

دانشگاه علوم پزشکی شاهرود

معاونت پژوهشی و فناوری

اجزای یک پروتکل مطالعه کارآزمایی بالینی

*.عنوان و خلاصه

*.مقدمه : زمینه علمی و توضیح منطق مطالعه .

معمولاً مقدمه شامل متنی بدون فرمت ساختاری است که در آن نویسندگان زمینه علمی و منطق تحقیق خود را توجیه می کنند. منطق ممکن است تفسیری (مثلاً ارزیابی احتمالی یک دارو بر فانکشن کلیه) یا واقع بینانه (pragmatic) (مثلاً راهنمایی پزشکان با مقایسه تأثیرات بالینی دوزهای مختلف) باشد. نویسندگان باید شواهدی از مزایای هرگونه مداخله فعال بکاررفته در کارآزمایی را گزارش کنند. همچنین یک توضیح معقول در مورد اینکه چگونه مداخله تحت مطالعه ممکن است عمل کند ارائه کنند بخصوص اگر تجربیات قبلی با مداخله مذکور کم یا هیچ باشد.

Helsinki Declaration بیان می کند که تحقیقات پزشکی بر روی انسانها باید بر اساس اطلاعات جامعی از مقالات علمی مرتبط با موضوع باشند. یعنی قرار دادن غیر ضروری انسانها در معرض خطرات تحقیق، غیر اخلاقی است. نشان داده شده که برخی کارآزمایی های بالینی ضروری نبوده اند چون مسئله ای که بدنبال آن بوده اند می توانست با یک بررسی سیستماتیک مقالات موجود، پاسخ داده شود بنابراین نیاز به یک کارآزمایی جدید بایستی در مقدمه، توجیهی شود. بطور ایده ال یک مقدمه باید شامل ارجاع به کارآزمایی هایی مشابه پیشین یا اشاره به نبود چنین تحقیقاتی باشد. در بخش اول مقدمه، نویسندگان باید مشکلی که کارآزمایی را ضروری نموده است توصیف کنند. باید با بحث در مورد ماهیت، محدوده و شدت مشکل، یک توضیح جذاب برای علت انجام مطالعه ارائه شود. این اطلاعات اغلب در گزارشات دیده نمی شود. سپس نویسندگان باید به صورت مختصر راهکارهای متنوع بررسی مشکل را بیان کنند و اهداف تحقیق را نیز ممکن است در این قسمت ذکر نمایند.

*. روشها: مورد ۲a: معیارهای واجد شرایط بودن برای شرکت کنندگان

هر RCT، مشکلی را که مرتبط با جمعیتی با شرایط مورد نظر است بررسی می کند. انجام دهندگان کارآزمایی ها معمولاً این جمعیت را با استفاده از معیارهای واجد شرایط بودن شرکت کنندگان و با انجام تحقیق در یک یا چند مرکز محدود می کنند. معیارهای انتخابی معمول ممکن است مربوط به سن، جنس، تشخیص بالینی و شرایط بیماری مشترک باشد. معیارهای خروج از مطالعه باید صراحتاً تعریف شود. اگر هرگونه عدم دقت در تشخیص بالینی بیماران وجود دارد باید شرح داده شود چون ممکن است بر قدرت تحقیق اثر بگذارد. باید توصیف دقیقی از شرکت کنندگان در مطالعه و زمینه ای که آنها در آن مورد بررسی قرار گرفتند ارائه شود تا خوانندگان بتوانند روایی خارجی (قابلیت تعمیم) نتایج کارآزمایی را ارزیابی کنند. بخصوص روش جمع آوری شرکت کنندگان مهم است مثلاً توسط ارجاع یا انتخاب توسط خود محققین (مثلاً از طریق آگهی). چون معیارهای ورود قبل از تصادفی کردن بکار می روند بر روایی داخلی کار آزمایی تا تیری ندارند ولی بر روایی خارجی اثر می گذارند.

معیارهای ورود شرکت کنندگان به مطالعه با وجود اهمیتشان به میزان کافی گزارش نمی شوند مثلاً ۲۵٪ از ۳۶۴ گزارش RCT در جراحی، معیار های ورود را مشخص نکردند. هشت کارآزمایی منتشر شده که منجر به ارائه هشدار کلینیکی از سوی موسسات ملی سلامت شدند بطور متوسط ۳۱ معیار ورود را مشخص کردند ام فقط ۶۳٪ معیارها در مقالات مجلات، ذکر شده و فقط ۱۹٪ در هشدار های کلینیکی به آن اشاره گردید. از طرفی تعداد معیار های ورود در کارآزمایی های مربوط به سرطان بین دهه های ۱۹۷۰ و ۱۹۹۰ بطور قابل توجهی افزایش پیدا کرده است.

مورد ۲: شرایط و محل های جمع آوری اطلاعات .

شرایط و مراکز جمع آوری اطلاعات برروایی خارجی یک کارآزمایی اثر می گذارند. موسسات مراقبت سلامت از لحاظ سازماندهی، مهارت و منابع و خطر پایه برای آن وضعیت پزشکی تحت بررسی بسیار با هم فرق دارند. آب وهوا و سایر فاکتور های فیزیکی، اقتصاد، جغرافیا و محیط فرهنگی و اجتماعی، همه می توانند بر روایی خارجی یک مطالعه اثر بگذارند.

نویسندگان باید تعداد و نوع شرایط و مراقبان سلامت درگیر در مطالعه را گزارش کنند تا خوانندگان قدرت ارزیابی روایی خارجی را داشته باشند. آنها باید محلها و محیطهای انجام مطالعه شامل کشور، شهر و محیط پیرامون (مثلاً اجتماع، مطب، کلینیک بیمارستانی یا واحد بستری بیمار در بیمارستان) را توصیف نمایند. بخصوص باید روشن شود که آیا کارآزمایی در یک یا چند مرکز صورت خواهد گرفت. این توصیف بایستی اطلاعات کافی را در اختیار خوانندگان بگذارد تا آنها بتوانند قضاوت کنند که آیا نتایج کارآزمایی به شرایط خود آنها مربوط می شود یا خیر. همچنین

نویسندگان باید هر اطلاعات دیگری در مورد محلها و محیط‌هایی که می‌تواند بر نتایج مطالعه اثر بگذارد را گزارش نمایند مانند مشکلات رفت و آمد که ممکن است بر شرکت بیمار اثر بگذارد.

