظیم شود. برای حفظ پنهان‌سازی، بایستی تغییرات دوز دارو در گروه دریافت کننده دارونما نیز به صورت متناظر با گروه مداخله انجام شود.

طراحی و اجرا کارآزمایی برای آزمودن یک مداخله درمانی به مراتب راحت تر از آزمودن ترکیبی از درمانها می‌باشد. با این حال در بسیاری از موارد مانند عفونت HIV و یا نارسایی احتقانی قلب، درمان به صورت ترکیبی از داروها و/یا درمان‌های غیردارویی انجام می‌شود. مهمترین عیب آزمون ترکیبی آنست که با ارزیابی پیامد، نتیجه گیری شفافی درباره نقش هر یک از عناصر مداخله نمیتوان داشت. به عنوان مثال در یکی از مطالعات برنامه سلامت زنان، زنان یائسه به وسیله استروژن به همراه پروژسترون و دارونما تحت درمان قرار گرفتند. نتایج نشان دهنده آن بود که خطر سرطان پستان افزایش یافته است؛ با این حال

می‌شوند، نظیر کیفیت زندگی مطالعه می‌تواند با تعداد شرکت‌کنندگان کمتری انجام شود. با اینحال پیامدهای بالینی خیلی مهم، نظیر عود سرطان، بطور غیرقابل اجتناب فقط می‌توانند دوتایی باشند.

**نشانه‌گرهای واسط،** نظیر تراکم استخوان، اندازه‌گیری‌هایی هستند که با پیامدهای بالینی مرتبط می‌باشند. کارآزمایی‌هایی که از پیامدهای واسط استفاده می‌کنند، می‌توانند درک ما را از پاتوفیزیولوژی بیشتر نموده و اطلاعاتی را برای انتخاب بهترین دوز دارو یا میزان تکرار درمان در کارآزمایی‌های با پیامدهای بالینی فراهم نمایند. ارتباط بالینی کارآزمایی‌ها با پیامدهای واسط به چگونگی و دقت تغییرات در این نشان‌گرها بستگی دارد. به ویژه تغییراتی که به دلیل درمان اتفاق می‌افتد نشان دهنده تغییرات در ریسک پیامدهای بالینی می‌باشد. نشانگر واسط می‌تواند به صورت نشانگر جانشین (surrogate marker) برای پیامدهای بالینی در نظر گرفته شود بطوریکه تغییرات ناشی از درمان در نشانگر واسط به طور پیوسته چگونگی تغییرات ناشی از درمان را در پیامد بالینی پیش‌بینی نماید. بطور کلی یک نشانگر جانشین خوب، تغییرات در یک فاکتور واسط در مسیر اصلی که منجر به پیامد بالینی می‌شود را اندازه‌گیری می‌کند.

بار ویروسی HIV یک نشانگر جایگزین خوب است. زیرا درمان‌هایی که بطور مداوم سبب کاهش بار ویروسی می‌شود، مرگ و میر بیماران مبتلا به HIV را کاهش خواهد داد. در مقابل تراکم مواد معدنی استخوان (BMD) یک نشانگر جایگزین ضعیف است. زیرا نشان‌دهنده مقدار مواد معدنی در یک بخش از استخوان است و درمان‌هایی که سبب بهبود BMD می‌شوند گاهی اوقات اثر ناچیزی بر خطر شکستگی دارند. مقدار افزایش مداوم تراکم استخوان با کاهش خطر شکستگی ناشی از درمان مرتبط نیست.

#### تعداد متغیرهای پیامد

وجود چند متغیر پیامد برای ارزیابی جنبه‌های مختلفی از یک پدیده مورد علاقه مطلوب می‌باشد. برای مثال اگر در مطالعه جایگزینی استروژن/پروژسترون و قلب، رویدادهای بیماری کرونری قلب به عنوان نقطه پایانی اصلی (primary endpoint) انتخاب شده باشند، بستری برای آنژین ناپایدار یا نارسایی احتقانی قلب، سکته، حمله ایسکیمیک گذرا، حوادث ترومبوآمبولیک وریدی و تمام مرگ‌ومیرها با علل غیرقطعی می‌توانند شرحی از تاثیرات هورمون درمانی روی قلب باشد. با این وجود، تنها از پیامد اصلی به منظور تعیین حجم نمونه و طول مدت مطالعه و همچنین برای جلوگیری از مشکلات ناشی از "تفسیر آزمایش با فرضیه‌های متعدد" استفاده می‌شود.

#### پیامدهای ترکیبی (مرکب)

در بعضی از کارآزمایی‌ها پیامدها به صورت ترکیب شده چند رویداد یا اندازه‌گیری مختلف تعریف می‌شوند. برای مثال، در بسیاری از کارآزمایی‌ها، کاهش ریسک بیماری کرونری قلب چندین رویداد خاص کرونری نظیر انفارکتوس میوکارد، مرگ و پروسیجرهای عروق کرونر را شامل می‌شود. پیامدهای ترکیبی زمانی معقول است که: هر یک از پیامدها به لحاظ بالینی مهم باشند، درمان بوسیله مکانیسم‌های مشابه برای هر رویداد انجام شود و انتظار داشته باشیم که مداخله ریسک هر رویداد را کاهش دهد. علاوه بر این، یک پیامد ترکیبی قدرت بیشتری را نسبت به پیامد تنها دارد زیرا رویدادهای آن بیشتر خواهد بود. با این وجود، اگر پیامدهای ترکیبی، رویدادهای ارزشمند بالینی (clinically meaningful) یا رویدادهایی که بیشتر از دیگر رویدادها اتفاق می‌افتند را شامل نشود سبب ایجاد یافته‌های گمراه‌کننده خواهد شد.

برای مثال، اگر بستری برای ارزیابی درد سینه به پیامد کرونری ترکیبی اضافه شود، در صورتیکه بستری برای درد سینه بیشتر از مرگ کرونری و انفارکتوس میوکارد اتفاق بیفتد، این رویداد بر رویداد ترکیبی برتری خواهد یافت [در حالیکه ارزش بالینی کمی دارد]. در یک مداخله که پیامد ترکیبی اینچنین تغییر پیدا کرده است، ممکن است کاهش ریسک رویدادهای کرونری گزارش داده شود، در حالیکه این مداخله تنها ریسک بستری شدن برای درد قفسه سینه را کاهش داده است.

پیامدهای ترکیبی باید با دقت زیادی طراحی شوند. اگر درمان اثر کمی روی یک پیامد ایجاد نماید، مخصوصاً اگر این پیامد به نسبت رایج باشد، ممکن است قدرت آماری کمی را اضافه نماید یا حتی حجم نمونه مورد نیاز برای شناسایی یک اثر را افزایش دهد. برای مثال، اگر سکتی به پیامد ترکیبی قلبی عروقی اضافه شود، مداخله ممکن است خطر رویدادهای کرونری را کاهش دهد، اما روی ریسک سکتی هیچ تاثیری نداشته باشد یا حتی آن را افزایش دهد و در اینصورت نتیجه‌گیری شود که مداخله هیچ تاثیری روی پیامد ترکیبی قلبی عروقی ندارد.

