

فصل ۹- مطالعه‌های مداخله‌ای و کارآزمایی‌ها

دکتر محمد تقی یاسمی*، دکتر مسعود یونسیان**

*دانشیار روان‌پزشکی،

دانشگاه علوم پزشکی و خدمات بهداشتی-درمانی شهید بهشتی

**استادیار اپیدمیولوژی،

دانشگاه علوم پزشکی و خدمات بهداشتی-درمانی تهران

فهرست

۱۹۹	هدف‌های یادگیری.....
۱۹۹	زمینه‌های لازم برای ورود به موضوع.....
۱۹۹	مقدمه.....
۲۰۱	جایگاه علمی مطالعه‌های مداخله‌ای.....
۲۰۲	تقسیم‌بندی مطالعه‌های مداخله‌ای.....
۲۰۳	مراحل کارآزمایی بالینی.....
۲۰۳	مشکل‌های اخلاقی کارآزمایی بالینی.....
۲۰۴	ضرورت وجود گروه شاهد.....
۲۰۵	طرح‌های مختلف کارآزمایی بالینی.....
۲۰۸	تخصیص تصادفی.....
۲۰۹	کورسازی.....
۲۱۱	نمونه‌گیری.....
۲۱۱	ملاک‌های ورود و خروج.....
۲۱۲	تعیین کم‌ترین اندازه نمونه.....
۲۱۲	تجزیه و تحلیل داده‌ها.....
۲۱۳	الف- تجزیه و تحلیل داده‌های پایه.....
۲۱۳	ب- تجزیه و تحلیل داده‌های پی‌گیری.....

۲۱۴	جمع‌بندی
۲۱۴	پرسش‌های پایانی
۲۱۵	پاسخ پرسش‌های پایانی
۲۱۶	پیوست ۱: روش‌های تصادفی‌سازی
۲۱۸	پیوست ۲: تجزیه و تحلیل بقا
۲۲۴	منابع
۲۲۶	منابع برای مطالعه بیشتر

مطالعه‌های مداخله‌ای و کارآزمایی‌ها

هدف‌های یادگیری

- خواننده باید پس از مطالعه این فصل بتواند:
- ۱- ضرورت انجام مطالعه‌های مداخله‌ای را شرح دهد.
 - ۲- جایگاه علمی مطالعه‌های مداخله‌ای را در کشف روابط علت- معلولی شرح دهد.
 - ۳- تفاوت‌ها و شباهت‌های مطالعه‌های مداخله‌ای و مشاهده‌ای را بیان کند.
 - ۴- انواع رایج مطالعه‌های مداخله‌ای را نام ببرد.
 - ۵- موضوع‌های اخلاقی در انجام مطالعه‌های مداخله‌ای و راه‌حل‌های آنها را بازگو کند.
 - ۶- اهمیت وجود گروه شاهد^(۱) را بیان کند.
 - ۷- روش‌های مختلف تخصیص تصادفی^(۲) را با هم مقایسه کند.
 - ۸- اهمیت اندازه نمونه و نقش آن در تفسیر نتایج پژوهش را بیان کند.
 - ۹- مزیت‌ها و معایب مطالعه‌های مداخله‌ای را بیان کند.

زمینه‌های لازم برای مطالعه این فصل

- دانستن مطالب زیر به فهم موضوع این فصل کمک می‌کند:
- ۱- کلیات انواع مطالعه‌ها، و نقاط قوت و ضعف هر کدام؛
 - ۲- روش‌های تجزیه و تحلیل داده‌ها؛
 - ۳- محاسبه کم‌ترین اندازه نمونه لازم در مطالعه‌های تحلیلی.

مقدمه

در دهه ۱۳۷۰، به دنبال چند بررسی اپیدمیولوژیک، گروه دکتر عزیزی نشان داد که اهالی بسیاری از مناطق روستایی ایران دچار کمبود شدید ید هستند و این کمبود منجر به اختلال شدید در رشد جسمی و ذهنی بسیاری از کودکان و نوجوانان این مناطق شده است (Azizi et al, 1995). یکی از اقدام‌های مهم پیش‌گیرانه کمیته کشوری مبارزه با کمبود ید، تولید، توزیع، و مصرف نمک یددار در کل کشور بود (عزیزی و همکاران، ۱۳۷۵). راه حل دیگر، تزریق محلول روغنی یددار به دانش‌آموزان منطقه‌های کوهستانی و هیپراندمیک، و سپس استفاده از آن در مناطق دیگر بود. گروه دکتر عزیزی

تزریق محلول روغنی یددار را در کاهش درجه گواتر و افزایش غلظت هورمون‌های تیروئید، و اثر آن را بر پیش‌رفت بهره‌هوشی و توانایی روانی-حرکتی^(۱) در مقایسه با گروه شاهد مناسب مطالعه کرد و به نتایج جالب توجه رسید (Azizi et al, 1996؛ میرمیران و همکاران، ۱۳۷۶، ۱۳۷۸). مطالعه‌های یادشده که اثر مداخله‌ای درمانی را ارزیابی می‌کردند، نمونه‌هایی از کارآزمایی مناسب برای کشوری در حال توسعه هستند که در عین حال بازتاب بین‌المللی مناسبی نیز پیدا کرد.

در کنار کوشش‌های سنجیده و موفقیت‌مانند مورد بالا، پیاپی با ادعاهایی از نوع دیگر نیز در جامعه روبه‌رو می‌شویم. درمانگرانی (که بیشتر آنان دارای صلاحیت پزشکی هم نیستند) با رسانه‌ها مصاحبه می‌کنند و مدعی می‌شوند که درمان‌هایی خارق‌العاده کشف کرده‌اند. وقتی از آنان خواسته می‌شود شواهد و دلایل خود را ارائه کنند، حداکثر گواهی کتبی یا نوار ویدیویی تعدادی بیمار را ارائه می‌کنند که از نتیجه درمان اظهار رضایت کرده‌اند. واقعیت تلخ این است که چنین ادعاهایی محدود به افراد فاقد صلاحیت پزشکی نمی‌شود، بلکه گاهی حتی پزشکان حاذق، بدون آن که از اصول روش‌شناسی علمی برای تولید علم استفاده کرده باشند، نتیجه تجربه‌های بالینی خود را به عنوان کشف یا پژوهش مطرح می‌کنند.

تجربه بالینی ارزش ویژه خود را دارد. ممکن است این تجربه‌ها فرضیه‌ها^(۲) باشند؛ بدین معنا که اندیشه‌ها و فرضیه‌هایی را به وجود آورند که بعداً با استفاده از روش‌شناسی علمی می‌توان آنها را ارزیابی کرد.

از نظر تاریخی، پژوهش‌های مداخله‌ای در پزشکی پس از گذراندن مرحله‌هایی به شرایط کنونی رسیده است. در این راه، تحول از روخوانی کتاب‌های قدما به استفاده از تجربه بالینی بسیار مهم بود؛ ولی تحول مهم‌تر انتقال از تجربه بالینی به کارآزمایی بالینی شاهددار^(۳) بود. نخستین بار جیمز لیند^(۴)، در سال ۱۷۵۳، مبتلایان به اسکوربوت^(۵) را جداگانه گروه‌بندی کرد و به گروهی از آنان آب‌لیمو داد که نتیجه درخشانی داشت. در قرن هجدهم، ویلیام ویتترینگ^(۶) دیجیتالیس را در ۱۶۳ مورد به صورت شاهددار به کار برد و نتایج آن را در سال ۱۷۸۵ منتشر کرد. ادوارد جنر^(۷) در سال ۱۷۹۸، اولین کارآزمایی پیش‌گیرانه شاهددار را با مایه کوبی^(۸) به کار برد. پس از درک اهمیت گروه شاهد، معرفی روش تصادفی‌سازی^(۹) به وسیله فیشر در سال ۱۹۲۲، نقطه عطفی بسیار مهم بود. نقطه عطف عملی، کارآزمایی بالینی^(۱۰) استرپتومایسین^(۱۱) در سال ۱۹۴۶ بود و از آن پس، یعنی از دهه ۱۹۵۰، جامعه پزشکی غالباً در اعتبار درمان‌هایی که با کارآزمایی شاهددار دوسوکور^(۱۲) تأیید نشده باشند، تردید می‌کند. بنابراین، می‌توان گفت تنها نیم قرن از ارزیابی وسواسی دخالت‌های

1. psychomotor

3. controlled

5. scurvy

7. Edward Jenner

9. randomization

11. streptomycin

2. hypothesis-generator

4. James Lind

6. William Withering

8. vaccination

10. clinical trial

12. double-blind

درمانی گذشته است، ولی به نظر می‌رسد پیش‌رفت‌هایی که در نتیجه تحول روش‌شناسی^(۱) صورت گرفته است، بیشتر از تحول فن‌شناسی^(۲) به کشف روش‌های جدید درمان و پیش‌گیری کمک کرده است. با توجه به مرور تاریخی بالا، می‌توان گفت کسانی که به تجربه بالینی خود اهمیت نمی‌دهند، چند قرن، و کسانی که تنها آن را مبنا قرار می‌دهند، دست‌کم نیم قرن از علم روز عقب مانده‌اند. ممکن است پرسیده شود با توجه به این که در حال حاضر هر ساله هزاران کارآزمایی بالینی در مراکز پیش‌رفته پزشکی و پژوهشی دنیا انجام می‌شود، آیا لازم است پژوهشگر پزشکی در کشوری در حال توسعه نیز چنین پژوهش‌هایی انجام دهد؟

دلایل زیادی وجود دارد که دست‌اندرکاران پزشکی و بهداشت در کشورهای در حال توسعه را برای آموختن و به‌کاربردن روش‌شناسی کارآزمایی ترغیب می‌کند. کم‌ترین فایده آشنایی با روش‌شناسی کارآزمایی بالینی مصون‌سازی درمانگران در برابر استنباط نابجا از تجربه‌های محدود بالینی است. فایده دیگر آن مجهز شدن به ابزار تجزیه و تحلیل انتقادی برای ارزیابی درمان‌های نوین است، چه درمان‌های جدید وارداتی که از کشورهای پیش‌رفته صنعتی وارد می‌شوند، و چه ادعاهای «پزشکی دگرشیوه یا جای‌گزین»^(۳) که بیشتر آنها نیز وارداتی است و اخیراً به‌طور جدی مطرح شده است. مهم‌تر از همه، آشنایی با این روش‌شناسی فرصتی فراهم می‌آورد تا متخصص پزشکی یا بهداشت در کشور در حال توسعه دخالت‌های پیش‌گیرانه یا درمانی مناسب با شرایط کشور خود را طراحی کند.

جایگاه علمی مطالعه‌های مداخله‌ای

گرچه در کار بالینی بارها از اصطلاح علت یا علیت استفاده می‌شود، اثبات ارتباط علیتی در پزشکی و بهداشت بسیار دشوار است. برای بررسی ارتباط میان متغیر مستقل و متغیر وابسته دور روش عمده وجود دارد: روش مشاهده‌ای و روش تجربی (مداخله‌ای). در نوع مشاهده‌ای، هیچ‌گونه دخالتی در کار طبیعت صورت نمی‌گیرد و تغییر در متغیر یا متغیرها در ارتباط با متغیرهای دیگر، بدون دخالت پژوهشگر، مطالعه می‌شود. حال آن‌که در یک تجربه^(۴) پژوهشگر عمداً دخالت می‌کند و با ایجاد تغییر در متغیر مستقل اثر آن را بر متغیر دیگر (متغیر وابسته) بررسی می‌کند و در جریان کار مراقب است تا سایر متغیرها بر متغیر مستقل اثر نگذارند. مثلاً می‌توان تعدادی موش را به دو گروه تصادفی تقسیم کرد و یک گروه را در معرض یک ماده سرطان‌زا قرار داد و بروز بعدی سرطان را در دو گروه با هم مقایسه کرد. چنین پژوهشی را به دلایل اخلاقی در انسان نمی‌توان انجام داد، ولی می‌توان مطالعه‌ای مشاهده‌ای از نوع هم‌گروهی ترتیب داد و در آن گروهی از افراد را که سیگاری هستند با گروهی از افراد غیرسیگاری پی‌گیری کرد و بروز بعدی سرطان را در این دو گروه مقایسه کرد. نوعی خاص از مطالعه‌های مشاهده‌ای، تجربه طبیعی^(۵) نام دارد که گاهی امکان انجام آن فراهم

1. methodology

2. technology

3. alternative medicine

4. experiment

5. natural experiment

می‌شود. در تجربه طبیعی وقوع یک حادثه به‌عنوان متغیر مستقل، در شرایطی که در سایر متغیرهای مهم تفاوتی به‌وجود نیامده است، امکان مشاهده تأثیر آن را روی متغیر وابسته فراهم می‌آورد. مطالعه اثرات بمباران هیروشیما^(۱) و تابش‌های هسته‌ای چرنوبیل^(۲) از این نوع بوده است.