*** مورد ۳:** جزئیات دقیق در مورد مداخله‌های مورد نظر برای هر گروه و چگونگی و زمان انجام اعمال آنها.

نویسندگان باید هر مداخله و از جمله مداخلات گروه کنترل را به طور کامل بیان کنند. اگر از پلاسبو استفاده خواهد شد، مشخصات پلاسبو و روشی که با آنها پنهان شده است باید گزارش شود. بخصوص مهم است که مراقبت معمول داده شده به یک-گروه کنترل یا یک مداخله که ممکن است ترکیبی از چند مداخله باشد بطور کامل توصیف شود. در برخی از موارد، توصیف انجام دهندگان درمانها بسیار مهم است چون ممکن است بخشی از مداخله را تشکیل دهد. مثلاً در مداخلات جراحی، توصیف تعداد، آموزش و مهارت جراحان علاوه بر روش جراحی، ضروری است. لازم است نویسندگان، جزئیات زمان بندی و مدت اعمال مداخله‌ها را گزارش کنند بخصوص اگر مداخله‌های چند جزئی مورد نظر باشد.

*** مورد ۴:** اهداف اختصاصی و فرضیه‌ها.

اهداف، سوالاتی هستند که کارآزمایی برای پاسخ به آنها طراحی شده است. آنها اغلب مرتبط با کارآیی یک مداخله خاص درمانی یا پیشگیری می‌باشند. فرضیه‌ها، سوالات از قبل مشخص شده‌ای هستند که برای کمک به رسیدن به اهداف مورد آزمایش قرار می‌گیرند.

فرضیه‌ها، اختصاصی‌تر از اهداف هستند و باید تحت ارزیابی آماری صریح قرار گیرند. درعمل اهداف و فرضیه‌ها همیشه به راحتی قابل افتراق نیستند.

*** مورد ۵a:** تعریف مشخص متغیرهای پیامد (outcome measure) اولیه و ثانویه.

تمام مطالعات کارآزمایی بالینی، متغیرهای پاسخ یا پیامدهایی که گروهها به خاطر آنها با یکدیگر مقایسه می‌شوند را ارزیابی می‌کنند.

بیشتر کارآزمایی‌ها چندین پیامد دارند که برخی از آنها بیش از بقیه موردتوجه هستند. متغیر پیامد اولیه، متغیر از پیش تعیین شده‌ای با بیشترین اهمیت است که معمولاً برای محاسبه سائز نمونه از آن استفاده می‌شود. برخی کارآزمایی‌ها ممکن است بیش از یک پیامد اولیه یا اصلی داشته باشد. اما بیش از یک یا دو پیامد اصلی داشتن منجر به ایجاد مشکلات تفسیری همراه با متعدد بودن آنالیزها می‌شود و توصیه نمی‌گردد. پیامد های اولیه بایستی صراحتاً در پروتکل RCT ذکر شوند. سایر پیامد های مورد توجه، پیامد های ثانویه هستند. ممکن است چندین پیامد

ثانویه وجود داشته باشد که اغلب شامل اثرات پیش بینی نشده یا ناخواسته مداخله هستند. تمام متغیر های پیامد، چه اولیه و چه ثانویه باید به طور کامل شناسایی و تعریف شوند. وقتی پیامد ها در زمان های متعددی پس از تصادفی کردن ارزیابی می شوند. نویسندگان باید نقطه زمانی ارزیابی پیامد اصلی را مشخص نمایند. گاهی بهتر است فرد ارزیابی کننده پیامد ها (مثلا اگر مهارت های خاصی برای آن لازم است) و تعداد ارزیابی کنندگان مشخص شود.

بسیاری از بیماری ها تعداد زیادی پیامد احتمالی دارند که می تواند با استفاده از معیارهای مختلف یا ابزار های مختلف اندازه گیری شود. در صورتی که معیار ها یا اصول توافقی معتبری از قبل در دسترس باشد، بایستی در کارآزمایی مورد استفاده قرار گیرد تا کیفیت اندازه گیری بهتر شود و به مقایسه با مطالعات مشابه کمک نماید. مثلا ارزیابی کیفیت زندگی با استفاده از یک ابزار معتبر، توصیه می شود. نویسندگان بایستی مبدأ و خصوصیات معیارها را ذکر نمایند.

مورد ۵b: ذکر روشهایی که برای بهبود کیفیت اندازه گیری ها به کار می رود (مثل مشاهدات متعدد، آموزش به ارزیابی کنندگان).

نویسندگان باید جزئیات کاملی در مورد نحوه اندازه گیری پیامدهای اولیه و ثانویه و هرگونه اقدامات خاص پیش بینی شده برای افزایش پایایی اندازه گیری ها ارائه نمایند. اندازه گیری برخی پیامدها از بقیه آسانتر است. ارزیابی مرگ (به هر علتی) معمولا آسان است در حالیکه اندازه گیری فشار خون، افسردگی یا کیفیت زندگی مشکل تر است. از برخی راهکارها برای بهبود کیفیت اندازه گیری ها می توان استفاده نمود. مثلا اگر بیش از یک بار فشار خون اندازه گیری شود ارزیابی آن قابل اطمینان تر خواهد بود. اگر شرکت کنندگان و ارزیابی کنندگان نسبت به نحوه تخصیص گروه ها کور باشند، احتمالا ارزیابی ها بیشتر عاری از bias خواهند بود. اگر یک کارآزمایی نیازمند اندازه گیری های ناآشنایی است، آموزش استاندارد و رسمی به افرادی که اندازه گیری را انجام می دهند مفید است.

*. مورد a ۶: حجم نمونه چگونه تعیین شده است.