#### عوارض جانبی

محقق باید پیامدهای قابل اندازه‌گیری جهت شناسایی عوارض جانبی ناشی از مداخله را در مطالعه وارد نماید. پاسخ به این سوال که آیا اثرات سودمند مداخله بیشتر از عوارض جانبی آن می‌باشد یکی از اهداف اصلی اغلب کارآزمایی‌های بالینی است. این هدف برای مداخلات کاملاً بی‌ضرر مانند یک برنامه آموزشی سلامت نیز باید در نظر گرفته شود. دامنه عوارض جانبی از علائم خفیف تا عوارض جدی و کشنده را در بر

می‌گیرد. نرخ وقوع، اثر درمان و حجم نمونه برای تشخیص عوارض جانبی کاملاً متفاوت از تشخیص مزایا است. متأسفانه تشخیص عوارض جانبی نادر حتی با کارآزمایی‌های بزرگ امکان‌پذیر نیست و تنها هنگام بکارگیری گسترده مداخله و با مطالعات مشاهده‌ای بزرگ یا گزارش‌های موردی قابل شناسایی می‌باشند.

در مراحل اولیه آزمون یک درمان جدید، زمانیکه عوارض جانبی بالقوه نامشخص هستند، محقق باید به طور گسترده سوالات بازی (open-ended) درباره همه انواع عوارض جانبی بپرسد. در کارآزمایی‌های بزرگ ارزیابی و کد کردن تمام رویدادهای جانبی بالقوه خیلی گران و وقت‌گیر است و معمولاً نتایج مهم کمی از آن حاصل می‌شود. محقق باید جهت مدیریت حجم کار و حفظ ارزیابی کافی از ضررهای بالقوه مداخله استراتژی‌هایی مناسب را در نظر بگیرد. برای مثال، در کارآزمایی‌های خیلی بزرگ، رویدادهای جزئی و رایج نظیر عفونت دستگاه تنفسی فوقانی و اختلالات دستگاه گوارش می‌تواند در یک زیرگروه از شرکت‌کنندگان [ارزیابی] و ثبت گردد. ممکن است به ثبت عوارض جانبی که جدی نیستند و مطالعات قبلی تفاوتی در بروز این علائم جزئی پیدا نکرده باشند نیازی نباشد. علاوه بر سوالات باز، برای کشف رویدادهای جانبی مهم که براساس تحقیقات قبلی یا تجربیات بالینی مورد انتظار می‌باشد، باید ارزیابی‌های ویژه طراحی شود. برای مثال چون myositis یک عارضه جانبی گزارش شده از درمان با استاتین است، علائم و نشانه‌های myositis باید در هر کارآزمایی از استاتین جدید بررسی شود.

عوارض جانبی که به عنوان علائم یا اصطلاحات بالینی گزارش می‌شوند، باید برای تحلیل به صورت کد شده و سازماندهی شده درآیند. MedDRA و SNOMED به عنوان لغت نامه اصطلاحی رایج استفاده می‌شوند که به چند صورت مختلف مانند علائم، تشخیص‌های خاص و مطابق با ارگانهای بدن گروه‌بندی شده‌اند. به عنوان مثال یک رویداد جانبی به صورت "تب و سرفه" و رویداد جانبی دیگری به صورت "برونشیت" ثبت شده است که با معیارهای دیگر مانند سینه پهلو در مجموع تحت عنوان "عفونت تنفسی" و در یک سطح بالاتر تحت عنوان "عارضه جانبی در سیستم تنفسی" گروه بندی خواهد شد. اینگونه طرح‌های طبقه‌بندی، خلاصه کلی خوبی را از عوارض جانبی فراهم می‌کنند و به طور منطقی برای بیماری‌هایی مانند شکستگی که به طور خاص تشخیص داده می‌شوند، صحیح می‌باشد. با این وجود در رابطه با رویدادهای جانبی مهمی که با چندین اصطلاح شرح داده می‌شوند، اگر این اصطلاحات با همدیگر گروه بندی نشده باشند، امکان از دست رفتن آنها وجود دارد. برای مثال، در یک کارآزمایی از denosumab برای پیشگیری از شکستگی استئوپرویک، MedDRA مواردی از cellulitis را بطور مجزا از موارد erysipelas کد می‌کند (دو نام برای انواع عفونت‌های یکسان). وقتی ترکیب می‌شوند ۱۲ مورد

جدی از cellulitis با erysipelas اتفاق افتاده است در مقابل ۱ موردی که با دارونما اتفاق افتاده است ( $p < 0.001$ ).

عوارض جانبی بر اساس میزان شدت طبقه‌بندی می‌شوند. حادثه مرگ یا تهدید کننده زندگی، حوادث نیاز به بستری شدن طولانی، معلولیت یا آسیب دائمی، نقص‌های مادرزادی، و سایر حوادث مهم پزشکی که نیاز به مداخله دارویی یا جراحی برای جلوگیری از پیامدهای دیگر دارند، به عنوان عوارض جانبی جدی (Serious adverse events) تعریف می‌شوند. عوارض جانبی خیلی جدی باید سریعاً به مسئول کارآزمایی گزارش داده شوند. وقتی کارآزمایی یک دارو جدید تصویب نهایی می‌شود، طراح کارآزمایی موظف است عوارض جانبی را گزارش نماید. برای برخی بیماری‌های خاص، مانند سرطان، روشهایی برای طبقه بندی کردن رویدادهای جانبی ایجاد شده است.

## انتخاب شرکت کنندگان

### تعریف معیار ورود

در یک کارآزمایی بالینی، هدف معیارهای ورود (inclusion) و خروج (exclusion) شناسایی و انتخاب یک جمعیت است که کارآزمایی در آنها قابل انجام، اخلاقی و مرتبط با تاثیر مداخله بر روی پیامد [انتخابی] باشد. معیار ورود باید تعداد مناسبی از شرکت‌کنندگان را که نرخ بالا و کافی از پیامد اصلی را دارند، ایجاد نماید. این امر سبب می‌شود تا مطالعه برای پیدا کردن یک تاثیر مهم از مداخله بر روی پیامد، قدرت کافی را داشته باشد. از سوی دیگر، معیار ورود باید قابلیت تعمیم یافته‌های حاصل از کارآزمایی را به حداکثر برساند.