مطالعه‌های تجربی، از این نظر که برای تعیین نتیجه نیازمند پی‌گیری افراد در حال مطالعه هستند، به مطالعه‌های هم‌گروهی شباهت دارند؛ ولی در مطالعه‌های هم‌گروهی هیچ‌گونه دخالتی در متغیر مستقل صورت نمی‌گیرد. در عین حال، تجربه‌ها بهترین آزمون رابطه علت و معلول شناخته می‌شوند و از این نظر بر مطالعه‌های هم‌گروهی برتری دارند. یکی از مهم‌ترین ضعف‌های مطالعه‌های مشاهده‌ای، و از جمله هم‌گروهی، تأثیر مخدوش‌کننده‌ها است. از مزیت‌های مهم مطالعه‌های تجربی این است که پژوهشگر می‌تواند از پیش کلیه تأثیرهای مخدوش‌کننده^(۳) را، به‌ویژه از طریق تخصیص تصادفی، کنترل کند.

باید توجه داشت که منظور از مداخله در روش‌شناسی تنها تخصیص مواجهه^(۴) به‌عنوان بخشی از طرح مطالعه است، و نه روش تهاجمی^(۵) که در فرایندهای بالینی مطرح است. برای مثال، گرچه در «بررسی رابطه سن با عوارض آنژیوگرافی»^(۶)، مداخله (دست‌کاری) جدی روی بیمار انجام می‌شود، طراحی این مطالعه از نوع مشاهده‌ای است. در عوض، دخالت ساده آموزشی که در طرح «تأثیر آموزش در میزان اضطراب مادران بعد از اولین تشنج همراه تب کودکان» صورت گرفته است (کلاهی و همکاران، ۱۳۷۸) در چارچوب طرح‌های مداخله‌ای بوده است.

تقسیم‌بندی مطالعه‌های مداخله‌ای

مطالعه‌های مداخله‌ای را به صورت‌های گوناگون می‌توان تقسیم‌بندی کرد. آنچه در این جا مبنای تقسیم‌بندی است، واحد تصادفی‌سازی^(۷) است. اگر تصادفی‌سازی، درباره افراد بیمار انجام شود، به این مطالعه کارآزمایی بالینی^(۸) و اگر روی افراد سالم انجام شود به آن کارآزمایی میدانی^(۹) گفته می‌شود. نمونه‌ای از کارآزمایی میدانی، «بررسی نتیجه مایه کوبی فلج اطفال در دانش‌آموزان سالم» بود که به وسیله سالک^(۱۰) انجام شد. وقتی مبنای تصادفی‌سازی به جای فرد، گروه یا جمعیت باشد (یعنی جمعیت‌های مختلف به گروه‌ها تخصیص داده شوند)، به آن کارآزمایی اجتماعی^(۱۱) گفته می‌شود. در کارآزمایی اجتماعی نوعی دخالت در عموم یک جمعیت صورت می‌گیرد، مانند افزودن فلوراید^(۱۲) به آب یا یددار کردن نمک برای یک یا چند جامعه و مقایسه آن با یک یا چند جامعه شاهد. گذشته از این تقسیم‌بندی و جزئیات آن، کلیات مطالعه‌های مداخله‌ای در همه این مطالعه‌ها

- | | |
|---------------------|-------------------|
| 1. Hiroshima | 2. Chernobyl |
| 3. confounding | 4. exposure |
| 5. invasive | 6. angiography |
| 7. randomization | 8. clinical trial |
| 9. field trial | 10. Salk |
| 11. community trial | 12. fluoride |

یکسان است.^(۱)

مراحل کارآزمایی بالینی

از آن‌جا که کارآزمایی‌های بالینی به فراوانی برای بررسی تندرستی و اثربخشی داروها به کار می‌روند، برای این نوع مطالعه مراحل^(۲) گوناگون در نظر می‌گیرند:

مرحله اول: این مرحله، که مرحله پیش‌بالینی^(۳) است، در آزمایشگاه و معمولاً درباره حیوانات انجام می‌شود. هدف از این مرحله به دست آوردن اطلاعاتی درباره فارماکولوژی داروها و تعیین سمیت آنها به منظور آماده‌شدن برای مطالعه‌های انسانی است.

مرحله دوم: در این مرحله، مطالعه درباره انسان صورت می‌گیرد و خود به چهار فاز تقسیم می‌شود: **فاز I:** هدف از این فاز ارزیابی اولیه بی‌خطر بودن درمان و تحمل بیمار نسبت به آن است که درباره تعداد معدودی افراد (معمولاً بین ۲۰ تا ۱۰۰ نفر داوطلب) انجام می‌شود.

فاز II: در این فاز اثرهای بالقوه درمان بررسی می‌شود. ممکن است در این فاز چندین روش با هم مقایسه شود تا روشی که احتمالاً بهتر از بقیه است، برای بررسی بعدی انتخاب شود. این فاز نیز معمولاً درباره تعداد مشخصی از افراد (بین ۱۰۰ تا ۲۰۰ نفر) انجام می‌شود.

فاز III: این فاز، که یک مطالعه مداخله‌ای تمام‌عیار است، به منظور بررسی کلیه اثرها و پی‌آمدهای درمان در گروه‌های در حال مقایسه انجام می‌شود. معمولاً هر روش جدید با روش معیار مقایسه می‌شود، افراد به طور تصادفی تخصیص داده می‌شوند و تلاش می‌شود بیمار و درمانگر از نوع درمان بی‌اطلاع باشند. بخش عمده این فصل مربوط به همین فاز است. از این جا به بعد، کارآزمایی بالینی مترادف مطالعه مداخله‌ای در نظر گرفته می‌شود.^(۴)

فاز IV: یا فاز پس از فروش دارو، اثر درازمدت درمان را بررسی می‌کند و معمولاً پس از تأیید روش درمانی توسط مقام‌های مسئول انجام می‌شود.

مشکل‌های اخلاقی کارآزمایی‌های بالینی

با توجه به این که در این مطالعه‌ها دخالت‌هایی در افراد انسانی صورت می‌گیرد، بیشترین حساسیت اخلاقی نیز در این مطالعه‌ها وجود دارد. توجه به این نکته لازم است که این پژوهش‌ها را باید افرادی

۱- گاهی مبنای تصادفی‌سازی حیوانات آزمایشگاهی است. به این مطالعه‌ها نام کلی "مطالعه‌های تجربی" اطلاق می‌شود.

2. stages

3. preclinical

۴. از آن‌جا که مطالعه‌های مداخله‌ای، به طور عمده، برای مقایسه اثر درمان‌های متفاوت به کار گرفته می‌شوند، از این پس در این فصل، درمان معادل مداخله و بیمار معادل واحد نمونه‌گیری به کار می‌رود.

انجام دهند که صلاحیت علمی کافی دارند. در مواردی که گروه شاهد وجود دارد، همیشه نمی‌توان از دارونما^(۱) استفاده کرد. در مواردی که از پیش درمان مؤثری وجود دارد، داروی جدید را، به جای دارونما، با داروی قبلی مقایسه می‌کنیم. فقط در مواردی که هنوز درمان مؤثری کشف نشده است، یا تداوم بیماری برای مدت محدود هیچ‌گونه خطر فوری و بعدی برای بیمار نداشته باشد، می‌توان به‌طور موقت برای گروه شاهد از دارونما استفاده کرد. درمان باید بی‌خطر باشد، ولی چون درمان بی‌خطر وجود ندارد باید عوارض احتمالی را با وخامت بیماری سنجید و بیمار را از همه مسایل آگاه کرد. باید از تمام بیماران رضایت آگاهانه^(۲) گرفته شود و بدون گرفتن رضایت بیماران حق نداریم آنان را در پژوهش شرکت دهیم. گروه پژوهشی نسبت به عوارض فوری و بعدی درمان مسؤول است و باید آن را جبران کند. بیمار می‌تواند در هر زمان که بخواهد از پژوهش خارج شود و نمی‌توان هیچ‌گونه تعهدی برای ادامه کار از او گرفت.

با توجه به این که در بیشتر این پژوهش‌ها بیماران باید از سرشت درمان بی‌اطلاع^(۳) باشند، قبل از تخصیص^(۴) بیماران به گروه درمان یا شاهد، برای همکاری از آنان موافقت گرفته می‌شود و به آنان گفته می‌شود که به‌طور تصادفی در یکی از این گروه‌ها جای خواهند گرفت.

ضرورت وجود گروه شاهد

اگر در مطالعه مداخله‌ای، گروه شاهد وجود نداشته باشد و فقط گروه مداخله در مطالعه شرکت داده شود، پژوهشگر نمی‌تواند اطمینان یابد که آیا اثرات مشاهده‌شده ناشی از مداخله انجام شده بوده‌اند یا عواملی دیگر (از جمله بهبود یا بدتر شدن خودبه‌خودی) این اثرها را پدید آورده‌اند. به‌عنوان مثال، تا مدت‌ها پس از این که مشخص شد نوزادانی که نارس بوده‌اند و در زمان تولد به آنان اکسیژن با دوز بالا می‌دهند، احتمال بیشتری برای ابتلا به فیبروپلازی پشت عدسی^(۵) دارند، این عارضه با کورتیکوستروئید^(۶) درمان می‌شود، چرا که مشاهده شده بود ۷۵ درصد این بیماران به درمان پاسخ مناسب می‌دهند. تا این که مدت‌ها بعد کارآزمایی‌های بالینی با گروه شاهد نشان دادند که حتی بدون درمان نیز ۷۵ درصد این کودکان خودبه‌خود به سمت بهبود می‌روند.

عیب دیگر نبودن گروه شاهد این است که گاهی پژوهشگر برای این که امکان موفقیت درمان جدید را بیشتر کند، ناخودآگاه بیمارانی را که بیماری خفیف‌تری دارند برای درمان جدید انتخاب می‌کند (سوگرایی انتخاب)^(۷)؛ یا موفقیت درمان جدید را بزرگ‌تر می‌بیند و شکست آن را نادیده می‌گیرد (سوگرایی مشاهده‌گر)^(۸).

مشکلی که گاهی در مورد انتخاب گروه شاهد وجود دارد این است که برخی از پژوهشگران که اصرار دارند روش جدید قطعاً بهتر از روش قدیمی است (حتی پیش از آن که مطالعه شاهداری در

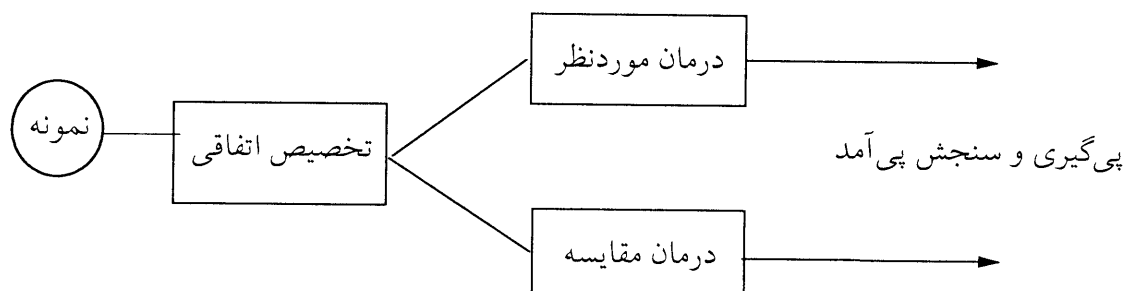
- | | |
|----------------------------|---------------------|
| 1. placebo | 2. informed consent |
| 3. blind | 4. allocation |
| 5. retrolental fibroplasia | 6. corticosteroid |
| 7. selection bias | 8. observer bias |

این زمینه صورت گرفته باشد) بیماران جدید خود را با روش جدید درمان می‌کنند و نتیجه آن را با بیماران قدیمی خود که قبلاً با روش معیار درمان می‌شده‌اند مقایسه کرده و نام این کار را مقایسه تاریخی^(۱) می‌گذارند. هر دو نوع مشکلی که دربارهٔ نبود گروه شاهد ذکر شد، در این جا نیز می‌تواند صادق باشد؛ یعنی از یک سو، بیمارانی که برای درمان جدید انتخاب می‌شوند از نظر شدت بیماری شبیه بیماران قبلی نباشند (بهتر باشند) و از سوی دیگر، ارزیابی نتایج درمان مشابه بیماران قبلی نباشد (بهتر ارزیابی شوند).

ساکس^(۲) و همکاران (Sacks et al, 1982) مطالعه‌هایی را که در مورد اثربخشی نوعی درمان خاص انجام شده و در آنها از تصادفی‌سازی یا کنترل تاریخی استفاده شده بود، مقایسه کردند. این مقایسه نشان داد مطالعه‌هایی که در آنها از شاهد تاریخی استفاده شده بود، به‌طور کلی، تمایل داشتند که اثر درمان را بهتر نشان دهند.