به دلایل علمی و اخلاقی، حجم نمونه برای یک کارآزمایی با در نظر گرفتن یک تعادل بین ملاحظات بالینی و آماری باید به دقت تعیین شود. بطور ایده آل یک مطالعه باید به قدر کافی بزرگ باشد تا احتمال بالایی (قدرت بالا) از کشف یک اختلاف معنی دار آماری برای یک تفاوت مهم بالینی داشته باشد. معمولا برای کشف تفاوت های کوچک به حجم نمونه زیاد احتیاج می باشد.

اجزاء محاسبه حجم نمونه شامل (۱) پیامدهای تخمین زده شده در هر گروه (که بیانگر تفاوت‌های مهم بالینی مورد نظر بین گروه‌های مداخله‌ای است)، (۲) سطح خطای α یا نوع I (۳) توان آماری (یا سطح خطای β (نوع II) ، (۴) انحراف معیار اندازه‌گیری‌ها برای پیامدهای مداوم، می‌باشد.

نویسندگان باید نشان دهند که چگونه حجم نمونه تعیین شده است. نویسندگان باید پیامد اولیه‌ای که محاسبه بر آن بنا شده، تمام اجزاء بکار رفته در محاسبه و حجم نمونه بدست آمده در هر گروه مقایسه را ذکر نمایند. در بعضی کارآزمایی‌ها، آنالیزهای میانی (interim) برای کمک به تصمیم‌گیری در مورد ادامه جمع‌آوری نمونه‌ها، انجام می‌شود. اگر به دلایلی این روش قرار است استفاده شود، باید علت و جزئیات آن توضیح داده شود. اغلب در گزارش‌های مربوط به مطالعات با حجم نمونه کم، این نتیجه‌گیری اشتباه دیده می‌شود که گروه‌های مداخله‌ای با یکدیگر فرق نداشتند در صورتی که تعداد افراد مورد مطالعه برای اتخاذ چنین نتیجه‌ای بسیار کم بوده است. مرور تحقیقات منتشر شده نسبت بالایی از کارآزمایی‌هایی را نشان می‌دهد که حجم نمونه بسیار کمی برای یافتن تاثیرات درمانی معنی‌دار بالینی داشتند. در حقیقت امکان یافتن تفاوت‌های واقعی کوچک اما ارزشمند بالینی وجود دارد، اگر کارآزمایی‌هایی با حجم نمونه زیاد برای یافتن آنها بکار رود.

مورد b ۶: توجه هر آنالیز میان مطالعه‌ای و قانونهای متوقف‌کننده بکار رفته

بسیاری از کارآزمایی‌ها، شرکت‌کنندگان را در طول یک دوره طولانی جمع‌آوری می‌کنند. اگر یک مداخله بسیار عالی یا بد عمل کند ممکن است لازم باشد مطالعه به دلایل اخلاقی زودتر از موقع پایان یابد. این نگرانی می‌تواند با امتحان نتایج بدست آمده وقتی اطلاعات جمع می‌شود، مورد توجه قرار گیرد. اما انجام آزمایشات آماری متعدد از اطلاعات در حال جمع‌آوری بدون تصحیح مناسب می‌تواند منجر به نتایج و تفاسیر اشتباه شود. روشهای آماری ترتیبی گروهی متعددی برای تطابق با آنالیزهای متعدد در دسترس است. استفاده از آنها بایستی از قبل در پروتکل کارآزمایی مشخص شود. با این روشها داده‌ها در هر آنالیز میان مطالعه‌ای مقایسه می‌شوند و یک مقدار بسیار کوچک p نمایانگر اهمیت آماری خواهد بود. مقداری از P که به دنبال آن کارآزمایی متوقف می‌شود باید از قبل مشخص باشد. نویسندگان باید گزارش کنند که آیا آنالیزهای متعدد را انجام می‌دهند و اگر اینطور است چند مورد خواهد بود و از چه روشهای آماری استفاده می‌شود.

*مورد Ya: روش بکار رفته برای تخصیص تصادفی شرکت‌کنندگان در مطالعه

بطور ایده آل شرکت کنندگان در یک کارآزمایی بایستی براساس روند شانس یا تصادف که غیر قابل پیش بینی است به گروه های مقایسه ای اختصاصی داده شوند. نویسندگان باید اطلاعات کافی در اختیار خواننده بگذارند تا بتواند روشهای به کار رفته برای ترتیب تخصیص تصادفی و احتمال bias در تخصیص گروه ها را ارزیابی نماید. بسیاری از این روشهای به کار رفته کافی هستند اما خواننده نمی تواند کفایت را از کلماتی مثل تخصیص تصادفی، تصادفی کردن یا تصادفی بدون توضیح بیشتر قضاوت نماید. نویسندگان باید روش ایجاد ترتیب را مشخص نمایند مثل یک جدول اعداد تصادفی یا یک تولید کننده کامپیوتری اعداد تصادفی. ترتیب ممکن است با روند minimization (کمینه سازی) که یک روش تصادفی محدود شده است، ایجاد گردد.

در برخی کارآزمایی ها شرکت کنندگان عمداً در تعداد نامساوی به هر مداخله اختصاص داده می شوند. به عنوان مثال برای کسب تجربه بیشتر با یک پروسه جدید یا محدود کردن هزینه های کارآزمایی. در چنین مواردی نویسندگان باید نسبت تصادفی کردن را گزارش کنند (مثلاً ۱:۲)

کلمه تصادفی، یک معنای دقیق تکنیکی دارد به این صورت که با تخصیص اتفاقی، هر شرکت کننده، یک احتمال مشخص از دریافت هر درمان یا مداخله را دارد اما درمان واقعی با یک روند شانس تعیین می شود و قابل پیش بینی نیست. اما اغلب از ((تصادف)) به اشتباه در مقالات برای توصیف کارآزمایی هایی که در آنها روشهای غیرتصادفی ((جبرگرایانه)) بکار رفته است مانند روش تناوبی، اعداد بیمارستان یا تاریخ تولد، استفاده می شود. وقتی محققان از چنین روشهای استفاده می کنند باید آن را دقیقاً توصیف کنند و از کلمه تصادفی و مانند آن استفاده نکنند. حتی کلمه نیمه تصادفی هم برای چنین مطالعاتی مورد سوال است. شواهد تجربی حاکی از آن است که چنین کارآزمایی هایی نتایج bias دارند. Bias در نتیجه ناتوانی در پنهان سازی کافی چنین سیستم های تخصیصی بوجود می آید.