برای مثال، اگر پیامد یک رویداد نادر است، نظیر سرطان سینه، استفاده از شرکت‌کنندگان در معرض خطر ضروری است. با این کار حجم نمونه و مدت زمان پیگیری را به سطوح امکان‌پذیر کاهش می‌دهیم. از سوی دیگر، انتخاب معیار ورود به زنان در معرض خطر قابلیت تعمیم نتایج را محدود می‌کند. اگر بخواهیم این مشکل برطرف شود باید تعداد شرکت‌کنندگان را افزایش دهیم.

برای انتخاب حجم نمونه صحیح، محقق باید یک تخمین قابل اعتماد از نرخ پیامد اصلی در افراد شرکت‌کننده داشته باشد. تخمین می‌تواند بر پایه اطلاعات آماری، مطالعات مشاهده‌ای طولی (longitudinal study)، یا نرخ مشاهده شده در گروه کنترل در کارآزمایی‌های با معیار ورود مشابه با کارآزمایی طراحی شده بدست آید.

برای مثال، نرخ مورد انتظار سرطان پانکراس در بزرگسالان را از داده های رجیستری ثبت سرطان می توان تخمین زد. محقق باید این نکته را در نظر بگیرد که نرخ رویداد، در میان افرادی که مورد غربالگری قرار گرفته و سالم هستند و در مطالعه وارد می شوند کمتر از جمعیت کلی می باشد. بنابراین ممکن است ترجیح داده شود که نرخ سرطان پانکراس از گروه کنترل مطالعات کارآزمایی دیگر با معیار ورود مشابه به دست آید.

اگر افراد در معرض خطر پیامد در مطالعه وارد شوند، تعداد شرکت کنندگان وارد شده در کارآزمایی می تواند کاهش یابد. اگر فاکتورهای خطر برای پیامد شناخته شده باشد، معیار انتخاب می تواند به گونه ای طراحی شود که افرادی که یک حداقلی از ریسک تخمینی را دارند در مطالعه وارد شوند.

در یک کارآزمایی، به بررسی اثر رالوکسیفن برای جلوگیری از بیماریهای قلبی و سرطان پستان پرداخته می شود. افرادی که در مطالعه وارد می شوند، زنان در معرض خطر بیماری قلبی می باشند. یک راه دیگر برای افزایش نرخ رویدادها آن است که افرادی که بیماری قلبی دارند در مطالعه وارد شوند. در مطالعه "قلب و جایگزینی استروژن/پروژسترون"، افرادی که بیماری قلبی داشتند در مطالعه وارد شده سپس به بررسی اثر استروژن و پروژسترون در کاهش رویدادهای قلبی جدید پرداخته شده است، حجم نمونه در این حالت ۲۷۶۳ نفر بود. این رویکرد نسبت به انتخاب زنان سالم، هزینه کمتری دارد. حجم نمونه با انتخاب افراد سالم به ۱۷۰۰۰ نفر می رسد.

اگرچه نمونه های احتمالی از جمعیت کلی، برای مطالعات مشاهده ای مزیت محسوب می شود، اما این روش نمونه گیری در مطالعات کارآزمایی تصادفی ضروری و شدنی نیست. ورود افراد با تنوع در مشخصات به کارآزمایی اطمینان برای بکارگیری وسیع نتایج را افزایش می دهد. با اینحال اگر اختلافات ژنتیکی و بیولوژیکی بین جمعیت ها روی درمان تاثیر نداشته باشد، معمولا نتایج کارآزمایی روی یک نمونه در دسترس (مثلا زنانی که بیماری قلبی دارند و برای ورود به مطالعه رضایت دارند) با نتایج یک نمونه احتمالی از جمعیت واجد شرایط (همه زنان که بیماری قلبی دارند) مشابه خواهد بود. گاهی اوقات کارآزمایی درمان به مشخصه های افراد وابستگی دارد: که اصطلاحا، اصلاح اثر یا تداخل (effect modification or interaction) نامیده می شود. برای مثال، برخی از درمانهای پوکی استخوان، خطر شکستگی را در زنان با تراکم استخوانی بسیار پایین، به طور قابل ملاحظه ای کاهش می دهد (T-score زیر ۲,۵-) و روی زنان با تراکم استخوانی بالاتر بدون تاثیر بوده و یا اثر ناچیز دارد (P=0.02 برای تداخل). در این مورد، تنها زنان با تراکم استخوانی کم در مطالعه وارد می شوند که در اینصورت میزان تاثیر (effect size) افزایش و حجم نمونه کاهش می یابد.

**طبقه بندی شرکت کنندگان** براساس یک ویژگی، نظیر گروه نژادی، به محقق این امکان را می دهد تا براساس مشخصه هایی که روی تاثیر درمان یا قابلیت تعمیم پذیری آن موثر است، به انتخاب صحیح افراد بپردازد. وقتی هدف تعیین شده برای شرکت کنندگان با یک خصوصیت تامین شده باشد ورود شرکت کنندگان به آن طبقه می تواند متوقف شود. با این وجود، به دلیل آنکه اغلب کارآزمایی ها جهت آزمون ناهمگنی اثر مداخله در چنین زیرگروه هایی با حجم نمونه مناسبی طراحی نمی شوند، این استراتژی ممکن است ارزش عملی محدودی داشته باشد.

### معیار خروج

معیار خروج باید درست و به صرفه باشد زیرا خروج غیرضروری سبب مشکل رسیدن به حجم نمونه لازم، افزایش پیچیدگی و هزینه می گردد. پنج دلیل اصلی برای خروج افراد از کارآزمایی وجود دارد (جدول ۱)

جدول ۱. دلایل خروج افراد از یک مطالعه کارآزمایی بالینی

مثال: کارآزمایی رالوکسیفن درمقابل دارونما جهت پیشگیری از بیماری قلبی	دلایل
واقعه ترومبوآمبولی وریدی قلبی (رالوکسیفن خطر ابتلا به ترومبوآمبولی وریدی را افزایش می دهد). سرطان پستان اخیر با گیرنده استروژن مثبت (درمان با مدولاتور انتخابی گیرنده استروژن موثر و درمان استاندارد است).	درمان مورد مطالعه ممکن است مضر باشد. • اگر ریسک غیرقابل قبولی از ضرر به گروه مداخله اختصاص داده شود. • اگر ریسک غیرقابل قبولی از ضرر به گروه کنترل اختصاص داده شود.
زنان نوجوان با خطر بسیار کم بیماریهای قلبی بیمار مبتلا به بیماریهای دریچه قلب، که به احتمال زیاد به علت درمان استروژنی نمی تواند به فرضیه اثرات رالوکسیفن پاسخ دهد. استروژن درمانی (که با رالوکسیفن رقابت می کند)	درمان فعال بعید است موثر باشد. • ریسک کم برای رسیدن به پیامد • داشتن نوعی بیماری که پاسخ به درمان را غیرممکن می کند. • گرفتن درمانی که با مداخله تداخل دارد.
پایبندی ضعیف در طول دوره اجرا	پایبند نبودن به مداخله
قبل از پایان کارآزمایی باید منتقل شود و نمی تواند برای ارزیابی پایانی پیامد در دسترس باشد. امید به زندگی کوتاه به دلیل داشتن یک بیماری جدی	کامل انجام ندادن پیگیری (پایش های دوره ای)
اختلالات روانی که سبب جلوگیری از پاسخ دقیق به سوالات گردد	مشکلات عملی با شرکت کردن در پروتکل