طرح‌های مختلف کارآزمایی بالینی

انواع گوناگونی از کارآزمایی بالینی وجود دارد که مهم‌ترین آنها را می‌توان در دو دسته طرح موازی^(۳) و طرح پی‌درپی^(۴) جای داد (Wingo et al, 1991). در طرح موازی، افراد به‌طور تصادفی در یکی از گروه‌های مداخله یا مقایسه قرار می‌گیرند و تا پایان مطالعه در همین گروه باقی می‌مانند و وضعیت آنان تجزیه و تحلیل می‌شود. ساختار این طرح را می‌توان به صورت زیر نشان داد:

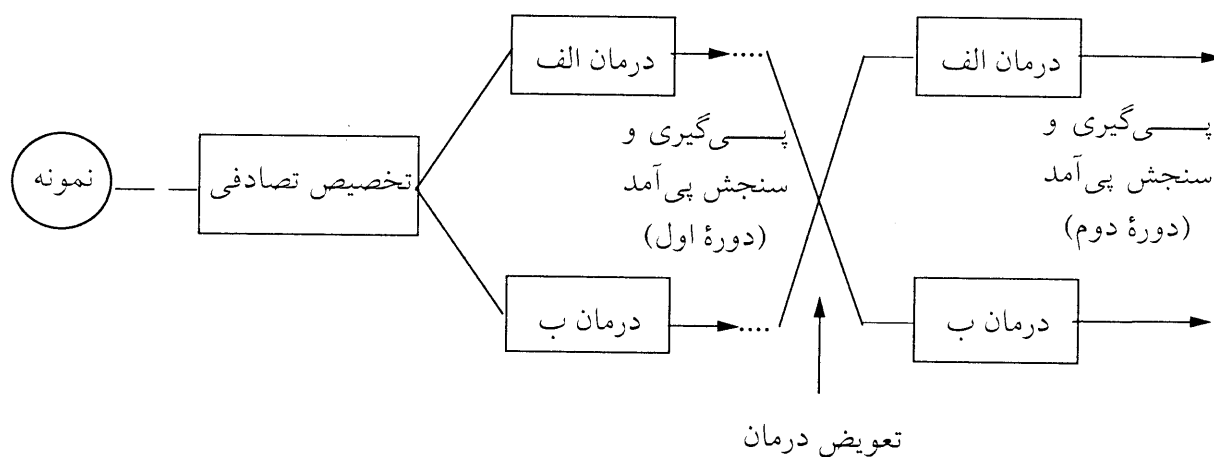


شکل ۱- طرح موازی در کارآزمایی بالینی

در طرح پی‌درپی، هر یک از افراد، در هر یک از گروه‌ها، از بیش از یک روش درمانی استفاده می‌کنند. معروف‌ترین نوع مطالعه پی‌درپی، طرح ضربدری^(۵) نام دارد که ساختار اصلی آن به صورت زیر است:

1. historical comparison
3. parallel design
5. crossover design

2. Sacks
4. successive design



شکل ۲- طرح پی در پی در کارآزمایی بالینی

همان‌گونه که مشاهده می‌شود، در دوره اول مطالعه، گروه اول از درمان الف و گروه دوم از درمان ب استفاده می‌کنند. پس از پی‌گیری کافی و بررسی پی‌آمد، درمان‌ها را عوض می‌کنند. گروه اول به روش ب و گروه دوم به روش الف درمان می‌شوند. در فاصله میان این دو دوره باید مدت کافی صبر کرد تا اثر درمان مرحله اول از بین برود (قسمت نقطه چین).

مزیت این طرح آن است که هر فرد شاهد خود به‌شمار می‌آید. این ویژگی از یک سو، باعث کاهش قابل توجه تعداد نمونه لازم می‌شود، و از سوی دیگر مقایسه پذیر بودن گروه‌های در حال درمان را تضمین می‌کند.

طرح رایج دیگری نیز وجود دارد که در آن هر فرد شاهد خود محسوب می‌شود. در این نوع طرح که طرح قبل - و - بعد^(۱) نامیده می‌شود، نخست متغیر وابسته اندازه‌گیری می‌شود و سپس مداخله مورد نظر در تمام افراد انجام می‌شود. سپس، مجدداً متغیر وابسته اندازه‌گیری می‌شود و مقدار (یا سطح یا حالت) متغیر وابسته پیش و پس از مداخله مقایسه می‌شود.

از آن‌جا که این گونه مطالعه‌ها دارای گروه شاهد واقعی نیستند، پژوهشگر نمی‌تواند با قطعیت اثر مشاهده شده را به درمان نسبت دهد و ممکن است عواملی دیگر نیز در این میان سهم داشته باشند. بنابراین، این نوع طرح، کارآزمایی بالینی واقعی نیست و جزو مطالعه‌های شبه‌تجربی^(۲) دسته‌بندی می‌شود. یک نوع از مطالعه‌های قبل - و - بعد، *ABA* نام گرفته است که بیشتر در پژوهش‌های علوم رفتاری مورد استفاده بوده است. در مرحله *A*، وضعیت پایه متغیر اندازه‌گیری می‌شود. در مرحله *B*، متغیر مستقل دخالت داده می‌شود و در مرحله بعدی *A*، بار دیگر، دخالت حذف می‌شود تا اثر متغیرهای فرعی به کم‌ترین میزان برسد. باید توجه داشت که طرح‌های قبل - و - بعد در بررسی بیماری‌هایی اجراشدنی است که سیری به نسبت ثابت داشته و نوسان‌های خودبه‌خودی نداشته باشند.

1. before and after

۲. مطالعه شبه‌تجربی (quasi-experimental) به مطالعه‌ای گفته می‌شود که در آن دست‌کم یکی از شرایط تخصیص اتفاقی یا گروه شاهد وجود نداشته باشد.

طرح متوالی:

در پزشکی زمانی استفاده از طرح متوالی^(۱) ضرورت می‌یابد که به علت عوارض جانبی خطرناک دارو یا احتمال تأثیر بسیار جدی درمان در بیماری‌های خطرناک، صلاح است که کم‌ترین تعداد بیمار در معرض پژوهش قرار بگیرند و در موارد تأثیر خوب درمان، بقیه بیماران از درمان اصلی محروم نمانند و نتیجه کارآزمایی نیز هرچه زودتر اعلام شود. در این موارد بدون الزام تعیین اندازه نمونه قطعی از پیش، از طرح متوالی استفاده می‌شود. با مشخص شدن نتیجه درمان هر فرد یا تعداد مشخصی از افراد، داده‌های موجود تجزیه و تحلیل می‌شوند و برحسب نتایج مشاهده شده یکی از تصمیم‌های زیرگرفته می‌شود:

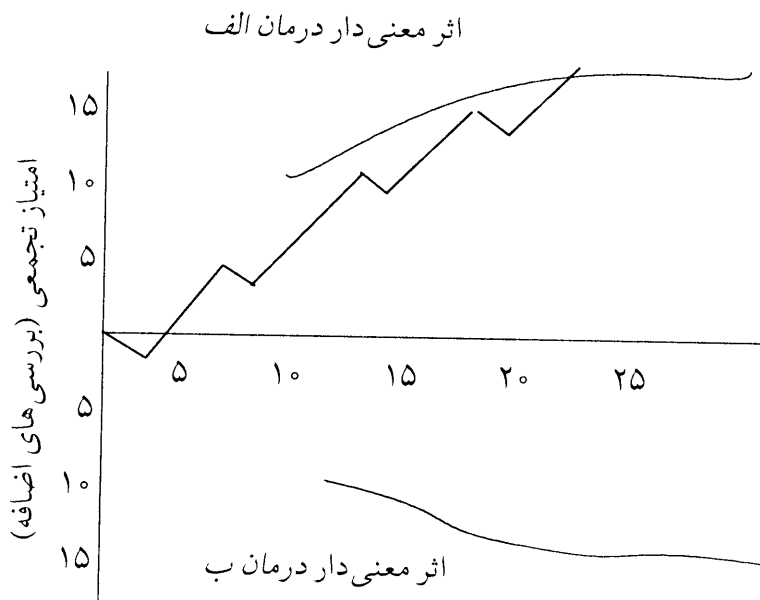
الف- میان درمان‌ها تفاوت معنی‌دار مشاهده می‌شود: در این صورت مطالعه در همین جا پایان می‌یابد.

ب- تفاوت معنی‌دار میان درمان‌ها مشاهده نمی‌شود ولی چون تعداد افراد شرکت‌کننده در مطالعه هنوز به اندازه‌ای نرسیده که قدرت (توان) مشاهده اختلاف بین روش‌های مختلف درمان (در صورت وجود) را داشته باشد، مطالعه ادامه داده می‌شود.

پ- تفاوت معنی‌دار میان درمان‌ها مشاهده نمی‌شود و چون مطالعه از توان کافی برای نشان دادن اختلاف (به فرض وجود) برخوردار است، مطالعه با نتیجه منفی پایان می‌یابد.

با استفاده از روش‌های آماری مانند توزیع دو جمله‌ای^(۲) می‌توان از پیش مرز تصمیم‌گیری را برای ادامه یا پایان درمان برای پژوهشگر تعیین کرد. به این منظور، معمولاً از نموداری مانند نمودار ۱ استفاده می‌شود. در این نمودار که برای مقایسه تأثیر دو داروی ضدسرفه کشیده شده مشخص شده که مطالعه در کدام شرایط، با نتیجه مثبت و در کدام شرایط با نتیجه منفی پایان داده می‌شود. محور افقی نشان‌دهنده زوج‌هایی از بیماران است که در آنها یکی از درمان‌ها بر دیگری برتری داشته و محور عمودی جهت آن را (غلبه درمان الف در جهت مثبت و غلبه درمان ب در جهت منفی) نشان می‌دهد. به این ترتیب، هر زوج از بیماران که در آن یکی از درمان‌ها بر دیگری غلبه داشته باشد یک واحد به سوی راست و یک واحد به سوی بالا یا پایین (برحسب این که کدام درمان برتری داشته باشد) حرکت کرده و نقطه موردنظر در شکل معین می‌شود. چنانچه خطی که به این ترتیب کشیده می‌شود از کناره‌های بالایی یا پایینی شکل عبور کند، اختلاف بین درمان‌ها معنی‌دار تلقی می‌شود و مطالعه پایان می‌یابد. چنانچه خط موردنظر از این کناره‌ها عبور نکند و به خط عمودی سمت راست شکل برسد، مطالعه منفی تلقی شده و پایان داده می‌شود.

همان‌گونه که مشاهده می‌شود در این مطالعه پس از آن که از ۱۷ مورد برتری، ۱۱ مورد برتری درمان الف بر ب مشاهده شد، تفاوت دو روش معنی‌دار تلقی گردید و به مطالعه پایان داده شد. البته این نوع طراحی برای مطالعه‌هایی که در آنها اثر درمان دیر ظاهر می‌شود، مناسب نیست. توصیه می‌شود برای به دست آوردن اطلاعات بیشتر در این مورد به منابع معتبر آماری مراجعه کنید.



نمودار ۱- نمودار تحلیل متوالی برای مقایسه دو درمان الف و ب

تخصیص تصادفی^(۱)

مجله علمی *British Medical Journal* از سال ۱۹۹۱ تصمیم گرفت، جز در مواردی خاص مانند بیماری‌های بسیار نادر، تنها کارآزمایی‌هایی را منتشر کند که در آنها تصادفی‌سازی به‌طور مناسب انجام شده باشد (Altman, 1999).

اگر در مطالعه‌ای درمانگر تصمیم بگیرد کدام درمان را به کدام بیمار تخصیص دهد، این احتمال وجود دارد که ویژگی‌های بالینی و جمعیت‌شناختی بیمارانی که در دو یا چند گروه قرار می‌گیرند، به‌طور منظم^(۲) متفاوت باشند. از آن‌جا که پزشک یا درمانگر نمی‌خواهند بیماران شرکت‌کننده در پژوهش زیان ببینند، این احتمال وجود دارد که، ناخودآگاه، درمان جدید را به بیمارانی که بیماری خفیف‌تری دارند یا، در مجموع، شرایط بالینی مناسب‌تری دارند اختصاص دهند. اگر خود بیمارانی تصمیم بگیرند که با کدام روش درمان شوند باز هم امکان بروز چنین خطایی وجود خواهد داشت (سوگرایی خود-انتخابی)^(۳). برای مثال، در مطالعه‌ای که درباره‌ی اثر مکمل ویتامینی در پیش‌گیری از بروز نقایص عصبی انجام شده بود (Smith et al, 1980)، گرچه کاهش چشم‌گیر بروز نقایص عصبی در نوزاد مادرانی که از مکمل ویتامینی استفاده می‌کردند، مشاهده شده بود، به‌دلیل این که تخصیص تصادفی انجام نشده بود، نتایج این مطالعه چندان مورد توجه قرار نگرفت؛ به‌گونه‌ای که بعدها شورای پژوهش پزشکی^(۴) بودجه‌ای درخور برای انجام مطالعه‌ی کارآزمایی بالینی بزرگ و پاسخ‌دادن

۱. random allocation یا randomization (هر دو به کار رفته است).