مورد b ۷: جزئیات هرگونه محدودیت تصادفی کردن (مثلاً بلوک بندی، طبقه بندی و ...)

باید بیان شود که آیا محدودیتی برای تصادفی سازی اعمال شده است یاخیر، مثلاً با بیان تصادفی کردن ساده در صورت عدم استفاده از محدودیت. در غیر این صورت روشهایی که برای محدود کردن تصادفی کردن به کار رفته همراه با روش استفاده شده برای انتخاب تصادفی (مورد 8a) باید مشخص گردد. در مورد تصادفی کردن بلوکی نویسندگان باید جزئیاتی در مورد اینکه چگونه بلوکها ایجاد شدند (مثلاً با استفاده از طرح permuted block) اندازه بلوکها و اینکه آیا اندازه بلوک بطور تصادفی متنوع بوده است یا خیر ذکر کنند. در صورتی که طبقه بندی بکار رفته باید مشخص شود و اینکه چه فاکتورهایی دخالت داده شدند و چه روشهایی برای بلوک کردن بکار رفته اند ذکر گردد. اگر چه روش طبقه

بندی بخصوص برای کارآزمایی های کوچکتر یک تکنیک مفید است اما اگر فاکتورهای طبقه بندی کننده متعددی بکار روند انجام آن پیچیده خواهد شد. اگر حداقل سازی (minimization) بکار رفته باید صراحتاً مشخص شود، همچنین متغیرهایی که در طرح آن وارد شده اند. استفاده از یک عنصر تصادفی باید بیان شود.

*. مورد ۸ : روش بکاررفته برای انجام ترتیب تخصیص تصادفی (مثلاً با محفظه های شماره گذاری شده یا تلفن مرکزی) و مشخص کردن اینکه آیا ترتیب مذکور تا زمان اعمال مداخله ها مخفی بوده است یا خیر .

در مورد ۷ درباره ایجاد یک ترتیب غیر قابل پیش بینی از تخصیص دادن افراد به گروهها بحث شد. اینکه چگونه این ترتیب در زمان شرکت افراد در کار آزمایی اعمال می شود بسیار مهم است . بطور ایده آل باید از طریق پنهان سازی تخصیص، اختصاص افراد به گروهها اعمال شود. در این پروسه مهم افرادی که شرکت کنندگان را وارد مطالعه می کنند از نحوه تخصیص درمان نباید آگاه شوند . تصمیم برای قبول یا رد یک شرکت کننده باید بدون در نظر گرفتن تخصیص بعدی در ترتیب موجود اتخاذ شود و رضایت نامه آگاهانه از او گرفته شود. مخفی سازی نحوه تخصیص ، نباید با blinding اشتباه شود. این پروسه (Allocation concealment) از bias در انتخاب جلوگیری می کند، از اختصاص دادن صحیح شرکت کنندگان تا زمان آن حفاظت می کند و همیشه می تواند با موفقیت انجام شود . در صورتی که blinding از bias در عملکرد و در نتایج کار آزمایی جلوگیری می کند، از صحت نتایج پس از اختصاص افراد حفاظت می کند و همیشه نمی تواند اعمال شود . بدون مخفی سازی کافی، تخصیص افراد و حتی ترتیبهای سهمیه بندی تصادفی و غیر قابل پیش بینی هم می تواند مخدوش شود.

تخصیص توسط « شخص ثالث » یا تمرکز زدایی شده ، بسیار مطلوب است- بسیاری از روشهای خوب برای مخفی سازی تخصیص یک دخالت خارجی را شامل می شوند. استفاده از یک سیستم تصادفی کردن تلفن مرکزی یا داروخانه دو تکنیک شایع هستند. استفاده از سیستمهای تخصیص اتوماتیک هم در حال افزایش است. وقتی استفاده از کمک خارجی امکان پذیر نیست ، یک روش عالی برای مخفی کردن تخصیص، استفاده از محفظه های شماره گذاری شده است .

مداخله ها (اغلب داروها) بسته به ترتیب تخصیص در محفظه های مساوی شماره گذاری می شوند. نگهداشتن مداخله ها در پاکتهای اپک محکم بسته شده و به تناوب شماره گذاری شده، در صورتی که با دقت انجام و نظارت شود مکانیسم خوبی برای مخفی کردن تخصیص افراد است. اما اگر انجام آن ضعیف باشد، می تواند خراب گردد.

محققان باید اطمینان حاصل کنند که پاکتها به تناوب و فقط پس از اینکه نام و دیگر مشخصات شرکت کننده روی پاکت مناسب نوشته شده باز شوند.

شواهد تجربی قوی حاکی از آن است که مخفی سازی ناکافی تخصیص در ایجاد bias در تخمین اثرات درمانی دخالت دارد.

مورد ۹: چه کسی ترتیب تخصیص را ایجاد می کند؟ چه کسی شرکت کنندگان را وارد مطالعه می کند؟ و چه کسی آنها را به گروههای خود اختصاص می دهد؟

همانطور که در مورد ۸ اشاره شد، مخفی سازی مداخله اختصاص داده شده در زمان وارد کردن افراد در مطالعه مهم است. بنابراین علاوه بردانستن روشهای بکار رفته، باید بدانیم چگونه ترتیب تصادفی انجام می شود. بخصوص اینکه چه کسی ترتیب تخصیص را ایجاد می کند؟ چه کسی شرکت کنندگان را در مطالعه می کند؟ و چه کسی آنها را به گروههای کار آزمایشی اختصاص می دهد؟ روند شرکت دادن افراد در یک کار آزمایشی دو جنبه بسیار متفاوت دارد: بوجود آوردن و انجام. گرچه ممکن است افراد مشابهی بیش از یک کار را تحت هر تیر انجام دهند اما محققان تا جایی که ممکن است باید تلاش کنند تا افراد در گیر در روند ایجاد و روند انجام تخصیص ها کاملا جدا از هم باشند.