افرادی که بطور بالقوه برای گروه کنترل و مداخله نامناسب هستند باید از مطالعه خارج گردند. درمان ممکن است برای افرادی که مستعد یا مظنون به عوارض جانبی هستند، نامناسب باشد. به عنوان مثال

انفارکتوس میوکارد یک عارضه جانبی نادر از درمان با سیلدنافیل است، بنابراین در بررسی این دارو برای درمان اسپاسم در بیماران مبتلا به بیماری رینود، باید بیماران مبتلا به بیماری قلبی از مطالعه حذف گردند. در مقابل، انتساب برخی افراد در گروه دریافت کننده دارونما نیز می‌تواند نامناسب باشد. برای مثال، در زنان مبتلا به شکستگی مهره، بیس فسفونات‌ها برای کاهش خطر ابتلا به شکستگی‌های بعدی، شناخته شده است. پس وارد کردن این افراد در گروه کنترل برای یک درمان جدید در پوکی استخوان قابل قبول نیست مگر آنکه بیس فسفونات برای افراد هر دو گروه فراهم شود. افرادی که در آنها موثر بودن درمان فعال، پایبند بودن آنها به مداخله یا کامل انجام شدن پیگیری ایشان بعید می‌باشد، باید از مطالعه کنار گذاشته شوند. گاهی نیز مشکلاتی از قبیل اختلالات ذهنی باعث می‌شود تا مشکلاتی در پیگیری دستورات ایجاد شود که در اینصورت این افراد از مطالعه خارج می‌شوند. محقق باید به دقت معیار خروجی را تعیین نماید زیرا ممکن است تاثیر بزرگی روی امکان پذیر بودن و هزینه کارآزمایی و تعمیم‌پذیری نتایج داشته باشد.

#### طراحی حجم نمونه و تعیین شرکت کنندگان بر اساس آن

طراحی کارآزمایی با حجم نمونه خیلی کم برای تشخیص اثرات مهم، بی‌فایده و غیراخلاقی است و ممکن است سبب ایجاد نتیجه‌گیری گمراه کننده شوند. یکی از مهمترین بخشها در طراحی کارآزمایی، برآورد حجم نمونه می‌باشد. در تعیین حجم نمونه باید به این نکته توجه نمود که نرخ پیامد در کارآزمایی بالینی نسبت به مقدار برآورد شده کمتر می‌باشد. این امر به دلیل ایجاد سوگیری است که با حضور افراد سالم در مطالعه به وجود می‌آید. علاوه بر این، جلب مشارکت افراد در مطالعات کارآزمایی نسبت به مطالعات مشاهده‌ای مشکل‌تر است چراکه شرکت کنندگان در کارآزمایی باید برای تصادفی‌سازی رضایت داشته باشند. بنابراین، محقق بایستی برای نمونه‌های زیاد از یک جمعیت بزرگ قابل دسترس، برنامه ریزی نماید. همچنین هنگامیکه موانع انجام کار بیشتر از حد انتظار است، هزینه و زمان کافی برای رسیدن به حجم نمونه مطلوب را داشته باشد.

#### اندازه‌گیری متغیرهای پایه

به منظور تسهیل در تماس با شرکت کنندگان و برای آنکه آنها را در طول پیگیری از دست ندهیم، ثبت کردن نام، شماره تلفن، آدرس و آدرس پست الکترونیکی از دو یا سه دوست یا خویشاوند که همیشه به افراد دسترسی دارند، مهم است. در صورت امکان ثبت شماره ملی برای تعیین وضعیت حیاتی شرکت کنندگان (از طریق فهرست مرگ ملی) و یا تشخیص نتایج کلیدی با استفاده از پرونده سلامت (به عنوان

مثال، سیستم بیمه درمانی) بسیار ارزشمند می‌باشد. اطلاعات سلامت افراد باید محرمانه نگه داشته شود و نباید برای موسسه یا فردی دیگر ارسال گردد.

#### توصیف شرکت کنندگان

محققان باید اطلاعات مربوط به عوامل خطر (قطعی/بالقوه) بر پیامد و ویژگی‌های شرکت کنندگان که ممکن است بر کارایی مداخله تاثیرگذار باشد، جمع‌آوری نمایند. این اندازه‌گیری‌ها قابلیت مقایسه پذیری دو گروه تصادفی‌سازی شده در شروع مطالعه را ارزیابی و اطلاعات کافی جهت ارزیابی تعمیم پذیری یافته‌ها را فراهم می‌کنند. هدف از توصیف شرکت کنندگان در شروع مطالعه اطمینان از آن است که اختلاف در مشخصه‌های پایه بین دو گروه بیشتر از اختلاف مورد انتظار ناشی از شانس، خطاهای تکنیکی یا بایاس در انجام تصادفی‌سازی نخواهد بود. در کارآزمایی‌های کوچک که بصورت شانس مشخصه‌های پایه به سمت توزیع ناعادلانه بین گروه‌های تصادفی‌سازی شده متمایل می‌باشند، اندازه‌گیری متغیرهای مهم پیش‌بینی پیامد، امکان اصلاح آماری (statistical adjustment) در مقایسه گروه‌های تصادفی شده را به منظور کاهش اثر توزیع ناعادلانه فراهم می‌کند. همچنین اندازه‌گیری متغیرهای پیش‌بینی پیامد، این امکان را برای محقق فراهم می‌کند که بتواند وجود اثر متفاوت مداخله در زیر گروه‌های طبقه بندی شده بر اساس متغیرهای پایه را بررسی نماید (اصلاح اثر یا effect modification).

#### اندازه‌گیری مقدار پایه از متغیر پیامد

اگر پیامدهای مورد ارزیابی شامل تغییر در مقادیر آنها باشد، متغیر پیامد بایستی در شروع مطالعه و به همان روش در پایان مطالعه اندازه‌گیری شود [تا تغییر در آنها قابل محاسبه باشد]. در مطالعاتی که متغیرهای پیامد پیوسته وجود دارد (اثرات درمان شناختی-رفتاری بر رتبه‌های افسردگی)، بهترین اندازه‌گیری، میزان تغییر در پیامد در ابتدا و انتهای کارآزمایی است. این روش معمولاً سبب می‌شود تا پراکندگی در پیامد بین افراد مورد مطالعه به حداقل برسد و نسبت به مقایسه ساده مقادیر در پایان کارآزمایی قدرت مطالعه را بیشتر می‌کند. در مطالعاتی با پیامدهای دوگانه (برای مثال، بروز CHD) ممکن است اثبات آنکه بیماری در ابتدا وجود نداشته (برای مثال با بررسی سوابق با گرفتن نوار قلب) مهم باشد. همچنین اندازه‌گیری متغیرهای پیامد ثانویه و پیامدهای مطالعات جانبی از پیش برنامه‌ریزی شده (ancillary studies) در ابتدا مطالعه می‌تواند مفید باشد.