2. systematic

3. self-selection bias

4. Medical Research Council

به این پرسش در نظر گرفت.

توجه داشته باشید که تخصیص تصادفی با نمونه‌گیری تصادفی فرق دارد. گرچه در مطالعه مقطعی باید نمونه‌ها به گونه‌ای انتخاب شوند که نماینده کل افراد جامعه باشند، در مطالعه‌های مداخله‌ای چنین ضرورتی وجود ندارد؛ بلکه پس از این که افراد انتخاب شدند، تقسیم آنها به دو گروه باید کاملاً تصادفی باشد. در واقع، حتی اگر انتخاب افراد برای شرکت در مطالعه مداخله‌ای به صورت تصادفی انجام شده باشد، پس از اعمال ملاک‌های ورود و خروج، دیگر، افراد شرکت‌کننده نمونه گویایی از جامعه نخواهند بود. با توجه به این که تعدادی از بیماران در طی مطالعه از آن خارج می‌شوند (که هم به توان مطالعه آسیب می‌رساند و هم با احتمال سوگرایی همراه است)، ترجیح داده می‌شود برای مدتی محدود به همه بیماران درمان یا داروی مشابه داده شود تا شرکت‌کنندگان پروپاقرص مشخص شوند. در تخصیص تصادفی، پژوهشگر متعهد می‌شود که احتمال حضور هر بیمار در هر یک از گروه‌ها معین (معمولاً مساوی) باشد، ولی نمی‌توان از پیش تعیین کرد که کدام بیمار با کدام روش درمان خواهد شد. ساده‌ترین روش تصادفی‌سازی استفاده از سکه (شیر یا خط) یا استفاده از جدول اعداد تصادفی است. برای تصادفی‌سازی می‌توان از روش‌هایی پیش‌رفته‌تر مانند تقسیم تصادفی با طبقه‌بندی^(۱) یا تقسیم تصادفی بلوکی^(۲) نیز استفاده کرد.

زمانی از تقسیم تصادفی بلوکی استفاده می‌شود که پژوهشگر بخواهد تعداد افراد در گروه‌های درمانی، در هر زمان، تقریباً یکسان باشند و با استفاده از طبقه‌بندی سعی می‌کند، تا آن‌جا که می‌توان، توزیع متغیرهای مهمی که با پی‌آمد رابطه دارند را در گروه‌های در حال مقایسه مشابه سازد (برای اطلاع بیشتر در مورد روش‌های تخصیص تصادفی به پیوست ۱ مراجعه کنید).

توجه داشته باشید که روش‌هایی مانند تقسیم‌بندی بیماران برحسب روز مراجعه (زوج یا فرد) یا رقم سمت راست شماره پرونده و... برای تصادفی‌سازی مناسب نیستند؛ زیرا پژوهشگر در زمان ورود فرد به مطالعه می‌داند بیمار در کدام گروه درمانی جای خواهد گرفت و این آگاهی می‌تواند بر تصمیم به شرکت دادن یا شرکت ندادن او در مطالعه مؤثر باشد (Altman, 1991).

کورسازی^(۳)

در طراحی یک مطالعه مداخله‌ای، باید کاری کرد که شرکت‌کنندگان از نوع مداخله‌ای که بر آنان صورت خواهد گرفت بی‌اطلاع باشند؛ در غیراین صورت، آگاهی از این که مداخله مؤثر یا مداخله بی‌اثر روی فرد صورت گرفته ممکن است بر پاسخ فرد به مداخله تأثیر بگذارد. این مشکل، گرچه در مورد پی‌آمدهای عینی (مانند اندازه‌گیری فشارخون با دستگاه) نیز اهمیت خاص خود را دارد و باید رعایت شود، در مورد پی‌آمدهایی که ماهیت ذهنی دارند (مانند احساس درد) اهمیت بیشتری پیدا می‌کند. علت این مسأله آن است که بسیاری از پاسخ‌های فیزیولوژیک و عضوی بدن نیز زیر تأثیر

1. stratified randomization

2. block randomization

3. blinding

پدیده‌های روان‌تنی^(۱) قرار دارند. خون‌ریزی از زخم گوارشی، التیام زخم دوازدهه، پیش‌گیری از سرماخوردگی، کاهش دفعات عود تب‌خال، کنترل دیابت، سیر آنژین صدری، اندازه‌کلسترول خون، و حتی مرگ‌ومیر بیماران می‌تواند تحت تأثیر دارونما قرار گیرد (Brown, 1998؛ یاسمی، ۱۳۷۲). مطالعه‌ای که در آن فرد شرکت‌کننده از نوع مداخله بی‌اطلاع باشد، یک سوکور^(۲) نامیده می‌شود. از سوی دیگر، پژوهشگری که بررسی پی‌آمدها و پی‌گیری بیماران را انجام می‌دهد نیز نباید از نوع مداخله آگاهی داشته باشد. این امر می‌تواند در برداشت وی از رخداد پی‌آمد (به‌ویژه در موارد مرزی) یا تصمیم‌گیری درباره‌ی خارج کردن فرد از مطالعه (مثلاً به دلیل رخداد عوارض) تأثیر بگذارد. یک راه برای مقابله با این مشکل آن است که انجام مداخله و ارزیابی پی‌آمد را دو فرد متفاوت انجام دهند. مطالعه‌هایی را که در آنها هم شرکت‌کننده و هم ارزیابی‌کننده از نوع مداخله بی‌اطلاع باشند دوسوکور^(۳) می‌نامند.

فراوانی آزمون‌ها و روش‌های آماری و تفاوت‌های هرچند جزئی در نتایج این روش‌ها و آزمون‌ها باعث می‌شود که حتی دید و داوری شخص آمارگر نیز، در صورت کنترل‌نشدن، بر نتیجه مطالعه تأثیر بگذارد. بنابراین ترجیح داده می‌شود، علاوه بر شرکت‌کننده و ارزیابی‌کننده، آمارگر نیز از نوع مداخله بی‌اطلاع باشد. این نوع مطالعه را سه سوکور^(۴) می‌نامند.

باید توجه داشت که گاهی کورسازی در عمل ممکن نیست. مثلاً، در مطالعه‌ای که اثر روش جراحی را با روش دارویی در درمان بیماری خاصی مقایسه می‌کند، به سختی می‌توان شرکت‌کننده یا حتی ارزیابی‌کننده را از نوع مداخله بی‌اطلاع نگاه داشت؛ گرچه باز هم می‌توان آمارگر را نسبت به این امر بی‌اطلاع نگه داشت.

کورسازی در کارآزمایی‌ها روشی خلاق است و پیروی از الگوی واحد برای این کارگمراه‌کننده است. بسیاری وقت‌ها با وجود آن‌که ظاهر دارو و دارونما یکسان است، به علت عوارض جانبی دارو (مثلاً خشک شدن دهان)، بیمار متوجه می‌شود که به او داروی اصلی داده شده است. در این موارد پژوهشگران ماده‌ای به دارونما می‌افزایند، که بدون تأثیر درمانی اختصاصی، این اثر را تقلید کند. به این نوع دارونما دارونمای فعال^(۵) می‌گویند. در مورد روش‌های غیردارویی مانند جراحی، روان‌درمانی، فیزیوتراپی و... می‌توان از درمان کاذب^(۶) استفاده کرد ولی بیشتر وقت‌ها نتیجه روش‌های درمانی گوناگون با هم مقایسه می‌شود.

با توجه به مطالبی که در مورد تخصیص تصادفی و کورسازی در کارآزمایی‌ها گفته شد، اگر در یک کارآزمایی امکان انجام این دو اقدام باشد و انجام نشود، اساساً آن پژوهش چندان قابل اعتنا نخواهد بود.

ساکس و همکاران در سال ۱۹۸۲، شش درمان گوناگون را در نظر گرفتند و مقاله‌هایی را که درباره‌ی آنها منتشر شده بود مرور کردند. پژوهش‌هایی که با تخصیص تصادفی صورت گرفته بودند، تنها در ۲۰ درصد موارد به تفاوت معنی‌دار دست یافته بودند؛ درحالی‌که پژوهش‌هایی که بدون تخصیص

1. psychosomatic
3. double-blind
5. active placebo

2. single-blind
4. tripple-blind
6. pseudotherapy or sham treatment

تصادفی صورت گرفته بودند، در ۷۹ درصد موارد به تفاوت معنی‌دار رسیده بودند. چالمرز^(۱) و همکاران، در مقایسه مشابهی، ۱۴۵ مقاله را بررسی کردند و به این نتیجه رسیدند که ۵۸ درصد پژوهش‌هایی که در آنها تخصیص تصادفی انجام نشده بود، ۲۴ درصد پژوهش‌های تصادفی ولی غیرکور، و تنها ۹ درصد با رعایت همه شرایط یعنی پژوهش‌هایی که هم کور و هم تصادفی بوده‌اند تأثیر درمانی را نشان داده‌اند. به بیان دیگر، هرکس که تنها به قاضی برود راضی برمی‌گردد.

نمونه‌گیری

در مطالعه مداخله‌ای نیز، همانند مطالعه مشاهده‌ای، نمونه‌گیری از اهمیت ویژه برخوردار است. اصول کلی نمونه‌گیری و تعیین اندازه نمونه لازم برای مطالعه مداخله‌ای مانند مطالعه هم‌گروهی است و توصیه می‌شود به آن فصل نیز مراجعه کنید.

ملاک‌های ورود و خروج

پژوهشگر ابتدا باید زیرگروهی از جامعه که می‌خواهد نتیجه پژوهش خود را به آن تعمیم دهد مشخص کند و نمونه خود را از این گروه انتخاب کند. برای این کار، پژوهشگر نخست مشخص می‌کند افراد باید چه ویژگی‌هایی داشته باشند تا بتوانند وارد مطالعه شوند. به این ویژگی‌ها، ملاک‌های ورود^(۲) گفته می‌شود. یکی از مهم‌ترین ملاک‌های ورود ملاک‌های تشخیصی^(۳) است. سپس افرادی که ویژگی‌های خاصی دارند از مطالعه حذف می‌شوند. به این ویژگی‌ها ملاک‌های خروج^(۴) گفته می‌شود.

حذف افراد به این علت انجام می‌شود که کسانی که هر کدام از درمان‌ها برایشان ضرر داشته باشد و کسانی که به هر دلیل امکان به پایان رساندن مطالعه را ندارند پیش از تصادفی‌سازی حذف شوند. در واقع، اگر ملاک‌های خروج را به کار نبریم و بعد از تصادفی‌سازی مشخص شود که فردی ویژگی لازم برای استفاده از درمان تخصیص داده شده را ندارد، یا درمان موردنظر برایش خطرناک است و از مطالعه حذف شود، امکان بروز سوگرایی در مطالعه وجود دارد و کار پژوهشگر را دشوار می‌سازد. برای مثال، در مطالعه‌ای که فتوحی و همکاران درباره «اثرات کاهنده چربی خون داروی گیاهی آنتوم» انجام دادند، ملاک‌های ورود به مطالعه سن بیشتر یا برابر با ۱۵ سال، کلسترول خون بیشتر از ۲۴۰ mg/dl در دو نوبت متوالی به فاصله دو هفته، و تری‌گلیسرید خون کم‌تر از ۲۵۰ mg/dl بود (فتوحی و همکاران، ۱۳۸۱). از این میان، افرادی که هر یک از ویژگی‌های قند خون ناشتای بیشتر از ۱۴۰ mg/dl یا سابقه دیابت، نارسایی قلبی، سابقه سکته قلبی یا مغزی در شش ماه گذشته، بیماری فعال کبدی، نارسایی کلیوی، زخم گوارشی، سابقه نقرس^(۵) یا اسید اوریک بیش از ۸ mg/dl، سابقه

1. Chalmers

2. inclusion criteria

3. diagnostic criteria

4. exclusion criteria

5. gout

آب‌سیاه^(۱)، بارداری، شیردهی، تداخل با داروهای خاص، و مصرف داروهای کاهنده چربی خون در شش ماه گذشته را داشتند از مطالعه حذف شدند. البته باید توجه داشت که هرچه ملاک‌های خروج بیشتر باشند عملاً زیرگروه ویژه‌تری از افراد انتخاب می‌شوند و تعمیم‌پذیری نتایج کاهش می‌یابد. به عبارت دیگر، با حذف افرادی که احتمال می‌دهیم نتوانند مطالعه را به پایان برسانند، پیش از آن که تخصیص تصادفی صورت گرفته باشد، احتمال سوگرایی را کاهش داده و روایی درونی^(۲) مطالعه را افزایش می‌دهیم که این امر با کاهش تعمیم‌پذیری یا کاسته شدن از روایی بیرونی^(۳) همراه است. باید تعادل مناسبی بین این دو برقرار شود، ولی اول باید به نتیجه‌درستی برسیم تا بعد درباره تعمیم‌پذیری آن داوری کنیم. بنابراین روایی درونی بر روایی بیرونی تقدم دارد.

تعیین کمترین اندازه نمونه

همان‌گونه که در ابتدای مبحث نمونه‌گیری اشاره شد، روش تعیین اندازه نمونه در مطالعه‌های مداخله‌ای شبیه مطالعه‌های هم‌گروهی است و خواننده برای اطلاع از دستورهای لازم در این زمینه به فصل مطالعه‌های هم‌گروهی ارجاع داده می‌شود. اهمیت و حساسیت اندازه نمونه در مطالعه‌های مداخله‌ای بیشتر است. از یک سو، در این نوع مطالعه به دلیل هزینه زیاد باید توان آماری^(۴) (اندازه نمونه) بیشتری داشته باشیم تا منابع را هدر ندهیم؛ و از سوی دیگر، پژوهشگر به دلیل محدودیت بودجه و ملاحظه‌های اخلاقی مجبور است کار خود را با کم‌ترین تعداد نمونه ممکن به انجام رساند. بنابراین، مسأله اندازه نمونه در مطالعه‌های مداخله‌ای از حساسیت ویژه برخوردار است و در عمل بسیاری از این مطالعه‌ها از نداشتن تعداد کافی نمونه رنج می‌برند. برای مثال، در بررسی ۷۱ کارآزمایی بالینی منفی^(۵) مشخص شد که تنها چهار مطالعه از این مطالعه‌ها توان آماری بیش از ۹۰ درصد داشته‌اند (Wingo et al, 1991).

تجزیه و تحلیل داده‌ها

داده‌های مطالعه مداخله‌ای در دو مرحله تجزیه و تحلیل می‌شوند:

الف- تجزیه و تحلیل داده‌های پایه؛

ب- تجزیه و تحلیل داده‌های پی‌گیری.

1. glaucoma

2. internal validity

3. external validity

۴. منظور از توان (power) یک مطالعه، احتمال کشف اختلاف به فرض وجود آن است. توان یک مطالعه به طور مستقیم با تعداد نمونه و به طور معکوس با اندازه اختلاف موردنظر رابطه دارد (برای اطلاعات بیشتر به فصل خطاهای پژوهش و نحوه برخورد با آنها مراجعه کنید).

۵. منظور از کارآزمایی بالینی منفی، مطالعه‌ای است که در آن اختلافی بین مداخله‌ها مشاهده نشده است.

الف- تجزیه و تحلیل داده‌های پایه

گرچه انتظار می‌رود ویژگی‌های افراد شرکت‌کننده در هر یک از گروه‌های درمانی، پس از تخصیص تصادفی، مشابه یکدیگر باشد، ولی همیشه این احتمال وجود دارد که از نظر برخی ویژگی‌های مهم مرتبط با درمان، که عوامل پیش‌آگهی^(۱) نامیده می‌شوند، نیز بین گروه‌ها اختلاف وجود داشته باشد. بنابراین، پژوهشگر باید ابتدا این ویژگی‌ها را در گروه‌های مداخله و مقایسه با یکدیگر مقایسه کند و چنانچه از این نظر تفاوتی در گروه‌ها مشاهده کرد، از روش‌های مناسب آماری برای برخورد با آنها استفاده کند. جزئیات مربوط به این امر مشابه تجزیه و تحلیل مطالعه مشاهده‌ای است و توصیه می‌شود خواننده برای آگاهی بیشتر به فصل تجزیه و تحلیل داده‌ها یا منابع مناسب (که در پایان همین فصل آورده شده است) مراجعه کند.

گرچه برای مقایسه داده‌های پایه نیز مرسوم است که از روش‌های آماری (انجام آزمون‌های آماری) استفاده شود، باید دانست نیازی به این کار وجود ندارد و داور چشمی در این مورد کفایت می‌کند (Wingo et al, 1991).

ب- تجزیه و تحلیل داده‌های پی‌گیری

برای تجزیه و تحلیل داده‌های پی‌گیری (که بخش اصلی تجزیه و تحلیل داده‌های مطالعه مداخله‌ای است) باید اندازه (های) پی‌آمد (های) از پیش تعیین شده را در گروه‌های مداخله و مقایسه آزمون آماری کنیم. به طور خلاصه، اگر گروه‌های در حال مقایسه مستقل از یکدیگر باشند (همسان‌سازی صورت نگرفته باشد)، برای مقایسه پی‌آمدهای کمی از آزمون مقایسه میانگین‌ها، و برای مقایسه پی‌آمدهای کیفی از آزمون مجذور کای^(۲) می‌توان استفاده کرد. در صورتی که همسان‌سازی انجام شده باشد، می‌توان از آزمون مقایسه میانگین‌های جور شده^(۳) برای پی‌آمدهای کمی، و از آزمون مک‌نمار^(۴) برای پی‌آمدهای کیفی استفاده کرد. جزئیات بیشتر این آزمون‌ها، و نیز معادل‌های غیر پارامتری^(۵) آنها در فصل تجزیه و تحلیل داده‌ها توضیح داده شده ولی به سبب اهمیت تحلیل بقا^(۶) در این نوع مطالعه‌ها، ضرورت انجام این نوع تجزیه و تحلیل و نیز معرفی یک روش کاربردی آسان برای انجام آن در پیوست ۲ همین فصل مورد بحث قرار گرفته است و علاقه‌مندان می‌توانند به آن مراجعه کنند. توجه به دو نکته مهم در تجزیه و تحلیل داده‌های پی‌گیری ضرورت دارد.

۱- تجزیه و تحلیل بر مبنای گروه‌های تخصیص داده شده. با وجود تلاش پژوهشگر برای این‌که افراد در مطالعه باقی بمانند، بیماران آزادند هر زمان که خواستند از مطالعه خارج شوند یا نوع درمان خود را عوض کنند. اگر پژوهشگر این افراد را به سادگی از مطالعه حذف کند احتمال بروز سوگرایی وجود

1. prognostic factors

2. chi-square test

3. paired t-test

4. McNemar's test

5. non-parametric

6. survival analysis

دارد، زیرا ممکن است بیماران به دلایلی (مثلاً بهبود قطعی یا بهبود نیافتن حتی به‌طور نسبی، یا عوارض جانبی) نوعی از درمان را ترک کنند؛ بنابراین باید در مطالعه منظور شوند. مطمئن‌ترین روش برای پرهیز از بروز سوگرایی در این موارد آن است که در زمان انتخاب افراد برای شرکت در مطالعه، افرادی انتخاب شوند که احتمال بیشتری برای به‌پایان رساندن مطالعه داشته باشند. ولی وقتی عدول از منشور پژوهش^(۱) رخ داد پژوهشگر باید همه افراد را (اعم از این که دوره درمان تخصیص داده شده را به‌طور کامل طی کرده باشند، به‌طور ناقص درمان شده باشند، یا حتی آن را تغییر داده و با روشی دیگر درمان شده باشند) تا پایان مطالعه پی‌گیری کند، پی‌آمد مورد نظر را در آنان بسنجد و در گروهی که در ابتدای مطالعه در آن قرار داشتند تجزیه و تحلیل کند. این روش، که تحلیل موارد با قصد درمان^(۲) نامیده می‌شود مورد نظر بیشتر صاحب نظران در زمینه مطالعه‌های مداخله‌ای است.

۲- توجه به دیگر متغیرهای مهم که با پی‌آمد رابطه دارند در زمان تجزیه و تحلیل (پیش از این اشاره شد).

جمع‌بندی

- ۱- گرچه تجربه‌های بالینی نقشی مهم در گسترش علم دارند، مطالعه‌های مداخله‌ای (با توجه به ماهیت آنها که دارای گروه شاهد و زیر کنترل پژوهشگر هستند) جایگاهی ویژه در توسعه علم بهداشت و پزشکی دارند، به‌گونه‌ای که قوی‌ترین نوع مطالعه‌ها برای بررسی رابطه علت- معلولی هستند.
- ۲- در تقسیم‌بندی مطالعه‌ها به مداخله‌ای و مشاهده‌ای باید به تفاوت میان مداخله^(۳) و تهاجم^(۴) توجه کافی شود.
- ۳- بزرگ‌ترین مشکل در انجام مطالعه‌های مداخله‌ای، ملاحظه‌های اخلاقی آنهاست.
- ۴- تخصیص تصادفی مهم‌ترین رکن مطالعه مداخله‌ای است که آن را از مطالعه مشاهده‌ای متمایز می‌کند. چنانچه تخصیص تصادفی به‌شکل شایسته‌ای انجام نشود، ممکن است کارآزمایی بالینی تا حد مطالعه هم‌گروهی تنزل یابد. در این صورت برای برخورد با این مشکل باید از روش‌های مناسب تجزیه و تحلیل مطالعه‌های مشاهده‌ای استفاده کرد.

پرسش‌های پایانی

- ۱- آیا آشنایی با روش‌شناسی کارآزمایی بالینی برای دست‌اندرکاران بهداشت و پزشکی کشورهای در حال توسعه ضرورت دارد؟ برای پاسخ خود چه دلایلی دارید؟
- ۲- به چه دلیل مطالعه‌های مداخله‌ای در کشف روابط علت و معلولی قوی‌تر از مطالعه‌های مشاهده‌ای‌اند؟

1. protocol

2. intention-to-treat analysis

3. intervention

4. invasion

- ۳- مطالعه‌های مداخله‌ای به کدام نوع مطالعه مشاهده‌ای نزدیکی بیشتری دارند؟
- ۴- دو دسته اصلی طرح‌های مداخله‌ای را از نظر نوع طراحی نام ببرید.
- ۵- میان گرفتن رضایت آگاهانه و غیرمطلع^(۱) نگاه داشتن بیمار از نوع دخالت، از نظر اخلاقی، تناقضی وجود دارد. این تناقض را چگونه حل می‌کنند؟
- ۶- اگر پژوهشگری دارویی را در گروهی از بیماران به کار ببرد و بهبود قابل ملاحظه‌ای گزارش کند، آیا این نتایج برای شما قانع‌کننده خواهد بود؟ چه پیشنهاد تکمیلی برای رفع این مشکل دارید؟
- ۷- پژوهشگری تصمیم می‌گیرد در جریان کارآزمایی بالینی، روزهای زوج داروی جدید و روزهای فرد دارونما تجویز کند. آیا این شیوه تخصیص بیماران را صحیح می‌دانید؟
- ۸- پژوهشگری برای ۱۰ بیمار درمان مورد قبول خود و برای ۱۰ بیمار درمان جدید رقیب را با رعایت همه جوانب کارآزمایی شاهددار تصادفی شده دوسوکور تجویز کرده است و در پایان با $p = 0/3$ به این نتیجه رسیده است که درمان جدید هیچ‌گونه برتری بر درمان قدیم ندارد. موضع شما در برابر این نتیجه‌گیری چیست و چه دلایلی برای آن دارید؟
- ۹- با وجود آن که مطالعه‌های مداخله‌ای قوی‌ترین مطالعه‌ها از نظر کشف رابطه علت- معلولی هستند، به چه دلیل در بیشتر مطالعه‌های اپیدمیولوژیک این روابط را به وسیله مطالعه‌های مشاهده‌ای بررسی می‌کنند؟

پاسخ پرسش‌های پایانی

- پاسخ پرسش ۱- آری، مصون‌سازی درمانگران در برابر استنباط نابجا از تجربه‌های محدود بالینی، توانایی تجزیه و تحلیل انتقادی برای داوری درباره درمان‌های نوین و حاشیه‌ای، انتخاب مناسب‌ترین روش اقدام‌های پیش‌گیرانه و درمانی با توجه به شرایط خاص جامعه.
- پاسخ پرسش ۲- مطالعه‌های مداخله‌ای تنها نوع مطالعه‌ای هستند که به‌طور خالص می‌توانند اثر متغیر مستقل را، پس از حذف اثر مخدوش‌کننده‌ها، تعیین کنند.
- پاسخ پرسش ۳- هم‌گروهی.
- پاسخ پرسش ۴- موازی و پی‌درپی.
- پاسخ پرسش ۵- رضایت آگاهانه پیش از تخصیص تصادفی گرفته می‌شود و درباره تخصیص تصادفی نیز کسب رضایت می‌شود.
- پاسخ پرسش ۶- خیر، این نتایج ممکن است مربوط به بهبود خودبه‌خودی یا اثرات غیراختصاصی (دارونما) باشد. استفاده از گروه شاهد و تجویز دارونما توصیه می‌شود.
- پاسخ پرسش ۷- خیر، سوگرایی در انتخاب بیمار ممکن است هم از سوی خود پژوهشگر و هم از سوی بیمار اتفاق افتد.
- پاسخ پرسش ۸- تا زمانی که این مطالعه از توان آماری مناسب برخوردار نباشد، که به‌ویژه در نمونه‌های با اندازه کم مطرح است، ادعای این پژوهشگر پذیرفتنی نیست. علاوه بر این، پژوهشگران

می‌توانند از دیدن یا ندیدن رابطه‌ای سخن بگویند، نه بودن یا نبودن رابطه. پاسخ پرسش ۹- به علت محدودیت‌های مطالعه مداخله‌ای، به‌ویژه محدودیت اخلاقی.

پیوست ۱: روش‌های تصادفی‌سازی

استفاده از جدول اعداد تصادفی

ساده‌ترین روش مناسب برای تصادفی‌سازی استفاده از جدول اعداد تصادفی است. در این روش، از پیش تعیین می‌شود که هر عدد (یا هر مجموعه از اعداد) به کدام روش درمانی تعلق دارد، سپس برای هر فرد از نمونه به‌طور تصادفی یک عدد از این جدول انتخاب شده و به درمان تعیین شده تخصیص داده می‌شود. برای مثال، اگر پژوهشگر بخواهد افراد در حال مطالعه به دو گروه مساوی تقسیم شوند، می‌تواند از اعداد یک‌رقمی استفاده کند و ارقام زوج و صفر را به یک روش، و ارقام فرد را به روش دیگر تخصیص دهد. سپس، برای هر یک از افراد شرکت‌کننده در مطالعه یک عدد انتخاب کرده و، بسته به این‌که در کدام گروه قرار دارد، درمان موردنظر را برای او به کار برد. هم‌چنین، ممکن است از اعداد دورقمی استفاده کند و اعداد ۰ تا ۴۹ را در یک گروه و ۵۰ تا ۹۹ را در گروه دیگر جای دهد. همان‌گونه که مشاهده می‌شود، حالت‌های زیادی برای این کار می‌توان در نظر گرفت که هیچ‌کدام بر دیگری برتری ندارد. آنچه اهمیت دارد این است که انتخاب اعداد به‌صورت تصادفی صورت بگیرد. برای این کار می‌توان یک نقطه شروع تصادفی در جدول انتخاب کرد و سپس در جهتی معین (مثلاً به طرف راست، چپ، بالا، پایین یا مورب) حرکت کرد و هر یک از اعدادی را که از این راه به دست می‌آیند، به ترتیب، به بیمارانی که دارای شرایط شرکت در مطالعه هستند، نسبت داد. اگر پژوهشگر بخواهد تعداد افراد در دو گروه مساوی نباشد و از نسبت خاصی تبعیت کند، می‌تواند اعداد جدول را هم به همین نسبت تقسیم کند. برای مثال، چنانچه بخواهیم یک سوم افراد در یکی از گروه‌ها و بقیه در گروه دیگر قرار بگیرند، می‌توانیم اعداد ۱ تا ۳۳ را در یک گروه و اعداد ۳۴ تا ۹۹ را در گروه دیگر قرار دهیم. به همین ترتیب، می‌توان افراد را در بیش از دو گروه نیز جای داد.

تقسیم تصادفی بلوکی^(۱)

گاهی پژوهشگر می‌خواهد نسبت افراد در دو گروه، در هر زمان از پژوهش تقریباً مساوی باشد. برای مثال، وقتی مداخله از نوعی باشد که به تدریج و با گذشت زمان مهارت پژوهشگر در انجام آن افزایش یابد، بهتر است افراد یک گروه زودتر از گروه دیگر تکمیل نشوند.

تنها با استفاده از جدول اعداد تصادفی، به‌ویژه در نمونه‌های با اندازه کم، نمی‌توان دست‌یابی به

این هدف را تضمین کرد. برای این کار، از تقسیم تصادفی بلوکی استفاده می‌شود. در این روش بلوک‌هایی با اندازه زوج (معمولاً چهارتایی یا شش‌تایی) ایجاد می‌شود و در آن، به‌طور تصادفی، نیمی از افراد هر بلوک در یک گروه و نیمی دیگر در گروه دیگر قرار داده می‌شوند (در صورتی که بیش از دو گروه داشته باشیم یا بخواهیم نسبت افراد در دو گروه مساوی نباشد، نیز می‌توان همین روش را تعمیم داد). سپس آن قدر بلوک انتخاب می‌شود تا به اندازه نمونه لازم برسیم. برای مثال، چنانچه با استفاده از بلوک‌های چهارتایی بخواهیم ۱۰۰ نفر را در دو گروه ۵۰ تایی الف و ب قرار دهیم، ابتدا کلیه حالت‌های چهارتایی را که در آن نیمی از افراد به گروه الف و نیم دیگر به گروه ب تخصیص داده می‌شوند مشخص می‌کنیم. این حالت‌ها در زیر نوشته شده‌اند:

- ب ب الف الف - ۱
- ب الف الف ب - ۲
- الف ب ب الف - ۳
- الف الف ب ب - ۴
- الف ب الف ب - ۵
- ب الف الف ب - ۶

سپس به هر یک از این ترکیب‌های چهارتایی (که شش حالت را شامل می‌شود) یکی از رقم‌های ۱ تا ۶ را نسبت می‌دهیم. در مرحله بعد باید ۲۵ بلوک چهارتایی را به‌طور تصادفی انتخاب کرده و ترکیب‌های آن را پشت سرهم بنویسیم. برای این کار، باید ۲۵ بار از این جامعه شش‌عضوی نمونه‌گیری با جای‌گذاری انجام دهیم، یعنی ۲۵ بار عدد تصادفی بین ۱ تا ۶ انتخاب کنیم. برای مثال، فرض کنید عدد‌های ۱، ۴، ۱ و ۶ چهار عدد اول از ۲۵ عدد انتخاب شده باشند. به این ترتیب، نوع درمان برای ۱۶ بیمار اول عبارت خواهد بود از:

ترکیب ششم	ترکیب اول	ترکیب چهارم	ترکیب اول
ب الف الف ب	ب ب الف الف	الف الف ب ب	ب ب الف الف

پژوهشگر با ادامه این روند مطمئن خواهد بود که در هر جایی از پژوهش که ناچار شود کار نمونه‌گیری را پایان دهد تفاوت بین دو گروه حداکثر از دو نفر (نصف اندازه بلوک) تجاوز نخواهد کرد.

تقسیم تصادفی با طبقه‌بندی

استفاده از تصادفی‌سازی ساده یا تقسیم تصادفی بلوکی، توزیع یکنواخت متغیرهای مرتبط با پی‌آمد را تضمین نمی‌کند. برای مثال، فرض کنید پژوهشگری بخواهد تأثیر دو نوع درمان سرطان پستان را با یکدیگر مقایسه کند و برای این کار ۴۰ بیمار دچار به سرطان را به دو گروه ۲۰ تایی تقسیم کند. اگر یائسه‌بودن عاملی مهم در پاسخ به درمان باشد و ۱۰ نفر از این ۴۰ بیمار یائسه باشند، انتظار می‌رود

نیمی از این تعداد، یعنی پنج نفر، در هر گروه درمانی قرار بگیرند؛ ولی تنها برحسب احتمال ممکن است این نسبت به جای ۵:۵ به صورت ۷:۳ در گروه‌ها ظاهر شود. یعنی، نسبت افراد یائسه در یک گروه بیش از دو برابر گروه دیگر باشد (گرچه این اختلاف از نظر آماری معنی‌دار نیست - $P > 0.1$ - ولی می‌تواند نتایج مطالعه را مخدوش کند).

برای برخورد با این مشکل، می‌توان ابتدا بیماران را از نظر یائسه‌بودن به دو گروه تقسیم کرد و سپس هر گروه را به دو روش درمان تخصیص داد. می‌توان دو روش تقسیم تصادفی بلوکی و تقسیم تصادفی با طبقه‌بندی را به‌طور هم‌زمان انجام داد. طبقه‌بندی را می‌توان برای بیش از یک متغیر انجام داد، ولی با افزایش تعداد متغیرها انجام این امر مشکل یا حتی ناممکن می‌شود. به عبارت دیگر، ممکن است در بسیاری از حالت‌های ایجادشده نتوان حتی دو مورد مشابه پیدا کرد. برای مقابله با این مشکل می‌توان از روشی به نام کمینه‌سازی^(۱) استفاده کرد که در آن تخصیص افراد به گروه‌ها تنها بر اساس تصادف صورت نمی‌گیرد، بلکه افراد به گونه‌ای به گروه‌های مختلف تخصیص داده می‌شوند که تفاوت میان گروه‌ها از نظر متغیرهای مهم حداقل باشد. در این روش، فرد اول به‌طور تصادفی به یکی از گروه‌ها تخصیص داده می‌شود. سپس، افراد بعدی به گونه‌ای به گروه‌ها تخصیص داده می‌شوند که تفاوت بین گروه‌ها از نظر این متغیرهای مهم حداقل باشد. برای اطلاع بیشتر در این زمینه، مطالعه کتاب *Altman DG. Practical Statistics for Medical Research. Chapman & Hall, London, 1994.* توصیه می‌شود.

پیوست ۲: تجزیه و تحلیل بقا

حادثه مورد نظر، در بسیاری از مطالعه‌های مداخله‌ای، رخ داد یا عدم رخ داد واقعه‌ای مانند مرگ، عود یا بهبود بیماری است. به‌طور ساده، در این مطالعه‌ها می‌توان نسبت افرادی که واقعه را تجربه می‌کنند در هر یک از گروه‌ها محاسبه کرد و این نسبت را در گروه‌های مختلف مقایسه کرد (به مبحث خطر نسبی^(۲) و آزمون‌های آماری مناسب در فصل تجزیه و تحلیل داده‌ها مراجعه نمایید).

هم‌چنین، ممکن است فردی از شاخص میزان بروز^(۳) (تعداد رویدادهای جدید تقسیم بر مجموع زمان‌هایی که افراد در مطالعه حضور داشته‌اند) در هر گروه استفاده کند و به مقایسه میزان بروز در گروه‌های گوناگون بپردازد. در کاربرد روش‌های تجزیه و تحلیل در این مطالعه‌ها باید به مفروضات و محدودیت‌های این روش‌ها توجه داشت:

۱- شاخص خطر نسبی زمانی ارزش دارد که همه افراد مورد مطالعه در یک زمان وارد مطالعه شده و در یک زمان نیز از مطالعه خارج شوند. برای مثال، چنانچه در زمان پی‌گیری، فردی از ادامه مطالعه خودداری کند یا به‌دلیلی در پی‌گیری گم شود^(۴)، آیا باید تعداد کل را تعداد افراد ابتدای مطالعه در نظر گرفت یا تعداد افراد باقی‌مانده؟ هیچ‌یک از این دو کار درست نیست و در چنین شرایطی نمی‌توان از شاخص خطر نسبی استفاده کرد. هم‌چنین، طول مدت پی‌گیری در گروه‌های متفاوت

1. minimization
3. incidence rate

2. relative risk
4. loss to follow-up

باید یکسان باشد (نمی‌توان خطر پنج‌ساله بروز عارضه در یک گروه را با خطر هفت‌ساله بروز آن در گروه دیگر مقایسه کرد). نکته قابل توجه دیگر این است که در پی آمده‌های شایع، اگر زمان پی‌گیری به اندازه کافی طولانی باشد، ممکن است تقریباً تمام افراد دو گروه پی‌آمد را تجربه کنند و نتوان به درستی اختلاف آنها را نشان داد.

۲- با استفاده از شاخص میزان بروز یا شاخص‌هایی که در مخرج کسر آنها (افراد در معرض خطر) از «شخص-زمان»^(۱) استفاده می‌شود، گرچه می‌توان این مشکل را حل کرد؛ ولی باز با مشکلات ویژه این روش‌ها روبه‌رو می‌شویم. برای مثال، در این روش‌ها ۱۰۰ نفر که هر کدام به‌طور متوسط به مدت یک سال پی‌گیری شده‌اند، معادل ۱۰ نفر که هر کدام به‌طور متوسط به مدت ۱۰ سال پی‌گیری شده‌اند، و حتی یک نفر که به مدت ۱۰۰ سال پی‌گیری شده باشد در نظر گرفته می‌شود. ولی آیا واقعاً نتیجه پی‌گیری این سه حالت یکسان است؟

۳- مشکل دیگری که در کاربرد شاخص‌های فوق پیش می‌آید، و تا حدودی با مشکل قبل هم‌پوشانی دارد، این است که در این روش احتمال رخداد واقعه (مانند دفع پیوند) در دوره‌های زمانی مشابه در طول مدت مطالعه یکسان فرض می‌شود. برای مثال، احتمال رد پیوند در ماه اول پس از پیوند معادل این احتمال در ماه ششم پس از پیوند در نظر گرفته می‌شود و واضح است که در بسیاری از حالت‌های واقعی نمی‌توان این فرض را صادق دانست.