عدم جدا سازی ایجاد ترتیب تخصیص از اختصاص دادن افراد به گروههای مطالعه می تواند bias ایجاد کند. مثلا کسی که یک ترتیب تخصیص را ایجاد کرده می تواند یک کپی را نگهدارد و حین مصاحبه با شرکت کنندگان برای یک کار آزمایشی به آن مراجعه نماید. بنابراین آن فرد می تواند صرف نظر از غیر قابل پیش بینی بودن ترتیب تخصیص، پروسه پذیرفتن افراد در مطالعه (enrollment) یا تخصیص آنها به گروهها (assignment) را دچار bias نماید. با این وجود گاهی یک فرد مجبور است هم طرح را آماده کند و هم در تخصیص افراد به گروهها درگیر شود. در اینصورت محققان باید مطمئن باشند که برنامه تخصیص غیر قابل پیش بینی و دور از دسترس حتی فردی است که آن را ایجاد کرده است. پروتکل کار آزمایشی باید مشخص کند که محققان لیست تخصیص را در کجا نگهداری می کنند.

مورد ۱۰a: آیا شرکت کنندگان، انجام دهندگان مداخله ها و ارزیابی کنندگان نتایج، نسبت به تخصیص گروهها blind خواهند بود یا خیر.

در کار آزمایشی های کنترل شده، blinding به ناگاهای شرکت کنندگان در مطالعه، مراقبان سلامت و گاهی افراد جمع آوری و آنالیز کننده اطلاعات بالینی از مداخله اشاره می کند تا آنها تحت تأثیر آن آگاهی قرار نگیرند.

این اقدام برای جلوگیری از bias در مراحل متعدد یک کار آزمایشی کنترل شده مهم است، اگر چه ربط آن بسته به شرایط متفاوت است .

Blinding بیماران مهم است چون آگاهی آنها از مداخله می تواند بر پاسخ به درمان اثر بگذارد. بیمارانی که تحت درمان استاندارد قرار نگرفته اند ممکن است انتظارات نامطلوب یا افزایش اضطراب داشته باشند.

بیمارانی که تحت درمان استاندارد قرار گرفته و از آن آگاهی دارند، ممکن است احساس اطمینان کنند یا فکر کنند بین آنها و دیگران فرق گذاشته شده است. استفاده از کنترلهای پلاسبو همراه با blinding بیماران برای جلوگیری از bias ناشی از اثرات غیر اختصاصی مربوط به دریافت مداخله (اثرات پلاسبو)، صورت می گیرد. ولی باید توجه داشت که استفاده از پلاسبو در موارد محدودی اخلاقی است.

Blinding بیماران و مراقبان سلامت از bias عملکردی جلوگیری می کند . این نوع bias ممکن است وقتی بوجود آید که مداخله های درمانی اضافی (Co-interventions) برای شرکت کنندگان در یکی از گروههای مقایسه صورت گیرد.

Blinding بیماران، مراقبان سلامت و سایر افراد درگیر در ارزیابی نتایج (مثلا رادیولوژیست ها) خطر ایجاد bias در ارزیابی را کاهش می دهد، که به آن observer bias, detection bias, ascertainment bias, هم گفته می شود. این نوع bias وقتی رخ می دهد که آگاهی از تخصیص بیمار بر روند ارزیابی نتایج اثر بگذارد. مثلا در یک کار آزمایشی مولیتپل اسکروز کنترل شده با پلاسبو، ارزیابی های انجام شده توسط نورولوژیست های غیر blind ، یک مزیت آشکار مداخله را نشان داد. در نهایت ، blinding آنالیز کننده اطلاعات هم از bias جلوگیری می کند. آگاهی از مداخله های دریافت شده ممکن است بر انتخاب روشها و راهکارهای آنالیز آماری اثر بگذارد.

کار آزمایشی های بدون هرگونه blinding به عنوان open یا اگر کار آزمایشی دارویی هستند، open-label شناخته می شوند. این طراحی در تحقیقات اولیه یک دارو (کار آزمایشی های فاز II) شایع است.

بر خلاف مخفی سازی تخصیص، blinding ممکن است همیشه مناسب یا ممکن نباشد. مثال آن یک کار آزمایشی مقایسه کننده سطوح درد همراه با خونگیری از گوش یا انگشت است. به خصوص وقتی سنجش پیامد تا حدی وابسته به ذهنیت فرد است مثل ارزیابی درد یا مرگ، blinding مهم می باشد. blinding برای معیارهای عینی مثل مرگ ناشی از هر علتی، وقتی امکان کمی برای bias در ارزیابی وجود دارد، کمتر اهمیت پیدا می کند . اگر چه نبود blinding در هر کار آزمایشی می تواند منجر به مشکلات دیگر مثل attrition شود.

در یکسری کار آزمایی ها بخصوص کار آزمایی های جراحی، **double blinding** مشکل یا غیر ممکن است. اما ارزیابی کور نتایج حتی در کار آزمایی های باز هم می تواند انجام شود. مثلاً می توان از ضایعات قبل و بعد از درمان فوتوگرافی تهیه نمود و توسط فردی که در انجام کار آزمایی دخالت ندارد مورد ارزیابی قرارداد. برخی درمانها اثرات ناخواسته ای دارند که به قدری اختصاصی است که وقوع آنها به طور اجتناب ناپذیر درمان دریافت شده را برای بیمار و پرسنل پزشکی معلوم خواهد کرد. وقتی خطر چنین مساله ای وجود دارد ارزیابی کور پیامد بسیار مفید خواهد بود. بسیاری از کار آزمایی ها به عنوان دوسوکور **double-blind** توصیف می شوند. گرچه این کلمه بیانگر این است که نه بیمار و نه درمان کننده از نوع درمان آگاهی ندارند، اما در ارتباط با کور بودن سایر افراد مثل افراد ارزیابی کننده پیامدها مبهم است. نویسندگان بایستی مشخص کنند چه کسی کور بوده است (شرکت کنندگان، مراقبت کننده ها، ارزیابی کنندگان، ناظران یا آنالیز کنندگان اطلاعات)، مکانیسم **blinding** (مثلاً با کپسول یا قرص) و شباهت خصوصیات درمانها (مثلاً ظاهر، طعم و روش اعمال دارو) همچنین باید توضیح دهند که چرا هر کدام از گروههای شرکت کنندگان، مراقبان سلامت یا ارزیابی کنندگان **blind** نبوده اند.