## صرفه جویی

اگرچه به استفاده‌های چندگانه از اندازه‌گیری‌های پایه اشاره شد، اما باید تاکید کرد که در طراحی یک کارآزمایی بالینی نیازی نیست که هرچیزی را اندازه‌گیری نمود. زیرا تصادفی‌سازی مشکل متغیرهای مخدوشگر (confounders) را به حداقل می‌رساند. افزایش تعداد اندازه‌گیری سبب افزایش پیچیدگی و هزینه خواهد شد. در یک کارآزمایی تصادفی شده با بودجه و زمان محدود، معمولاً بهتر است روی موارد اساسی مانند کافی بودن حجم نمونه، موفقیت در تصادفی‌سازی و پنهان‌سازی و پایندی کامل به پیگیری وقت صرف گردد.

## بانک نمونه

ذخیره‌سازی تصاویر، سرمها، DNA و ... امکان اندازه‌گیری بعدی تغییرات ناشی از درمان، نشانگرهایی که پیامد را پیش‌بینی می‌کنند و فاکتورهایی مانند ژنوتایپ که سبب تعریف افرادی که پاسخ خوب یا ضعیف به درمان می‌دهند را فراهم می‌کند. نمونه‌های ذخیره شده می‌توانند برای مطالعه سوالات پژوهشی دیگر، که مستقیماً به پیامد اصلی مورد مطالعه مربوط نیست، یک منبع غنی باشند.

## تصادفی‌سازی و پنهان‌سازی

چهارمین مرحله در تصویر ۱، انتصاب تصادفی افراد در دو گروه می‌باشد. در یک طراحی ساده، یک گروه درمان فعال یا مداخله و گروه دیگر دارونما را دریافت می‌کند. انتصاب تصادفی اطمینان می‌دهد که به جز پراکندگی ناشی از شانس سایر مخدوشگرها شامل سن، جنس و دیگر مشخصه‌های پایه که ممکن است یک ارتباط مشاهده شده را مخدوش نمایند (حتی آنهایی که ناشناخته‌اند و یا اندازه‌گیری نشده‌اند) در ابتدا کار به صورت یکسانی بین گروه‌های تصادفی شده توزیع خواهند شد. پنهان‌سازی جهت حفظ مقایسه‌پذیری گروه‌های مورد مطالعه در طول کارآزمایی و اطمینان از نتایج بی‌طرفانه مهم است.

## تصادفی‌سازی

از آنجا که تصادفی‌سازی، سنگ بنای یک کارآزمایی بالینی است، درست انجام شدن آن بسیار مهم است. دو ویژگی بسیار مهم در تصادفی‌سازی وجود دارد. اول آنکه اختصاص تصادفی درمانها کاملاً واقعی است و دوم آنکه انتصاب به صورت دخالت ناپذیر (tamperproof) است بطوریکه هیچکدام از عوامل عمدی و غیرعمدی نمی‌توانند تصادفی‌سازی را تحت تاثیر قرار دهند.

[قبل از تصادفی‌سازی] تکمیل مجموعه اطلاعات پایه توسط شرکت‌کنندگان مهم است زیرا برای بررسی واجدالشرایط بودن آنها برای ورود به کارآزمایی ضروری است. همچنین اعلام رضایت افراد برای شرکت در کارآزمایی باید قبل از تصادفی‌سازی صورت گیرد. سپس افراد [واجد شرایط و رضایت] بصورت تصادفی توسط الگوریتم‌های کامپیوتری و/یا با استفاده از مجموعه‌ای از اعداد تصادفی انتصاب داده می‌شوند. هنگامی که یک توالی برای انتصاب تصادفی گروه‌های مطالعه تولید می‌شود باید به ترتیب برای شرکت‌کنندگان که به کارآزمایی وارد می‌شوند برای انتصاب استفاده گردد.

طراحی روش انتصاب تصادفی بطوریکه اعضا تیم تحقیق تأثیری روی تخصیص‌ها نداشته باشند بسیار ضروری است. برای مثال، برای کارآزمایی‌هایی که در یک سایت انجام می‌شود، انتصاب‌های تصادفی درمان می‌تواند در یک مجموعه پاکت مهر و موم شده بوسیله شخصی ثالثی که در تخصیص تصادفی و باز کردن پاکتها مشارکت ندارد، انجام شود. هر پاکت باید شماره گذاری شود (بطوریکه در پایان مطالعه شمرده شود)، مات باشد (برای جلوگیری از transillumination توسط یک نور قوی) و بطورکلی به صورت دخالت ناپذیر باشد. هنگامیکه یکی از شرکت‌کنندگان به صورت تصادفی انتخاب شد، نام او و شماره پاکت باز نشده در توالی در حضور یک کارمند دیگر ثبت شده و هر دو کارمند پاکت را امضا می‌کنند. سپس پاکت باز می‌شود و شرکت‌کننده به گروه مندرج در آن انتصاب داده می‌شود.

کارآزمایی‌های چند مرکزی بطور معمول از تسهیلات تصادفی دخالت ناپذیر مجزا استفاده می‌کنند. در این حالت، هنگامیکه یک شرکت‌کننده واجد شرایط برای انتخاب تصادفی اعلام آمادگی می‌کند، کارمند کارآزمایی تماس می‌گیرد. سپس، نام و کد شرکت‌کننده جدید را ثبت می‌کند. سپس بر اساس یک برنامه کامپیوتری که شماره انتصاب تصادفی مرتبط با مداخله را ایجاد می‌کند، انتصاب تصادفی صورت می‌گیرد. همچنین مادامیکه که برنامه کامپیوتری دخالت ناپذیر در دسترس باشد درمان می‌تواند به صورت تصادفی به وسیله برنامه کامپیوتری در یک مطالعه تک مرکزی انتصاب داده شود. اقدامات احتیاطی دقیقی جهت جلوگیری از دستکاری روی تصادفی‌سازی مورد نیاز می‌باشد زیرا محققان گاهی اوقات تحت فشار برای نفوذ در روند تصادفی‌سازی قرار می‌گیرند (به عنوان مثال، مشاهده فردی در گروه کنترل که به نظر می‌رسد بطور خاصی برای گروه مداخله مناسب می‌باشد).