به سبب وجود مشکلات فوق (و مشکلات دیگر که در این جا به آنها اشاره نشده) در این مطالعه‌ها ترجیح داده می‌شود پی‌آمد به جای رخداد واقعه^(۲) به صورت زمان لازم تا رخداد واقعه^(۳) تعریف شود. چنین مطالعه‌هایی روش تجزیه و تحلیل ویژه‌ای با نام تجزیه و تحلیل بقا^(۴) دارند. باید توجه داشت که واژه «بقا» لزوماً به این معنی نیست که واقعه مورد نظر مرگ یا حتی یک واقعه منفی (مانند عود سرطان) است، بلکه این واقعه می‌تواند یک واقعه مثبت (مانند بهبود بیماری) یا خنثی (مانند پایان شیردهی) باشد.

ممکن است در نگاه اول چنین به نظر برسد که برای تجزیه و تحلیل چنین مطالعه‌ای می‌توان میانگین مدت بقا در گروه‌های گوناگون را با یکدیگر مقایسه کرد. چنین روش تجزیه و تحلیل، گرچه به صورت نظری مشکلی ندارد، مستلزم آن است که همه افراد، در مدت مطالعه، واقعه مورد نظر را تجربه کرده باشند. برای مثال، چنانچه در مطالعه‌ای ده نفر را به مدت ده سال پی‌گیری کنیم و تا پایان مطالعه هنوز دو نفر واقعه را تجربه نکرده باشند، «زمان لازم تا رخداد واقعه» برای این دو نفر چه قدر است؟

در مورد مقایسه میان مدت بقا در گروه‌ها، گرچه لازم نیست حتماً همه افراد واقعه مورد نظر را تجربه کنند و اگر دست‌کم نیمی از افراد هر گروه واقعه را تجربه کرده باشند میانه محاسبه‌شدنی است، خروج افراد از مطالعه^(۵) موضوع را قدری پیچیده می‌سازد.

1. person-time

2. event

3. time to event

4. survival analysis

5. censor

تجزیه و تحلیل داده‌های بقا با روش کاپلان-مایر^(۱)

یکی از روش‌های متداول تجزیه و تحلیل داده‌های بقا، که مشکل‌ها و محدودیت‌های بالا را ندارد، استفاده از روش کاپلان-مایر است. در این روش در هر لحظه از زمان پی‌گیری می‌توان احتمال بقا (عدم رخداد واقعه) را محاسبه کرد. برای آشنایی با این روش بد نیست ابتدا با چند اصطلاح مرسوم در این زمینه آشنا شویم:

– **خروج از مطالعه:** هر فردی که وارد مطالعه شود به یکی از این دو صورت مطالعه را ترک می‌نماید: یا به دلیل رخداد واقعه فرد از مطالعه خارج می‌شود، یا پیش از رخداد واقعه و به دلایلی که در زیر به آنها اشاره می‌شود از مطالعه خارج می‌شود.

– **خروج از مطالعه بدون رخداد واقعه:** به طور نظری سه دلیل می‌توان برای این وضعیت در نظر گرفت: یا فرد در پی‌گیری گم می‌شود، یا مدت مطالعه به پایان می‌رسد و فرد هنوز واقعه را تجربه نکرده، یا واقعه دیگری برای او اتفاق افتاده و امکان رخدادن واقعه مورد نظر وجود ندارد. دو حالت اول نیاز به توضیح ندارند. در مورد حالت سوم به این مثال توجه کنید: اگر گروهی از افراد دچار سرطان، در حال درمان با روش خاصی باشند و واقعه مورد نظر مرگ در اثر عود سرطان باشد، فردی که در اثر سانه‌راندگی فوت کرده، دیگر نمی‌تواند واقعه مورد نظر را تجربه کند. در عمل معمولاً این سه حالت یکی در نظر گرفته می‌شود.

– **فاصله^(۲):** منظور فاصله بین خروج هر فرد تا خروج فرد بعدی است. اولین فاصله عبارت است از فاصله زمانی بین آغاز مطالعه و خروج اولین فرد. فاصله بین خروج اولین فرد و دومین فرد، دومین فاصله را تشکیل می‌دهد و ...

اساس کار در این روش را محاسبه احتمال وقوع رخداد (و در نتیجه احتمال بقا که برابر است با: «احتمال وقوع - ۱») در هر یک از این فاصله‌ها تشکیل می‌دهد. به این ترتیب که در هر دوره، احتمال وقوع عبارت است از تعداد افرادی که در این فاصله واقعه را تجربه کرده‌اند تقسیم بر تعداد افراد ابتدای فاصله. هم‌چنین، احتمال بقا در این فاصله، همان‌گونه که گفته شد، از کم کردن مقدار محاسبه شده از عدد یک به دست می‌آید. بقا در فاصله دوم مشروط به بقا در فاصله اول است. به عبارت دیگر، احتمال این که فردی تا پایان فاصله دوم بقا داشته باشد عبارت است از: «احتمال بقا در دوره اول ضرب در احتمال بقا در دوره دوم». به این ترتیب، می‌توان احتمال بقا تا هر نقطه زمانی دل‌خواه را محاسبه کرد:

ابتدا فاصله‌هایی را که در فاصله زمانی بین آغاز مطالعه و نقطه مورد نظر قرار دارند مشخص نموده و سپس احتمال بقا در هر یک از این فاصله‌ها را در یکدیگر ضرب می‌کنیم.

مثال: جدول زیر داده‌های مطالعه فرضی پی‌گیری بیماران مبتلا به نوعی سرطان که شیمی‌درمانی

می‌شوند را نشان می‌دهد. واقعه مورد نظر پس‌رفت^(۱) سرطان است. بنابراین، منظور از بقا در این مطالعه «نرسیدن به پس‌رفت» است و واقعه مورد نظر یک واقعه مثبت است. در مورد هر بیمار مدت حضور در مطالعه (به صورت تعداد هفته‌های سپری شده از آغاز مطالعه تا خروج از مطالعه) و وضعیت فرد هنگام خروج ثبت شده و بر این اساس احتمال بقای هر فاصله و سپس احتمال بقای تجمعی محاسبه شده است (برای کوتاه کردن نوشته تنها داده‌های مربوط به ۱۸ بیمار اول نشان داده شده است).

جدول شماره ۱- داده‌های مطالعه فرضی پی‌گیری بیماران مبتلا به سرطان که شیمی‌درمانی می‌شوند

تعداد باقی‌مانده در پایان دوره	بقای تجمعی	احتمال بقا	احتمال رخداد واقعه	رخداد واقعه	زمان به هفته
۱۷	۰/۹۴۴	۰/۹۹۴	۰/۰۵۶	+	۱
۱۶	$۰/۹۴۴ \times ۰/۹۴۱ = ۰/۸۸۹$	۰/۹۴۱	۰/۰۵۹	+	۵
۱۵	$۰/۸۸۹ \times ۰/۹۳۸ = ۰/۸۳۳$	۰/۹۳۸	۰/۰۶۲	+	۵
۱۴	$۰/۸۳۳ \times ۱ = ۰/۸۳۳$	۱	۰	-	۶
۱۳	$۰/۸۳۳ \times ۰/۹۲۹ = ۰/۷۷۴$	۰/۹۲۹	۰/۰۷۱	+	۷
۱۲	$۰/۷۷۴ \times ۱ = ۰/۷۷۴$	۱	۰	-	۹
۱۱	$۰/۷۷۴ \times ۱ = ۰/۷۷۴$	۱	۰	-	۱۰
۱۰	$۰/۷۷۴ \times ۱ = ۰/۷۷۴$	۱	۰	-	۱۰
۹	$۰/۷۷۴ \times ۰/۹ = ۰/۶۹۶$	۰/۹	۰/۱	+	۱۳
۸	$۰/۶۹۶ \times ۰/۸۸۹ = ۰/۶۱۹$	۰/۸۸۹	۰/۱۱۱	+	۱۶
۷	$۰/۶۱۹ \times ۱ = ۰/۶۱۹$	۱	۰	-	۱۷
۶	$۰/۶۱۹ \times ۱ = ۰/۶۱۹$	۱	۰	-	۱۹
۵	$۰/۶۱۹ \times ۱ = ۰/۶۱۹$	۱	۰	-	۲۲
۴	$۰/۶۱۹ \times ۰/۸ = ۰/۴۹۵$	۰/۸	۰/۲	+	۲۳
۳	$۰/۴۹۵ \times ۱ = ۰/۴۹۵$	۱	۰	-	۳۲
۲	$۰/۴۹۵ \times ۱ = ۰/۴۹۵$	۱	۰	-	۳۳
۱	$۰/۴۹۵ \times ۱ = ۰/۴۹۵$	۱	۰	-	۳۳
۰	$۰/۴۹۵ \times ۱ = ۰/۴۹۵$	۱	۰	-	۳۵

همان‌گونه که مشاهده می‌شود در پایان هفته اول یکی از بیماران پس‌رفت را تجربه کرده، بنابراین نسبت بیمارانی که در این زمان هنوز پس‌رفت را تجربه نکرده‌اند (به عبارت دیگر هنوز «باقی مانده‌اند») عبارت است از $\frac{۱۷}{۱۸} = ۰/۹۴۴$ - ۱. در پایان هفته پنجم، که دومین بیمار به پس‌رفت

رسیده، احتمال بقا بین هفته‌های اول تا پنجم عبارت است از $\frac{16}{17} = 0/941$ ، $1 - \frac{1}{17} = \frac{16}{17}$. از سوی دیگر، بیمار سوم نیز در پایان هفته پنجم واقعه را تجربه کرده است. بنابراین، در عمل احتمال بقا بین هفته‌های اول و پنجم عبارت است از $\frac{15}{17} = 0/882$ ، $1 - \frac{2}{17} = \frac{15}{17}$. عدد $0/882$ احتمال بقا بین هفته‌های اول تا پنجم را نشان می‌دهد.

برای تعیین احتمال بقا در پایان هفته پنجم، فرد باید در هفته اول بقا داشته باشد تا بتواند بین هفته‌های اول تا پنجم بقا داشته باشد، بنابراین:

$$\text{احتمال بقا در پایان فاصله دوم} = \text{احتمال بقا در فاصله اول} \times \text{احتمال بقا در فاصله دوم}$$

هم‌چنین، توجه داشته باشید که در این مثال، احتمال بقا در فاصله دوم و سوم، روی هم ریخته شده‌اند، زیرا هر دو فاصله به پایان هفته پنجم ختم شده‌اند. به عبارت دیگر، عدد $0/882$ را از دوراه می‌توان به دست آورد:

$$\frac{16}{17} \times \frac{15}{16} \text{ (ضرب احتمال بقا در فاصله‌های دوم و سوم) و}$$

$\frac{15}{17}$ (احتمال بقا در فاصله دوم و سوم که در عمل روی هم افتاده‌اند). برای سادگی کار در جدول بالا، برای خروج هر فرد یک فاصله تعریف شده است. به این ترتیب، احتمال بقا در پایان هفته پنجم برابر است با:

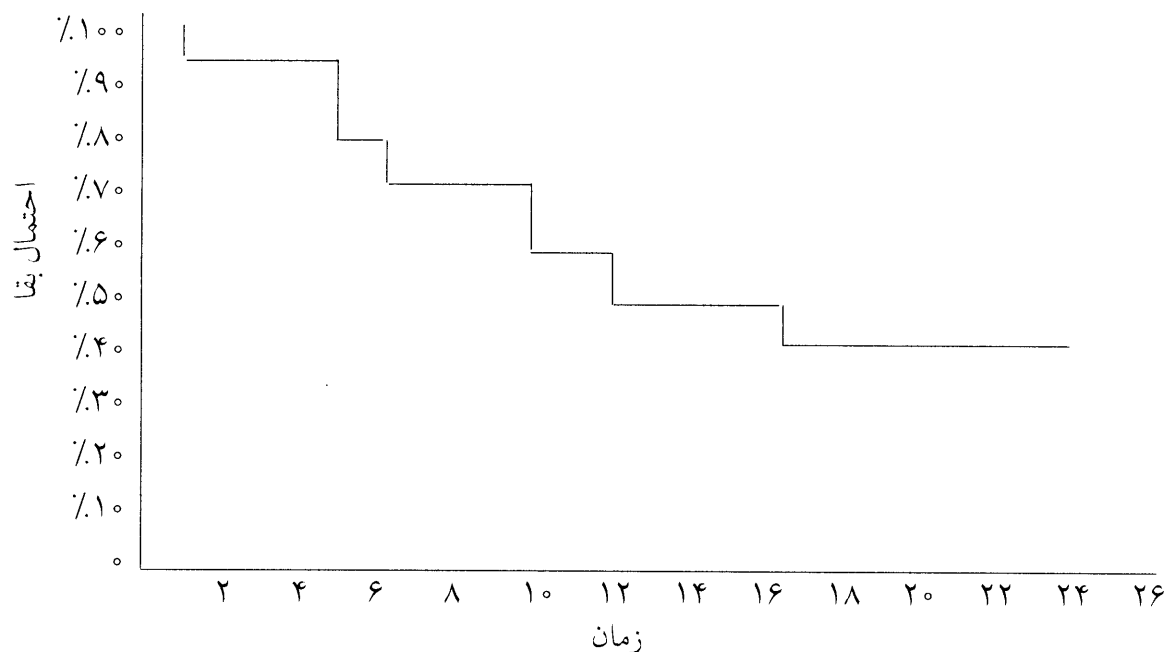
$$0/944 \times 0/941 \times 0/938 = 0/833$$

چهارمین فرد، در پایان هفته ششم، بدون این‌که واقعه را تجربه کند از مطالعه خارج شده است.

توجه داشته باشید که احتمال رخداد واقعه در فاصله چهارم (یعنی بین هفته پنجم و ششم) صفر و در نتیجه احتمال بقا در این دوره ۱۰۰٪ است. بنابراین، تغییری در بقای تجمعی ایجاد نمی‌شود، تنها تعداد افراد باقی‌مانده که به فاصله‌های بعدی راه می‌یابند کاهش یافته است. سه نفری که بین هفته‌های نهم و دهم از مطالعه خارج شده‌اند نیز همین وضعیت را دارند.

پس از کامل کردن این جدول می‌توان احتمال بقا در هر نقطه زمانی دلخواه را مشخص کرد. برای مثال، احتمال بقای ۱۲ هفته عبارت است از $0/774$ ، زیرا بین هفته‌های ۱۰ تا ۱۳ هیچ فردی از مطالعه خارج نشده و بقای ۱۰ هفتگی با بقای ۱۲ هفتگی برابر است.

در عمل برای بهتر نشان دادن نتیجه این تجزیه و تحلیل، معمولاً آنها را به صورت یک نمودار نمایش می‌دهند که محور افقی آن نشان‌دهنده زمان و محور عمودی آن نشان‌دهنده احتمال بقاست. به این نمودار منحنی بقا^(۱) گفته می‌شود. نتیجه مطالعه فرضی بالا را می‌توان به شکل زیر نشان داد:



نمودار ۲- نمونه منحنی بقا

چنانچه پژوهشگری بخواهد بقای دو یا چند گروه از افراد مورد مطالعه را با یکدیگر مقایسه کند (به عبارت دیگر عوامل مؤثر بر بقا را بررسی کند)، می‌تواند برای هر یک از این گروه‌ها منحنی کاپلان-مایر را رسم کند و سپس بقای دو گروه را در هر نقطه‌ای از زمان مقایسه کند. عیب این روش آن است که به این ترتیب مقایسه‌ای کلی از بقای دو گروه به دست نمی‌آید و چنانچه پژوهشگر بقای دو گروه را در زمان‌های مکرر، مثلاً در پایان هر هفته، با یکدیگر مقایسه کند، مشکل مقایسه‌های چندگانه^(۱) پیش می‌آید که در آن احتمال خطای نوع اول (یافتن رابطه معنی‌دار کاذب) افزایش می‌یابد (برای آگاهی بیشتر در این زمینه به فصل تجزیه و تحلیل داده‌ها و کتاب‌های معتبر آماری مراجعه کنید).

برای حل این مشکل، روش‌هایی ابداع شده که بقای کلی گروه‌ها را (بدون در نظر گرفتن زمان) با یکدیگر مقایسه می‌کنند. یکی از رایج‌ترین این روش‌ها استفاده از آزمون Logrank است که فرضیه برابری بقای گروه‌های مختلف را آزمون می‌کند. این روش، همانند سایر آزمون‌های فرضیه، برآورد مناسب از اندازه اختلاف بقا در دو یا چند گروه به دست نمی‌دهد و تنها معنی‌داری اختلاف را نشان می‌دهد.

توجه داشته باشید که اگر منحنی‌های بقا یکدیگر را قطع کنند (یعنی در آغاز مطالعه بقای گروه اول بیشتر از گروه دوم باشد، ولی از یک نقطه زمانی خاص بقای گروه دوم بیشتر شود)، انجام این آزمون بدون خطا نیست و باید نتیجه آن را با احتیاط تفسیر کرد.

نکته مهم که درباره نمودار کاپلان-مایر باید به آن اشاره کرد این است که گرچه لفظ «منحنی» به فراوانی برای آن به کار می‌رود ولی ماهیت این نمودار به صورت خط شکسته است (به نمودار ۲ مراجعه شود) و ترسیم آن به صورت یک خط منحنی درست نیست.

روش دیگر برای مقایسه بقای گروه‌ها استفاده از مدل‌های آماری مانند مدل تناسبی مخاطره‌ها (مدل کاکس)^(۱) است که علاوه بر آزمون فرضیه، برآوردی از اندازه اختلاف بین گروه‌ها نیز به دست می‌دهد. باید توجه داشت که استفاده از این مدل‌ها، مانند استفاده از هر مدل آماری دیگر، نیازمند فهم دقیق و روشن مفروضات و منطق آن است و بدون داشتن چنین دانشی استفاده از آن توصیه نمی‌شود.

منابع

- ۱- بشردوست ن، اردلان ع. طراحی انواع مطالعه‌های اپیدمیولوژیک. یزد. انتشارات طب‌گستر و معاونت امور پژوهشی دانشگاه علوم پزشکی شهید صدوقی یزد، ۱۳۷۸.
- ۲- عزیزی ف و شیخ‌الاسلام ر. برنامه کشوری مبارزه با کمبود ید. طب و تزکیه، شماره‌های ۱۹ و ۲۰، تابستان ۱۳۷۴ و بهار ۱۳۷۵، صص ۱۸ تا ۲۲.
- ۳- فتوحی ا، اصغری ف، میرزاده ع، علاءالدینی ف، یزدانی ک، آریا آ، صفایی ا، وفادار ج. بررسی مقایسه‌ای اثرات درمانی و عوارض کوتاه‌مدت آنتوم با اسید نیکوتینیک، کلستیرامین، مجموع آنها و دارونما در افراد مبتلا به هیپرکلسترولمی با متد دوسوکور و انتخاب تصادفی نمونه‌ها. فصلنامه گیاهان دارویی. ۱۳۸۱، سال اول، شماره سوم، صص ۵۱ تا ۶۰.
- ۴- کلاهی ع، فرسارار، یاسمی م. تأثیر آموزش در میزان اضطراب مادران بعد از اولین تشنج همراه تب.
- ۵- میرمیران پ، کیمیاگر م، نفرآبادی م ط و همکاران. اثرات تزریق محلول روغنی یددار در دانش‌آموزان مبتلا به اختلالات ناشی از کمبود ید، بررسی سه ساله. پژوهش در پزشکی، سال ۲۱، شماره ۱، فروردین و خرداد ۱۳۷۶، صص ۵۴ تا ۶۶.
- ۶- میرمیران پ، سرشار مقدم ا، مداح پ و همکاران. تغییرات شدت گواتر، ضریب هوشی، آزمون
- سایکوموتور و غلظت هورمون‌های تیروئید پس از تزریق محلول روغنی یددار در روستاهای کشاور. مجله غدد درون‌ریز و متابولیسم ایران، سال اول، شماره ۱، بهار ۱۳۷۸، صص ۴۸ تا ۵۵.
- ۷- یاسمی م. انتخاب بیماران و جای‌گذاری بیماران در گروه‌های درمانی. سمپوزیوم آشنایی با کارآزمایی بالینی داروها. دانشگاه علوم پزشکی تهران، ۱۳۷۲.
- ۸- یاسمی م. برخی از مسایل روش‌شناختی در ارزیابی درمان‌های غیردارویی در روان‌پزشکی. پژوهش‌های روان‌شناختی. ۱۳۷۳، دوره ۲، شماره ۳ و ۴، صص ۷۰ تا ۸۰.

9. Altman DG. Randomization. *BMJ*, 1991; 302: 1481-82.

10. Altman DG. *Practical Statistics for Medical Research*. Chapman & Hall, London, 1994.

11. Altman DG, Bland M. Time to Event (survival) sata. *BMJ* 1998; 317: 468-69.

12. Altman DG, Bland M. Treatment Allocation in Controlled Trials: Why Randomise? *BMJ*, 1999; 318:1209.
13. Altman DG, Bland M. How to Randomise. *BMJ*, 1999; 319: 703-4.
14. Armitage P, Berry G. *Statistical Methods in Clinical Research*. Oxford, Blackwell, 1987.
15. Azizi F, Kalani H, Kimiagar M et al. Physical, Neuromotor and Intellectual impairment in Non-cretinous School Children with Iodine Deficiency. *Int J. Nut Res.* 1995; 65: 199-205.
16. Azizi F, Kimiagar M, Ghazi AA. et al. Treatment of Goitrous Hypothyroidism with Iodized Oil Supplementation in an Area of Iodine Deficiency. *Exp. Clin. Endocrinol. Diab.* 1996; 104: 387-391.
17. Brown W.A. The Placebo Effect. *Scientific American*, 1998, 278: 90-95.
18. Collet D. *Modeling Survival Data in Medical Research*. Chapman & Hall, London, 1994.
19. Chalmers TC. *Data Analysis for Clinical Medicine*. Rome International University Press. 1988.
20. Fleiss JL. *The Design and Analysis of Clinical Experiments*. John Wiley & Sons, New York, 1986.
21. Friedman GD. *Primer of Epidemiology*. (4th edition). New York, McGraw Hill, 1994.
22. Mausner JS, Kramer S. *Epidemiology, an Introductory Text*. (2nd edition). Philadelphia, WB Saunders Co. 1985.
23. Pocock SJ. *Clinical trials, a Practical Approach*. John Wiley & Sons, London, 1983.
24. Socks H, Chalmers TC, Smith H. Randomization versus Historical Control for Clinical Trials. *AMJ Med*, 1982; 72: 233-240.
25. Smithells RW, Shepard S, Schoran CS et al. Possible Prevention of Neural tube Defects by Periconceptual Vitamin Supplementation. *Lancet*, 1980; i: 339-340.
26. Spreit A, Simon P. *Methodology of Clinical Drug Trials*, Basel, Karger, 1985.
27. Treasure T, Mac Roc KD. Minimization: The Platinum Standard for Trials? *BMJ*, 1998, 317: 362-363.
28. Vogt DD, Montague M. Historical Review in *Materson GM. (Ed) The Clinical Research Process in the Pharmaceutical Industry*. New York, Marcel Dekker, 1984, 51-65.
29. Wingo PA, Higgins JE, Rubin CL, Zahniser SC. *An Epidemiologic Approach to Reproductive Health*. CDC publication, Atlanta, 1991.

منابع برای مطالعه بیشتر

1. Altman DG. *Practical Statistics for Medical Research*. Chapman & Hall, London, 1994.

نویسنده این کتاب افرادی را مخاطب قرار داده است که اطلاعات آماری چندانی ندارند. از همین نویسنده سلسله مقاله‌هایی در مجله *British Medical Journal* به چاپ رسیده است که می‌تواند مکمل مطالب کتاب باشد. به ویژه توصیه می‌شود پزشکان، این کتاب و سلسله مقاله‌های یادشده را در زمینه مطالعه مداخله‌ای مطالعه کنند.

2. Fleiss JL. *The Design and Analysis of Clinical Experiments*. John Wiley & Sons, New York, 1986.

این کتاب به جنبه‌های تخصصی تر طراحی و تحلیل آماری پرداخته است که برای فهم آن به اطلاعات پایه درباره روش‌های آماری نیاز است.

3. Kahn HA. *Statistical Methods in Epidemiology*. Oxford University Press, New York, 1989.

در این کتاب با دید اپیدمیولوژی به آمار، بیشتر بر نکته‌هایی تأکید شده که در مطالعه‌های اپیدمیولوژیک (که به طور عمده با پی‌آمدهای دو حالته سروکار دارند) کاربرد دارد. دو فصل از این کتاب (فصل‌های پنجم و ششم) روش‌های حذف اثر متغیرهای مزاحم با و بدون استفاده از مدل‌های آماری را به خوبی شرح داده است.