نویسندگان اغلب گزارش نمی کنند که آیا **blinding** بکار رفته یا خیر و اگر **blinding** مشخص شده، اغلب جزئیات آن را ذکر نمی کنند. مثلاً ۵۱٪ از گزارش ۵۰۶ کار آزمایی در فیبروز کیستیک، ۳۳٪ از ۱۹۶ کار آزمایی در آرتروز روماتوئید و ۳۸٪ از ۶۸ کار آزمایی در درماتولوژی مشخص نکردند که آیا **blinding** صورت گرفته است. از ۳۱ گزارش کار آزمایی دوسوکوردر زنان و زایمان، فقط ۱۴ مورد شباهت رژیمهای درمانی و کنترل را نشان دادند. بعلاوه فقط ۵ مورد آشکارا بیان کردند که **blinding** موفق بوده است.

گاهی بجای **blinding** از اصطلاح **masking** استفاده می شود تا با وضعیت پزشکی نابینایی اشتباه نشود. اما **blinding** در معنای متدولوژیک خود در سرتا سر جهان شناخته شده و برای گزارش کار آزمایی های بالینی مورد قبول می باشد.

مورد b ۱۰: در صورت انجام **blinding**، موفقیت آن چگونه ارزیابی خواهد شد؟

همانگونه که ما به دنبال شواهدی از مخفی سازی می گردیم تا از انجام واقعا تصادفی تخصیص افراد اطمینان حاصل کنیم، به دنبال شواهدی از موفقیت **blinding** هم هستیم. اگرچه توصیف مکانیسم به کاررفته برای **blinding** می تواند چنین اطمینانی را حاصل کند، گاهی موفقیت **blinding** رامی توان با پرسش مستقیم از شرکت کنندگان، مراقبان یا ارزیابی کنندگان نتایج درمورد اینکه فکر می کنند چه درمانی اعمال شده است، ارزیابی نمود.

Prasad و همکارانش یک کار آزمایشی کنترل شده با پلاسبو در مورد آب نبات مکیدنی حاوی روی برای کاهش دوره علائم سرماخوردگی معمولی انجام دادند. آنها یک مطالعه جداگانه در داوطلبان سالم برای امتحان قابلیت مقایسه طعم آب نباتهای پلاسبو یا روی انجام دادند. همچنین از شرکت کنندگان در کار آزمایشی اصلی خواستند که سعی کنند تا بفهمند چه درمانی دریافت کردند. در پایان کار آزمایشی گزارش شد که ۵۶٪ دریافت کنندگان روی و ۲۶٪ از دریافت کنندگان پلاسبو به درستی گروه مربوط به خود را شناسایی کردند.

اساساً اگر blinding موفقیت آمیز بود، قدرت حدس دقیق شرکت کنندگان در مورد گروه خود نباید بهتر از شانس می بود. اما در عمل اگر شرکت کنندگان بتوانند مداخله مربوط به خود را بیش از میزان مورد انتظار از شانس، به درستی تشخیص دهند، ممکن است به معنای عدم موفقیت blinding نباشد.

مورد a ۱۱: روشهای آماری به کار رفته برای مقایسه گروهها از لحاظ پیامد اولیه.

داده ها ممکن است از راههای مختلفی قابل بررسی باشد که برخی از آنها احتمالاً در یک شرایط خاص کاملاً مناسب نیست. باید مشخص شود که کدام روش آماری برای هر آنالیز به کار می رود. و چگونه گزارش خواهند شد.

در روشهای استاندارد آنالیز فرض بر این است که اطلاعات مستقل هستند. در مورد کارآزمایی های کنترل شده این معمولاً بدین معناست که یک مشاهده به ازای هر شرکت کننده وجود دارد. اگر با مشاهدات متعدد از یک بیمار به صورت اطلاعات مستقل برخورد شود یک خطای جدی خواهد بود. چنین اطلاعاتی وقتی ایجاد می شوند که نتایج را به توان در قسمت های مختلف بدن اندازه گیری کرد مثل دندانپزشکی یا روماتولوژی آنالیز اطلاعات باید براساس فقط یکبار شمارش هر شرکت کننده باشد یا باید با استفاده از روشهای آماری پیچیده تر صورت گیرد. آنالیز نادرست مشاهدات متعدد در ۱۲۳ مورد از ۱۹۶ کارآزمایی مربوط به آتریت روماتوئید دیده شده است.

مورد b ۱۱: روشهای آنالیز اضافی مثل آنالیزهای زیرگروهی و آنالیزهای adjusted

روش آنالیز زیرگروهی هم مانند آنالیزهای اصلی باید به روشنی مشخص شود. قویترین آنالیزها آنهایی هستند که بر پایه جستجوی شواهدی از یک اختلاف در اثر درمانی در زیر گروه های مکمل باشند (مثلاً شرکت کنندگان مسن تر و جوانتر)، مقایسه ای که به آن ((امتحان اثرات متقابل)) گفته می شود. یک راهکار شایع ولی نامناسب مقایسه مقادیر p از آنالیزهای جداگانه اثر درمان در هر گروه است. این نادرست است که تاثیر درمانی یک زیرگروه از روی یک مقدار p معنی دار و یک p غیرقابل توجه استنباط شود. چنین استنباط هایی مقادیر مثبت کاذب بالایی دارند.

بخاطر خطر زیاد یافته های نادرست استفاده از آنالیزهای زیرگروهی اغلب توصیه نمی شود. مقایسه های تعقیبی (post hoc) (آنالیزهایی که پس از مشاهده اطلاعات انجام می شوند) بخصوص با مطالعات بیشتر به احتمال زیاد تایید نمی شوند چنین آنالیزهایی اعتبار زیادی ندارند.