## تکنیک‌های خاص تصادفی‌سازی

رویکرد ترجیحی آن است که تصادفی‌سازی روی هر فرد شرکت‌کننده برای هر گروه مداخله صورت گیرد. در کارآزمایی‌های با اندازه کوچک تا متوسط برای ایجاد توازن در تعداد شرکت‌کنندگان در هر گروه (تصادفی‌سازی بلوکی) و توزیع متوازن متغیرهای پایه که برای پیش‌بینی پیامد شناخته شده می‌باشند (تصادفی‌سازی طبقه‌بندی شده<sup>۱</sup>) از روشهای تصادفی‌سازی خاصی استفاده شود. این روشها قدرت مطالعه را کمی افزایش می‌دهند.

**تصادفی‌سازی بلوکی (blocked randomization)**، تکنیک رایجی است برای اطمینان از آنکه شرکت‌کنندگان بطور مساوی بین گروه‌های مطالعه توزیع شده باشند. تصادفی‌سازی روی بلوک‌هایی با تعداد از قبل مشخص شده انجام می‌گردد. برای مثال اگر اندازه بلوک ۶ باشد، به طور نرمال در هر بلوک شش نفر، سه نفر به صورت تصادفی به یک گروه منتسب می‌شوند و سه نفر به گروه دیگر. این بدین معناست که در یک مطالعه با ۳۰ نفر شرکت‌کننده، دقیقاً ۱۵ نفر به هر گروه انتساب داده خواهد شد. یا در یک مطالعه با ۳۳ نفر شرکت‌کننده، عدم تناسب بزرگتر از ۱۸:۱۵ نمی‌باشد. تصادفی‌سازی بلوکی با اندازه بلوک‌های ثابت برای مطالعات پنهان‌سازی نشده، خیلی مناسب نمی‌باشد. زیرا انتساب افراد در پایان هر بلوک می‌تواند قابل پیش‌بینی و دستکاری کردن باشد. این مشکل می‌تواند با تعریف بلوک‌هایی با اندازه‌های متغیر (برای مثال بلوک‌های چهار تا هشت‌تایی) مطابق با ترتیبی که برای محقق نامشخص است به حداقل برسد.

**تصادفی‌سازی بلوکی طبقه‌بندی شده (stratified blocked randomization)**، تضمین می‌کند که یک متغیر پیشگویی کننده مهم برای پیامد، بطور مساوی بین گروه‌های مطالعه توزیع شده است. در یک کارآزمایی برای بررسی اثر یک دارو جهت جلوگیری از شکستگی، سابقه شکستگی قبلی یک متغیر پیش‌بینی کننده قوی برای پیامد است. بنابراین بایستی افرادی که سابقه شکستگی قبلی دارند به تعداد یکسانی به دو گروه اختصاص داده شوند. این کار با انجام تصادفی‌سازی بلوکی مجزا برای افراد با شکستگی و بدون شکستگی می‌تواند انجام شود. تصادفی‌سازی بلوکی طبقه‌بندی شده می‌تواند در کارآزمایی‌های کوچک به دلیل کاهش تغییرپذیری در پیامد به سبب عدم تناسب شانس در متغیرهای پیشگویی کننده پایه، قدرت مطالعات را کمی افزایش دهد. این کار در کارآزمایی‌های بزرگ (بیشتر از ۱۰۰۰ شرکت‌کننده) کمتر سودمند خواهد بود زیرا انتساب تصادفی به تنهایی توزیع متغیرهای پایه را تضمین می‌کند.

یک محدودیت مهم در تصادفی‌سازی بلوکی طبقه‌بندی شده آنست که فقط هنگامیکه تعداد متغیرهای پایه کم است، نه بیشتر از دو یاسه متغیر، توازن را می‌تواند ایجاد نماید. یک روش برای جلوگیری از این محدودیت، **تصادفی‌سازی تطبیقی (adaptive randomization)** است. در این روش از یک "سکه مغرضانه" برای تغییر احتمال انتساب هر یک از شرکت‌کنندگان جدید استفاده می‌شود. بطوریکه برای مثال فردی که بر اساس هر تعداد از متغیرهای پیش‌آگهی پایه، نمره ریسک بالایی دارد بیشتر احتمال دارد به گروه مطالعه‌ای ملحق شود که ریسک کلی آن بر اساس همه شرکت‌کنندگان تصادفی تخصیص داده شده به آن تاکنون کمتر می‌باشد. از معایب این روش دشواری توضیح نحوه انتساب در گروه‌های مطالعه برای شرکت‌کنندگان پیش از اخذ رضایت آگاهانه است. دیگر آنکه اجرای آن بوسیله یک سیستم کامپیوتری تعاملی که احتمالات سکه مغرضانه را با هر بار تصادفی‌سازی دوباره محاسبه نماید، بسیار پیچیده است.

معمولاً بهترین تصمیم انتصاب افراد بطور مساوی در گروه‌های مطالعه می‌باشد. در اینصورت قدرت مطالعه برای حجم نمونه مفروض افزایش می‌یابد. با این وجود میزان تضعیف قدرت مطالعه با حتی نسبت ۲:۱ اندک است. گاهی تخصیص نابرابر شرکت‌کنندگان به گروه‌های کنترل و مداخله می‌تواند مناسب باشد:

- افزایش گروه مداخله نسبت به کنترل سبب می‌شود تا کارآزمایی برای شرکت‌کنندگان جذاب‌تر گردد. مانند افراد مبتلا به عفونت HIV که می‌خواهند شانس بیشتری برای دریافت درمان فعال در صورت شرکت در کارآزمایی داشته باشند.
- کاهش گروه مداخله نسبت به کنترل می‌تواند کارآزمایی را هنگامیکه مداخله بسیار گران است مقرون به صرفه نماید (مانند کارآزمایی رژیم غذایی کم چرب در برنامه سلامت زنان).
- افزایش سهم اختصاص یافته به گروهی که کنترل برای چندین گروه درمان فعال است، می‌تواند با افزایش دقت در تخمین گروه کنترل، قدرت هر مقایسه را افزایش دهد.

**تصادفی‌سازی جفت‌های تطبیق یافته (randomization of matched pairs)**، یک استراتژی برای حفظ تعادل در متغیرهای مخدوشگر پایه می‌باشد که به انتخاب جفت‌هایی از شرکت‌کنندگان، که در ویژگی‌های مهم مانند سن و جنس همسان است، نیاز دارد. سپس به طور تصادفی یک عضو از هر جفت به هر گروه مطالعه اختصاص داده می‌شود. یکی از معایب این تصادفی‌سازی آن است که شرکت‌کننده واجد شرایط بایستی تا زمانی که جفت مناسبی پیدا شود منتظر بماند. علاوه بر این، در کارآزمایی‌های بزرگ هنگامیکه انتساب تصادفی سبب توازن گروه‌ها روی متغیرهای پایه می‌گردد، تطبیق ضروری نمی‌باشد.

استفاده از این تصادفی‌سازی زمانی جالب می‌شود که اثرات متضاد مداخله و کنترل را در افراد مشابه بخواهیم بررسی نماییم. در مطالعه رتینوپاتی دیابتی، برای مثال، هر شرکت کننده با بینایی یک چشم بطور تصادفی به گروه مداخله و کنترل انتساب داده می‌شوند.

## پنهان‌سازی

در صورت امکان محقق باید مداخله را به گونه ای طراحی نماید که شرکت کنندگان، کارکنان، ارزیاب‌ها و تحلیلگران از چگونگی انتصاب در گروه‌های مطالعه اطلاعی نداشته باشند. اگر امکان پنهان‌سازی همه این افراد وجود ندارد، مطلوب آن است که تا آنجا که امکان دارد پنهان‌سازی انجام شود (برای مثال پنهان‌سازی ارزیاب‌ها). در یک کارآزمایی تصادفی شده، پنهان‌سازی مانند تصادفی‌سازی مهم است. تصادفی‌سازی، اثر متغیرهای مخدوشگر را در زمان تصادفی‌سازی کاهش می‌دهد اما روی تفاوت‌هایی که بین گروه‌ها در طول پیگیری وجود دارد هیچ تاثیری ندارد (جدول ۲). پنهان‌سازی سوگیری (bias) با منشا پس از تصادفی‌سازی را به حداقل می‌رساند مواردی نظیر مداخلات همزمان و ارزیابی و تصمیم‌گیری سوگیرانه در خصوص پیامدها.

استفاده از پنهان‌سازی از سوگیری ناشی از مداخلات همزمان داروها، درمانها و رفتارهای خاص نسبت به مداخله‌ای که ریسک پیامد مطلوب را تغییر می‌دهد، جلوگیری می‌کند. دومین هدف مهم پنهان‌سازی به حداقل رساندن اثبات و قضاوت مغرضانه نتایج می‌باشد. در یک کارآزمایی پنهان‌نشده محقق ممکن است وسوسه شود که با دقت بیشتری به نتایج در گروه درمان نشده توجه کند یا پیامدهای بیشتری را تشخیص دهد. برای مثال در یک کارآزمایی پنهان نشده از استاتین درمانی، محققان ممکن است فقط از شرکت کنندگان در گروه مداخله در مورد دردهای عضلانی و لمس بودن به منظور ارزیابی وجود میوزیت سوال بپرسند. پنهان‌سازی شرکت‌کننده‌ها مخصوصاً زمانی مهم است که پیامدها مبتنی بر علائم گزارش شده توسط خود شرکت کننده (self-reported symptoms) می‌باشد.

بعد از آنکه وقوع یک پیامد قطعی شد، ممکن است نیاز به قضاوت باشد. برای مثال، اگر پیامد یک کارآزمایی سکت قلبی باشد، محققان به جمع‌آوری اطلاعات درباره علائم، یافته EKG و آنزیم قلبی می‌پردازند. سپس افراد خبره که پنهان‌سازی گروه مداخله برایشان صورت گرفته است از این داده‌ها استفاده کرده و درمورد آنکه آیا سکت قلبی رخ داده است یا خیر قضاوت می‌کنند.

جدول ۲. در یک کارآزمایی تصادفی پنهان‌سازی شده، تصادفی‌سازی اثر متغیرهای مخدوشگر را به حداقل می‌رساند و پنهان‌سازی مداخلات همزمان و اثبات مغرضانه نتایج را به حداقل می‌رساند.

توضیح ارتباط		استراتژی برای غیرمتمثل شمردن توضیح رقیب
شانس (Chance)	فاز طراحی	فاز تحلیل
افزایش حجم نمونه و استراتژیهای دیگر برای افزایش دقت.	محابسه pvalue و بازه اطمینان و تفسیر آنها در بستر شواهد قبلی.	
سوگیری (Bias)	در نظر گرفتن عواقب بالقوه هر اختلاف بین سوال تحقیق و طراحی مطالعه. تغییر طراحی مطالعه در صورت لزوم	بررسی سازگاری با مطالعات دیگر (مخصوصاً مطالعاتی که طراحی متفاوتی دارند).
	جمع‌آوری داده‌های اضافی که اجازه ارزیابی سوگیری احتمالی را بدهد.	تحلیل داده‌های اضافی برای فهمیدن آنکه آیا سوگیری اتفاق افتاده است.
اثر - علت	عدم استفاده از متغیرهایی که تحت تاثیر متغیرهای پیشگویی قرار دارند به عنوان معیار ورود یا متغیر تطبیق.	متغیرهایی که تحت تاثیر متغیرهای پیشگویی هستند را کنترل نکنید.
	در کارآزمایی غیرممکن است.	
مخدوشگر (Confounder)	متغیرهای مخدوشگر پیش از تصادفی‌سازی	تصادفی‌سازی
	متغیرهای مخدوشگر پس از تصادفی‌سازی	پنهان‌سازی
علت - اثر	مطلوب	

نتایج کارآزمایی Canadian Cooperative Multiple Sclerosis نشان دهنده اهمیت پنهان‌سازی در داورى غیرمغرضانه نتایج می‌باشد. افراد مبتلا به اسکروز چندگانه (MS) در دو گروه تعویض پلاسما ترکیبی+سیکلوفسفامید/پروپیل تیورین و تعویض پلاسما غیرواقعی+دارونما بصورت تصادفی قرار گرفتند. در پایان کارآزمایی، شدت اسکروز چندگانه از طریق آزمایشات ساختاریافته بوسیله دو گروه نورولوژیست مورد ارزیابی قرار گرفت. گروه اول شامل نورولوژیست‌هایی که چگونگی انتصاب برای آنها پنهان شده بود و گروه دوم شامل نورولوژیست‌هایی که عمل پنهان‌سازی برای آنها انجام نشده بود. براساس ارزیابی نورولوژیست‌های گروه اول درمان موثر نبود. در حالیکه گروه دوم تاثیر قابل توجه درمان را گزارش کرده بودند. نورولوژیست‌های پنهان‌سازی نشده به صورت برنامه‌ریزی شده دچار داورى مغرضانه در نتایج نشده

بودند اما انسان قویا متمایل به دیدن بهبود در بیماران پس از درمان است خصوصا اگر درمان دردناک و بالقوه مضر باشد. پنهان‌سازی، داوری مغرضانه نتایج را به حداقل می‌رساند.

در صورتیکه پیامد کارآزمایی، پیامدی نظیر "مرگ" باشد، پنهان‌سازی کمتر اهمیت پیدا می‌کند. همچنین در مورد ارزیابی‌های خودکار که فرصتی برای داوری مغرضانه وجود ندارد، پنهان‌سازی اهمیتی ندارد. در سایر پیامدها نظیر علت مرگ، تشخیص بیماری، ارزیابی‌های فیزیکی، مقیاس‌های پرسشنامه‌ای، و شرایط خود گزارش شده (self-reported)، امکان قضاوت و داوری مغرضانه نتایج وجود دارد.

پس از پایان کارآزمایی، ایده خوبی است که به ارزیابی آنکه آیا واقعا پنهان‌سازی روی شرکت‌کنندگان و محققان انجام شده بود بپردازیم. برای این کار از محقق خواسته می‌شود که حدس بزنند که فرد شرکت‌کننده به کدام گروه تعلق داشته است. اگر نسبت صحیح به دست آمده بالاتر از حد مورد انتظار باشد در بحث یافته‌ها باید قید گردد: ارزیابی سوگیری بالقوه دارد و عدم پنهان‌سازی جزئی دلیل آن می‌تواند باشد.

**در صورتیکه پنهان‌سازی ممکن نباشد چه کاری باید انجام دهیم؟**

در بعضی موارد پنهان‌سازی به دلایل تکنیکی یا اخلاقی مشکل یا غیرممکن است. برای مثال، پنهان کردن افرادی که قرار است تحت مداخله آموزشی، رژیم، و یا تمرینی قرار بگیرند. همینطور مداخلات جراحی در اغلب موارد نمی‌تواند بصورت پنهان‌سازی انجام شود زیرا غیراخلاقی است. با اینحال، به دلیل آنکه عمل جراحی با خطر همراه می‌باشد، تعیین آنکه آیا این روش واقعا موثر است بسیار مهم است. برای مثال، یک کارآزمایی تصادفی نشان داد که دبریدمان آرتروسکوپی از غضروف زانو از آرتروسکوپی با دبریدمان غیرواقعی برای تسکین درد آرتروز زانو مؤثرتر نمی‌باشد. در این مورد، پذیرش ریسک برای شرکت‌کنندگان در گروه کنترل ارزش داشته باشد اگر نتایج کارآزمایی مانع از آن شود که هزاران بیمار تحت درمان غیر موثر قرار گیرند.

اگر پنهان‌سازی روی مداخله نتواند صورت گیرد، محقق بایستی تا جایی که امکان دارد، مداخلات همزمان بالقوه را به حداقل برساند و اطمینان دهد پنهان‌سازی برای ارزیابان و تحلیلگران انجام شده است. برای مثال، یک محقق به بررسی اثر بوگا برای رهایی از گرگرفتگی در زمان یائسگی می‌پردازد. شرکت‌کنندگان هر دو گروه کنترل و مداخله از شروع درمان‌های جدید نظیر فعالیت‌های آرامش بخش دیگر یا درمان‌های دیگر برای جلوگیری از گرگرفتگی تا پایان کارآزمایی خودداری کنند. همچنین مجریان کارآزمایی که

جمع‌آوری اطلاعات مربوط به شدت گرگرفتگی را برعهده دارند می‌توانند کارمندانی به غیر از افرادی که آموزش دهنده بوگا می‌باشند، در نظر گرفته شوند.

## منبع

در صورت استفاده از مطالب این جزوه به منبع زیر ارجاع شود.

*Designing Clinical Research (2013), Third Edition by Stephen B. Hulley.*

بودند اما انسان قویا متمایل به دیدن بهبود در بیماران پس از درمان است خصوصا اگر درمان دردناک و بالقوه مضر باشد. پنهان‌سازی، داوری مغرضانه نتایج را به حداقل می‌رساند.

در صورتیکه پیامد کارآزمایی، پیامدی نظیر "مرگ" باشد، پنهان‌سازی کمتر اهمیت پیدا می‌کند. همچنین در مورد ارزیابی‌های خودکار که فرصتی برای داوری مغرضانه وجود ندارد، پنهان‌سازی اهمیتی ندارد. در سایر پیامدها نظیر علت مرگ، تشخیص بیماری، ارزیابی‌های فیزیکی، مقیاس‌های پرسشنامه‌ای، و شرایط خود گزارش شده (self-reported)، امکان قضاوت و داوری مغرضانه نتایج وجود دارد.

پس از پایان کارآزمایی، ایده خوبی است که به ارزیابی آنکه آیا واقعا پنهان‌سازی روی شرکت‌کنندگان و محققان انجام شده بود بپردازیم. برای این کار از محقق خواسته می‌شود که حدس بزند که فرد شرکت‌کننده به کدام گروه تعلق داشته است. اگر نسبت صحیح به دست آمده بالاتر از حد مورد انتظار باشد در بحث یافته‌ها باید قید گردد: ارزیابی سوگیری بالقوه دارد و عدم پنهان‌سازی جزئی دلیل آن می‌تواند باشد.

#### در صورتیکه پنهان‌سازی ممکن نباشد چه کاری باید انجام دهیم؟

در بعضی موارد پنهان‌سازی به دلایل تکنیکی یا اخلاقی مشکل یا غیرممکن است. برای مثال، پنهان کردن افرادی که قرار است تحت مداخله آموزشی، رژیم، و یا تمرینی قرار بگیرند. همینطور مداخلات جراحی در اغلب موارد نمی‌تواند بصورت پنهان‌سازی انجام شود زیرا غیراخلاقی است. با اینحال، به دلیل آنکه عمل جراحی با خطر همراه می‌باشد، تعیین آنکه آیا این روش واقعا موثر است بسیار مهم است. برای مثال، یک کارآزمایی تصادفی نشان داد که دبریدمان آرتروسکوپی از غضروف زانو از آرتروسکوپی با دبریدمان غیرواقعی برای تسکین درد آرتروز زانو مؤثرتر نمی‌باشد. در این مورد، پذیرش ریسک برای شرکت‌کنندگان در گروه کنترل ارزش داشته باشد اگر نتایج کارآزمایی مانع از آن شود که هزاران بیمار تحت درمان غیر موثر قرار گیرند.

اگر پنهان‌سازی روی مداخله نتواند صورت گیرد، محقق بایستی تا جایی که امکان دارد، مداخلات همزمان بالقوه را به حداقل برساند و اطمینان دهد پنهان‌سازی برای ارزیابان و تحلیلگران انجام شده است. برای مثال، یک محقق به بررسی اثر یوگا برای رهایی از گرگرفتگی در زمان یائسگی می‌پردازد. شرکت‌کنندگان هر دو گروه کنترل و مداخله از شروع درمان‌های جدید نظیر فعالیت‌های آرامش بخش دیگر یا درمان‌های دیگر برای جلوگیری از گرگرفتگی تا پایان کارآزمایی خودداری کنند. همچنین مجریان کارآزمایی که

جمع‌آوری اطلاعات مربوط به شدت گرگرفتگی را برعهده دارند می‌توانند کارمندانی به غیر از افرادی که آموزش دهنده یوگا می‌باشند، در نظر گرفته شوند.

#### منبع

در صورت استفاده از مطالب این جزوه به منبع زیر ارجاع شود.

**Designing Clinical Research (2013), Third Edition by Stephen B. Hulley.**