در برخی مطالعات عدم تعادل در خصوصیات شرکت کنندگان (متغیرهای پیش بینی کننده) با استفاده از فرمی از آنالیز **multiple regression** جبران می شود. اگر چه نیاز به تطبیق در **RCT** ها بسیار کمتر از مطالعات اپیدمیولوژیک است، اما به خصوص اگر یک یا چند متغیر پیش بینی کننده به نظر مهم می رسند استفاده از یک آنالیز تطبیق داده شده منطقی خواهد بود. بطور ایده آل آنالیزهای تطبیق داده شده **adjusted** باید در پروتکل مطالعه مشخص شوند. مثلاً اغلب برای هرگونه متغیر طبقه بندی تطبیق توصیه می شود. در **RCT** تصمیم برای تطبیق نباید براساس اینکه آیا تفاوت های پایه، معنی آماری دارند یا نه گرفته شود. منطق هر آنالیز تطبیقی و روشهای آماری به کار رفته باید مشخص گردد. نویسندگان باید انتخاب متغیرهایی که تطبیق برای آنها صورت گرفته را مشخص کنند. بیان کنند که چگونه با متغیرهای پیوسته رفتار می کنند

نتایج

مورد a ۱۲: گردش شرکت کنندگان از هر مرحله (قویا یک نمودار توصیه می شود)

ریزش شرکت کنندگان از مطالعه در نتیجه عدم امکان پیگیری که اغلب اجتناب ناپذیر است، باید از حذف بیماران با تصمیم محققین به دلایلی چون واجد شرایط نبودن، عقب نشینی از درمان و عدم رعایت درست پروتکل کارآزمایی افتراق داده شود. اگر شرکت کنندگان از آنالیز حذف شوند ممکن است نتیجه گیری های اشتباهی صورت گیرد و عدم تعادل در چنین حذف های سی-بین گروه ها می تواند نشاندهنده **bias** باشد.

اگرچه **CONSORT** قویا استفاده از این ابزار گرافیکی را برای نمایش گردش شرکت کنندگان در طول مطالعه توصیه می کند اما یک فرمت اختصاصی مشخص وجود ندارد. قرار دادن نمودار در کارآزمایی های ساده ای که بدون موارد حذف شده در پیگیری یا حذف شده از مطالعه هستند ممکن است ضرورتی نداشته باشد.

مورد B ۱۲: توصیف انحرافات پروتکل از طرح اولیه مطالعه به همراه دلایل آن

نویسندگان باید اعلام نمایند که تمام انحرافات پروتکل در زمینه تغییرات برنامه ریزی شده در مداخله ها، آزمایشات، جمع آوری اطلاعات و روشهای آنالیز را گزارش خواهند کرد. برخی از این انحرافات پروتکل ممکن است در نمودار گردش گزارش شوند. مثلاً شرکت کنندگانی که مداخله مورد نظر را دریافت نکرده اند. اگر شرکت کنندگان پس از

تصادفی کردن بخاطر نداشتن معیارهای ورود از مطالعه حذف شده باشند، می توانند در نمودار گردش وارد شوند. فقط استفاده از کلمه انحراف پروتکل در مقالات چاپ شده برای توجیه خروج افراد پس از تصادفی کردن کافی نیست.

مورد ۱۳: تاریخهای بیانگر زمانهای جمع آوری نمونه ها و پیگیری باید ذکر شود.

دانستن زمان انجام یک مطالعه و دوره زمانی جمع آوری شرکت کنندگان اهمیت خاصی دارد. درمانهای پزشکی و جراحی و از جمله درمانهای همزمان بطور مداوم در حال ایجاد هستند و ممکن است بر درمان معمول داده شده به بیماران در طی یک کارآزمایی اثر بگذارند. دانستن سرعت جمع آوری شرکت کنندگان نیز بخصوص برای سایر محققان مفید است.

طول دوره پیگیری همیشه یک دوره ثابتی پس از تصادفی کردن مطالعه نیست. در بسیاری از RCTs پیگیری تمام شرکت کنندگان در یک تاریخ خاص پایان می پذیرد. این تاریخ باید ذکر شود و ذکر متوسط دوره پیگیری هم مفید است.

اگر کارآزمایی باید بسته به نتایج آنالیز بین مطالعه ای متوقف شود این باید اعلام گردد. توقف زود هنگام منجر به عدم تطابق بین حجم نمونه برنامه ریزی شده و واقعی می شود. بعلاوه کارآزمایی هایی که زود پایان می پذیرند احتمالاً اثر درمانی را بیش از حد تخمین می زنند.

مورد ۱۴: خصوصیات پایه ، دموگرافیک و کلینیکی هر گروه

اگر چه معیارهای ورود نمایانگر افراد واجد شرایط برای کارآزمایی است اما دانستن خصوصیات شرکت کنندگان در مطالعه هم مهم است و به خوانندگان گزارش، به خصوص کلینیسین ها اجازه می دهد قضاوت کند که نتایج کارآزمایی چقدر به یک بیمار خاص مربوط می شود.

اگر چه تخصیص تصادفی صحیح از bias در انتخاب جلوگیری می کند اما تضمین نمی کند که گروه ها در ابتدا با هم معادل باشند. اما هر تفاوتی در خصوصیات پایه در نتیجه شانس است نه bias. گروه های مطالعه باید از لحاظ خصوصیات پایه ای مهم، دموگرافیک و بالینی با هم مقایسه شوند تا خواننده ارزیابی کند که گروه ها چقدر شبیه بوده اند. اطلاعات پایه بخصوص وقتی متغیر پیامد در ابتدای کارآزمایی مهم قابل اندازه گیری است مهم هستند. روش ثبت و گزارش این اطلاعات پایه باید در پروتکل مزالعه ثبت و روشن شود.

وقتی اطلاعات پیوسته توزیع ناقص دارند روش ترجیحی ذکر میانه و یک دامنه درصدی است. خطاهای استاندارد و فواصل اطمینان برای توصیف تنوع مناسب نیستند چون آنها آمار استنباطی هستند نه توصیفی، با متغیرهایی که شامل

اعداد کوچکی از گروه های ترتیبی هستند (مثل مراحل بیماری از I تا IV) نباید بصورت متغیرهای پیوسته برخورد شود. در عوض اعداد و نسبتها برای هر گروه ترتیبی باید گزارش شود.

علیرغم هشدارهای متعدد در مورد نامناسب بودن آزمایش معنی دار بودن تفاوت‌های پایه- انجام آنها هنوز شایع است. بطور ایده آل در پروتکل کارآزمایی لزوم انجام تطبیق برای متغیرهای پایه اسمی با استفاده از آنالیز کوواریانس باید مشخص شود.

مورد ۱۵: تعداد شرکت کنندگان در هر گروه که در هر آنالیز وارد می شوند و اینکه آیا آنالیز مربوط به افراد با قصد درمان بوده است (intention to treat analysis)

عدم ورود تمام شرکت کنندگان در آنالیز ممکن است نتایج را دچار bias نماید. اما بیشتر کارآزمایی ها اطلاعات کامل ندارند. مشکلاتی در پروتکل ممکن است رخ دهد مثلا وقتی بیماران مداخله کامل یا صحیح را دریافت نمی کنند یا وقتی تعداد کمی بیماران غیر واجد شرایط به اشتباه بطور تصادفی در گروه ها قرار می گیرند. یک راه توصیه شده برای برخورد با این مسائل آنالیز تمام شرکت کنندگان بر طبق گروه اولیه ای است که به آن نسبت داده شده اند صرف نظر از این که بعدا چه رخ داده است. این راهکار intention-to-treat همیشه هم انجامش آسان نیست. بسیار پیش می آید که برخی بیماران مطالعه را کامل نکنند ممکن است از مطالعه خارج شوند یا از درمان فعال عقب نشینی کنند و بنابراین در پایان ارزیابی نمی شوند. اگر چه این شرکت کنندگان را نمی توان در آنالیز وارد کرد هنوز مرسوم است که به آنالیز تمام شرکت کنندگان در دسترس، آنالیز intention-to-treat گفته شود. وقتی برخی شرکت کنندگان که اطلاعات آنها در دسترس است از مطالعه خارج می شوند این کلمه به صورت نادرست استفاده می شود. مثلا بیمارانی که هیچکدام از درمانهای مورد نظر را به دلیل عدم رعایت پروتکل دریافت نکردند. برعکس آنالیز می تواند فقط به شرکت کنندگانی محدود شود که تمام شرایط پروتکل را از لحاظ واجد شرایط بودن مداخله ها و ارزیابی نتایج دارند. به این آنالیز on treatment یا per protocol گفته می شود. گاهی هر دو نوع آنالیز ارائه می شود. خارج کردن شرکت کنندگان از آنالیز می تواند سبب نتیجه گیری های اشتباه شود. آنالیز intention-to-treat معمولا مطلوب است چون از bias همراه با از دست رفتن غیرتصادفی شرکت کنندگان جلوگیری می کند.

مورد ۱۶: ارائه نتایج برای هر پیامد اولیه و ثانویه،

برای هر پیامدی نتایج مطالعه باید بصورت خلاصه ای از آن پیامد در هر گروه (مثلا نسبت شرکت کنندگان بیا بدون وقوع پیامد، یا میانگین و انحراف معیار اندازه گیری ها) همراه با تمایز بین گروه ها که به آن ((اندازه تأثیر)) گفته می شود گزارش گردد. این موارد باید به روشنی در پروتکل مطالعه توضیح داده شود

برای پیامدهای دوتایی اندازه اثر ممکن است نسبت خطر (خطر نسبی)،

odds ratio یا تفاوت خطر باشد. برای اطلاعات مدت زمان بقا اندازه اثر ممکن است hazard ratio یا تفاوت در زمان بقای میانه باشد و برای اطلاعات پیوسته اندازه اثر همان تفاوت در میانگین هاست. فواصل اطمینان باید برای تمایز بین گروه ها در نظر گرفته شود. یک اشتباه شایع ارائه فواصل اطمینان جداگانه برای پیامد در هر گروه بجای ارائه آن برای اثر درمانی است.

استفاده از فاصله اطمینان در سالهای اخیر بسیار افزایش یافته است اگر چه نه در تخصص های پزشکی. با وجود اینکه مقادیر P ممکن است علاوه بر فواصل اطمینان ارائه شوند، اما نتایج نباید فقط بصورت مقادیر P گزارش شوند. نتایج باید برای تمام پیامدهای اولیه و ثانویه طرح ریزی شده گزارش شود، نه فقط برای آنالیزهایی که از لحاظ آماری معنی دار می شوند.

مورد ۱۷: هرگونه آنالیز انجام شده از جمله آنالیزهای زیرگروهی و تطبیقی از قبل طراحی شوند

آنالیزهای متعدد داده های یکسان، خطریافته های مثبت کاذب را افزایش می دهد. نویسندگان باید در برابر وسوسه انجام آنالیزهای زیرگروهی متعدد مقاومت کنند. آنالیزهایی که در پروتکل تحقیق از قبل تعیین شده اند قابل اعتمادتر از آنالیزهای پیشنهاد شده با داده ها هستند. نویسندگان باید معلوم کنند که کدام آنالیزها از قبل طراحی شده اند.

مورد ۱۸: تمام وقایع یا عوارض جانبی نامطلوب مهم در هر گروه مداخله باید ذکر گردد.

اکثر مداخله ها علاوه بر اثرات مورد نظر عوارض جانبی ناخواسته و اغلب نامطلوب دارند. خوانندگان به اطلاعاتی در مورد ضررها و مزایای مداخله ها نیاز دارند تا بتوانند تصمیم های منطقی و متعادل بگیرند. وجود و ماهیت اثرات نامطلوب تاثیر زیادی بر میزان مفید و قابل قبول در نظر گرفته شدن یک مداخله خاص دارد. تمام وقایع نامطلوب مشاهده شده در حین یک کارآزمایی ضرورتا نتیجه مداخله نیستند. برخی ممکن است نتیجه وضعیتی باشند که تحت درمان قرار گرفته است. کارآزمایی های کنترل شده تصادفی بهترین راه را برای تامین ایمنی و کارآیی اطلاعات اطلاعات فراهم می کنند اگر چه نمی توانند اثرات بد نادر را بیابند. نویسندگان باید مشخص نمایند که

چگونه اثرات نامطلوب و عوارض مطالعه را ثبت و گزارش خواهند کرد.

بحث: