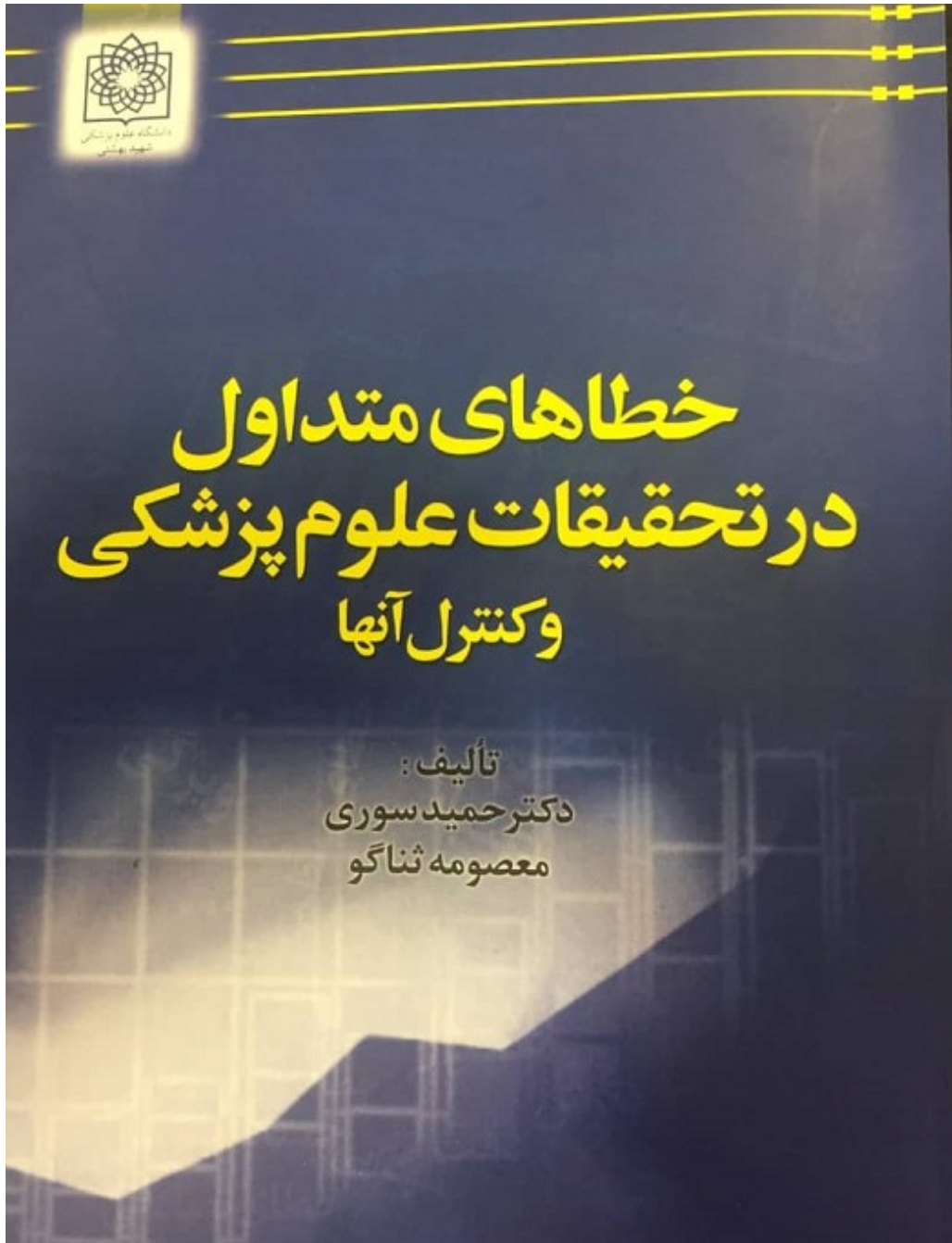


evidence.ir

t.me/irevidence



خطاهای متداول در تحقیقات علوم پزشکی و کنترل آنها

دکتر حمید سوری، استاد اپیدمیولوژی

سرشناسه	: سوری، حمید ۱۳۳۸ -
عنوان و نام پدیدآور	: خطاهای متداول در تحقیقات علوم پزشکی و کنترل آن‌ها/ تألیف حمید سوری، معصومه ثناگو
مشخصات نشر	: تهران: دانشگاه علوم پزشکی و خدمات بهداشتی درمانی شهید بهشتی، ۱۳۸۶.
مشخصات ظاهری	: ۲۹۳ ص: جدول، نمودار.
شابک	: ۹۷۸-۹۶۴-۸۸۵۸-۴۲-۶
وضعیت فهرست‌نویسی	: فیبا
یادداشت	: ص.ع. به انگلیسی: Hamid Soori, Masoomeh Sanagoo common errors in biomedical research & their control
یادداشت	: چاپ قبلی: مرکز ملی تحقیقات علوم پزشکی کشور، ۱۳۸۴.
یادداشت	: کتاب‌نامه
موضوع	: پزشکی—تحقیق—روش‌شناسی.
موضوع	: اشتباه‌های علمی
شناسه افزوده	: ثناگو، معصومه، ۱۳۵۲ -
شناسه افزوده	: دانشگاه علوم پزشکی و خدمات بهداشتی درمانی شهید بهشتی
رده‌بندی کنگره	: ۱۳۸۶ قخ ۹ س / R ۸۵۰
رده‌بندی دیویی	: ۶۱۰/۷۲۴
شماره کتابشناسی ملی	: ۱۰۳۶۰۵۳



نام کتاب: خطاهای متداول در تحقیقات علوم پزشکی و کنترل آن‌ها

تألیف:

دکتر حمید سوری

معصومه ثناگو

ناشر: دانشگاه علوم پزشکی شهید بهشتی

نوبت چاپ: دوم - ۱۳۸۶

بها: ۳۵۰۰۰ ریال

شمارگان: ۱۵۰۰ جلد

چاپ و صحافی:

صبا

شابک: ۹۶۴-۸۶۴۹-۰۵-۷

تهران، اوین، جنب بیمارستان طالقانی، دانشگاه علوم پزشکی شهید بهشتی

تلفن: ۲۲۴۰۰۲۰۷ دورنگار: ۲۲۴۰۸۹۹۵

کلیه حقوق مادی این کتاب متعلق به دانشگاه علوم پزشکی شهید بهشتی است.

خطاهای متداول در تحقیقات علوم پزشکی و کنترل آنها

چاپ دوم

تألیف:

دکتر حمید سوری، استاد اپیدمیولوژی

آمار حیاتی



دانشگاه علوم پزشکی شهید بهشتی

دبآچه چاپ دوم

فهرست مطالب

فهرست شكلها

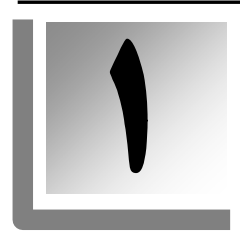
فهرست نمودارها

فهرست جداول

- جدول ۱-۱: شمای کلی از پروپوزال طرح‌های تحقیقاتی ۱۵
- جدول ۲-۱: نواقص و راه بهبود سؤال پژوهش و طرح مطالعه ۱۷
- جدول ۳-۱: نحوه طراحی یک پژوهش بالینی برای مطالعه اثر هورمون درمانی پس از یائسگی برای جلوگیری از CHD ۱۸
- جدول ۴-۱: تنظیم معیارهای انتخاب برای یک کارآزمایی بالینی فرضی ۲۰
- جدول ۵-۱: حالت‌های مختلف قضاوت‌های آماری و واقعیت‌های جامعه ۲۵
- جدول ۱-۲: دسته‌بندی طرح‌های مطالعات در پزشکی ۳۵
- جدول ۱-۳: دلایل حذف افراد تحت مطالعه در کارآزمایی‌های بالینی ۶۶
- جدول ۲-۳: نمونه‌های فهرست عوارض جانبی برای ارزیابی یک دارو جدید ۶۷
- جدول ۳-۳: جدول اعداد تصادفی ۷۳
- جدول ۴-۳: چک‌لیست دستورالعمل CONSORT برای گزارش کارآزمایی‌های بالینی تصادفی شده ۷۶ و ۷۱
- جدول ۵-۳: چهار گروه‌درمانی با حجم نمونه برابر در هر گروه در یک مطالعه فاکتوریل ۲×۲ ۷۵
- جدول ۶-۳: طرح فاکتوریل برای مطالعه تأثیر آسپیرین و بتا-کاروتن ۷۶
- جدول ۱-۴: راهبردهای مختلف برای بالا بردن دقت ۹۶
- جدول ۲-۴: اعتبار و دقت در اندازه‌گیری یک ویژگی ۹۷
- جدول ۳-۴: راهبردهای مختلف در کاهش خطای منظم به‌منظور افزایش اعتبار ۹۹
- جدول ۴-۴: نحوه ارزیابی آزمایش‌های پزشکی ۱۰۱
- جدول ۵-۴: تفسیر نتایج عکس‌های پاتولوژی توسط دو پاتولوژیست ۱۰۳
- جدول ۶-۴: نحوه تنظیم جدول ۲×۲ در آزمون‌های تشخیصی ۱۰۷
- جدول ۷-۴: شاخص‌های ارزیابی آزمون‌های تشخیصی ۱۰۸
- جدول ۸-۴: مقایسه حافظه والدین با پرونده‌های بهداشتی در تزریق واکسن ب ت ژ ۱۰۸
- جدول ۹-۴: رابطه بین مننژیت باکتری یا وجود باکتری در خون و تعداد گلبول‌های سفید نوزادان ۱۱۷
- جدول ۱۰-۴: مشکلات ناشی از خطای اندازه‌گیری یا طبقه‌بندی ۱۱۹
- جدول ۱-۵: پنج استنباط ممکن که پس از مشاهده رابطه بین نوشیدن قهوه و سکت قلبی در نمونه مورد مطالعه امکان وقوع دارند ۱۲۷
- جدول ۲-۵: روش‌های کاهش مشاهده روابط تصنعی در مطالعات تحلیلی ۱۲۸
- جدول ۳-۵: کنترل و یا حذف اثر مخدوش‌کنندگی در مرحله طراحی مطالعات ۱۳۱
- جدول ۴-۵: تعدیل اثر مخدوش‌کنندگی در مرحله تجزیه و تحلیل اطلاعات ۱۳۲
- جدول ۵-۵: ارزیابی شواهد ارائه شده برای وجود رابطه سببی بین عفونت با هلی‌کوباکتریلوری و زخم دئودنوم ۱۳۴
- جدول ۱-۶: وضعیت عامل خطر و وجود یا عدم وجود بیماری در یک مثال فرضی ۱۳۷
- جدول ۲-۶: وضعیت عامل خطر و بیماری به تفکیک مخدوش‌کننده فرضی در داده‌های مربوط به جدول ۱-۶ ۱۳۸
- جدول ۳-۶: وضعیت عامل خطر و بیماری در داده‌های فرضی ۱۳۸
- جدول ۴-۶: جدول توافقی سه‌طرفه مربوط به عامل خطر، وضعیت بیماری و مواجهه با یک مخدوش‌کننده فرضی ۱۳۹
- جدول ۵-۶: وضعیت مسکن و وقوع CHD در مردان ۱۳۹
- جدول ۶-۶: داده‌های به تفکیک استعمال و عدم استعمال سیگار آورده شده است ۱۴۰
- جدول ۱-۷: تحلیل بر مبنای روش طبقه‌بندی برای ارزیابی رابطه جنس و بیماری مالاریا ۱۵۳
- جدول ۲-۷: تحلیل طبقه‌بندی رابطه قرص‌های ضدبارداری خوراکی ((OCP و سکت قلبی در زن‌ها ۱۵۴
- جدول ۳-۷: استفاده از روش استاندارد مستقیم برای مقایسه میزان بروز ((I در دو گروه مورد مطالعه ۱۵۷
- جدول ۴-۷: وقوع بیماری گرفتگی عروق کرونر به تفکیک گروه‌های سنی و رده‌های اجتماعی - اقتصادی ۱۵۷
- جدول ۵-۷: استفاده از استاندارد کردن غیرمستقیم برای محاسبه میزان مرگ‌ومیر استاندارد شده با جمعیت مرجع ۱۵۸
- جدول ۶-۷: نمادگذاری‌های به کاررفته برای محاسبه نسبت شانس تعدیل یافته به روش مانتل-هانزل ۱۵۹
- جدول ۷-۷: نحوه نمایش نتایج در مطالعات زوج شده ۱۶۰

جدول ۷-۸:	مثال فرضی از نحوه انجام تعدیل مستقیم برای کنترل هم‌زمان جنس، نژاد و میزان تحصیلات.....	۱۶۱
جدول ۸-۱:	جدول توافقی 2×2 نمایش یک متغیر پاسخ دوحالتی و متغیر مواجهه دوسطحی.....	۱۶۷
جدول ۸-۲:	شاخص‌های مشهور برای مقایسه دو نسبت.....	۱۶۹
جدول ۸-۳:	وضعیت سیگار کشیدن بر اساس جنس.....	۱۷۰
جدول ۸-۴:	وضعیت سیگار کشیدن و ایجاد بیماری‌های قلبی و عروقی در طول پیگیری.....	۱۷۱
جدول ۸-۵:	نحوه نمادگذاری و تفسیر ضرایب رگرسیونی در مدل‌های رگرسیون چندگانه.....	۱۷۳
جدول ۸-۶:	خروجی کامپیوتری برآورد ضرایب رگرسیونی چند گانه برای اندازه‌گیری FEV.....	۱۷۵
جدول ۸-۷:	معادله‌های مدل رگرسیون پواسن برای مقایسه دو گروه مواجهه.....	۱۷۵
جدول ۸-۸:	ده سطر اول بانک اطلاعاتی مطالعه بررسی عوامل خطر بیماری‌های قلبی و عروقی.....	۱۷۶
جدول ۸-۹:	میزان‌های سکنه قلبی در دو گروه سیگاری و غیرسیگاری.....	۱۷۶
جدول ۸-۱۰:	رگرسیون پواسن برای محاسبه میزان سکنه قلبی در دو گروه سیگاری و غیرسیگاری.....	۱۷۷
جدول ۸-۱۱:	معادله‌های مدل رگرسیون لجستیک برای مقایسه دو گروه مواجهه.....	۱۷۷
جدول ۸-۱۲:	تعداد و درصد افراد آلوده به کوری رودخانه به تفکیک مکان سکونت.....	۱۷۹
جدول ۸-۱۳:	رابطه بین گروه‌های سنی و mf.....	۱۷۹
جدول ۸-۱۴:	خروجی کامپیوتری برای مدل آلودگی به mf با محل سکونت و گروه‌های سنی.....	۱۷۹
جدول ۸-۱۵:	محاسبه برآورد کاپلان-مایر.....	۱۸۴
جدول ۸-۱۶:	خلاصه‌ای از روش‌های تحلیل داده‌ها.....	۱۸۶ الی ۱۸۸
جدول ۸-۱۷:	روش تحلیل داده‌ها در مطالعات کارآزمایی تصادفی.....	۱۹۰
جدول ۸-۱۸:	روش تحلیل داده‌ها در مطالعات متقاطع و طولی.....	۱۹۱
جدول ۸-۱۹:	مطالعه روش‌های مناسب برای تحلیل داده‌های مورد-شاهدی.....	۱۹۰
جدول ۹-۱:	مطالعه مورد-شاهدی فرضی که تمام موردها و شاهد‌های واجد شرایط بررسی شده‌اند.....	۱۹۵
جدول ۹-۲:	مطالعه مورد-شاهدی فرضی بر مبنای نمونه‌ای بدون تورش.....	۱۹۵
جدول ۹-۳:	مطالعه مورد-شاهدی با وجود تورش انتخاب.....	۱۹۵
جدول ۹-۴:	مطالعه مورد-شاهدی با وجود تورش انتخاب.....	۱۹۶
جدول ۹-۵:	مثال فرضی از اثر سو طبقه‌بندی بدون جهت برای دو رده دارای مواجهه و بدون مواجهه با میزان خطای ۳۰٪.....	۲۰۲
جدول ۹-۶:	مثال فرضی از نحوه اثر سو طبقه‌بندی جهت‌دار بر تشخیص مواجهه.....	۲۰۳
جدول ۱۰-۱:	مراحل کنترل کیفیت در پژوهش‌های بالینی.....	۲۲۰
جدول ۱۰-۲:	کنترل کیفیت مراحل آزمایشگاهی.....	۲۲۲
جدول ۱۰-۳:	کنترل کیفیت داده‌ها در مرحله طراحی.....	۲۲۳
جدول ۱۰-۴:	کنترل کیفیت داده‌ها در مرحله اجرا.....	۲۲۴
جدول الف-۱:	عناصر اصلی نوشتن پروپوزال.....	۲۳۰
جدول ج-۱:	انواع خطا در آزمون فرضیه.....	۲۵۵
جدول ج-۲:	محاسبه حجم نمونه.....	۲۵۹
جدول ج-۳:	مقادیر مختلف توزیع نرمال به ازای مقادیر معین خطای نوع اول و به ازای مقادیر معین توان و خطای نوع دوم.....	۲۶۱
جدول ج-۴:	عامل تعدیل برای محاسبه حجم نمونه هنگامی که حجم نمونه در گروه‌ها برابر نیست.....	۲۶۲
جدول ج-۵:	کاربردهای توان مطالعه برای تفسیر آزمون‌های آماری.....	۲۷۰

مفاهیم پایه



۱- مقدمه

داشتن تجربیات خام برای درک حقیقت کافی نیست و نکته مهم داشتن درکی صحیح و عمیق از واقعیت (یافته‌ها) و نتیجه‌گیری منطقی از آن‌هاست. این امر نکته‌ای اساسی و بااهمیت در کشف مجهولات علمی و تولید علم محسوب می‌شود. زمانی که نیوتن در زیر درخت سیبی نشسته بود و سیبی از درخت به طرف پایین سقوط کرد آنچه او را به کشف جاذبه زمین هدایت کرد همین درک صحیح و عمیق از چنین واقعیتی بود. مسلماً نه آن سیب اولین سیبی بود که از درخت سقوط می‌کرد و نه نیوتن اولین شاهد این اتفاق بود. میلیون‌ها نفر قبل از او چنین اتفاقی را نظاره‌گر بودند ولی ما اسحاق نیوتن را کاشف جاذبه زمین می‌دانیم. بستر اصلی برای درک عمیق و دقیق از حقایق، اندیشه است. از این‌رو است که تفکر و اندیشه را پایه تحقیق برمی‌شمارند و محقق یعنی کسی که به حقیقتی رسیده باشد فردی است که دائماً به حقایق پیرامون خود می‌اندیشد و از زاویه تخصصی خود سعی در کشف مجهولات و درک حقایق دارد.

نکته دیگر در تحقیق شجاعت اندیشیدن و شجاعت ابراز اندیشه است. بنا بر آنچه گفته شد مجموعه دستاوردهای علمی بشر حاصل اندیشه پژوهشگران گذشته و حال است. اندیشه‌ای که خطاپذیر است و دستاوردهایی که می‌توانند اصلاح‌پذیر باشند. آنچه در طول تاریخ موجب پیدایش و تکامل علم شده است نیز نتیجه شجاعت اندیشیدن و ابراز اندیشه بوده است. ارسطو بر این باور بود که زمین‌مرکز کائنات است و تمامی کرات جهان به دور زمین می‌چرخند؛ اما تفکر و تحقیق عمیق‌تر باعث شد این باور اصلاح شود، زیرا درک ما از حقیقت ممکن است از زمانی به زمان دیگر متفاوت باشد. ابداع ابزارهای جدید، کشف یافته‌های علمی مرتبط با یک موضوع و تغییر زاویه دید ممکن است زوایای پنهان دیگری از حقیقت را بنمایاند. برای مثال ابداع دستگاه ام-آر-آی به پزشکان کمک کرد تا آنچه قبلاً در معاینات معمول بر ایشان قابل مشاهده و درک نبود کشف کنند و یا اختراع میکروسکوپ امکان مشاهده موجوداتی را فراهم کرد که قبل از آن بشر درک ملموسی از آن‌ها نداشت. در درک حقیقت منظر یا دیدگاه نیز بسیار بااهمیت است. چون حقیقت چندوجهی است درک آن نیز ممکن است از وجوه مختلف امکان‌پذیر باشد و هرکس از نظرگاهی آن را استنباط نماید. تصحیح نظرگاه و منظر و گاه تغییر آن نیز می‌تواند در درک عمیق‌تر ما از حقایق مؤثر باشد.

از مطالب بیان شده تاکنون چنین استنتاج می‌شود که در کار پژوهش باید فرهنگ تحقیقی را جایگزین فرهنگ تقلیدی ساخت، ضمن آن که تفکر عمیق و دقیق و انتخاب منظر مناسب برای آزمودن فرضیات علمی، حل مسائل یا پاسخ دادن به مجهولات و سؤالات پژوهشی اجتناب‌ناپذیر است. چون علوم موجود حاصل اندیشه بشری است و این اندیشه خطاپذیر است علم نیز پویا و دینامیک است. اگر کسی منظر درستی در تحقیق برنگزیند منظور او با سوگرایی و خطا همراه خواهد بود. این امر گاه در بررسی‌های علمی اجتناب‌ناپذیر است و لذا دریچه‌ای را برای نقد و بررسی مطالب علمی بر روی ما باز می‌کند.

امروزه بیش از ۶۰ هزار عنوان مجله علمی، تنها به زبان انگلیسی در دنیا منتشر می‌شود که حاوی تازه‌های علمی و حاصل تلاش‌های پژوهش دانشمندان و پژوهشگران کشورهای مختلف است. از این تعداد بین ۲۰ تا ۲۵ هزار مجله مربوط به علوم پزشکی است که سالانه در حدود دو میلیون مقاله را شامل می‌شود. این مقالات پس از گذراندن مراحل مختلف بررسی و داوری امکان چاپ شدن یافته‌اند. هرچند تلاش بسیاری برای به حداقل رساندن خطاهای مربوط به نگارش، روش تحقیق و سایر نکات شده است، اما هنوز باب نقد و داوری به روی خواننده مطلب باز است [۱]. طبق بررسی‌های انجمن پزشکی انگلستان^۱ بیش از دوسوم مقالات ارائه شده به سری مجلات BMJ^۲ در مراحل اولیه برای چاپ رد می‌شود، از باقیمانده حدود ۹۸ درصد مقالات نیز نیاز به اصلاح دارند و تنها ۱۵ درصد مقالات تجربی علوم پزشکی که در این مجلات به چاپ رسیده‌اند مبنای علمی محکمی دارند و مابقی علیرغم تجدیدنظرهای مکرر قبل از چاپ واجد اشکالات و خطاهایی هستند که عمدتاً به خاطر خطاهای متدولوژیک و به علت عدم انطباق یافته‌های تحقیق با بحث و نتیجه‌گیری حاصل شده است. از مهم‌ترین معیارهای سنجش یک مقاله باارزش آن است که این مقاله پس از چاپ چند بار در مقالات دیگر پژوهشگران استفاده شده است. بررسی‌های فوق نشان داده است که حدود ۸۵ درصد مقالات چاپ شده در این مجلات هرگز در مقالات دیگر استفاده نشده یا به آنها استناد نشده است.

به‌طور کلی پنج ویژگی برای بیان و ارائه مطالب علمی می‌بایست مورد توجه قرار گیرد و هر مطلب علمی که از این ویژگی‌ها بهره‌مند باشد دارای استحکام و اعتبار بیشتری خواهد بود.
یک مطلب علمی باید:

- ۱- از ویژگی‌ها یا خصوصیات قابل اثبات شدن برخوردار باشد.
 - ۲- کلیت داشته و بتواند به تمامی مطالب یا اجزایی که ویژگی مشابهی دارند تعمیم داده شود.
 - ۳- از احساسات، هیجانات، تعصبات و نظریات شخصی، گروهی و قومی به دور باشد.
 - ۴- پویا بوده و سیر تکاملی داشته باشد.
 - ۵- منطبق بر یافته‌های علمی ثابت شده پژوهشگران قبلی بوده و بدیهیات علمی قبلی را نقض نکند.
- دست‌اندرکاران علوم پزشکی به علل مختلفی از جمله استفاده برای تهیه متون آموزشی، آشنا شدن با تازه‌های علمی، استفاده در برنامه‌ریزی‌ها و تصمیم‌گیری‌ها، تدوین، اجرا یا گزارش یک طرح تحقیقاتی و گاه

۱. British Medical Association

۲. British Medical Journals

داوری یک مقاله بنا به درخواست یک مجله علمی با مقالات تخصصی و پژوهشی سروکار دارند^۲. خواندن یک مقاله به هر منظوری انجام شود نگارش متون علمی و تنظیم مقالات علمی نیاز به شناخت نظاممند و اصولی آن دارد. جهت فراهم آوردن شرایط مناسب برای برآورده شدن منظور خواننده، اصول و مقرراتی در نگارش علمی وجود دارند که شناخت آنها ضروری است. یکی از مهم‌ترین ابزارهای لازم برای نقد مقالات شناخت خطاهای تحقیق است. هیچ پژوهشگری نمی‌تواند تحقیق خود را مطلقاً عاری از خطا بداند و در صورت هر مطالعه‌ای دارای درجه‌ای از خطا است که محقق با استفاده از تمهیدات لازم سعی می‌کند آن را به حداقل ممکن برساند [۲]. البته باید توجه داشت که خطاهای پزشکی نیز مبحث بااهمیت دیگری است که پرداختن مبسوط به آن نیازمند توجه جدی بوده و خارج از منظور این کتاب است. خطاهای پزشکی زمانی رخ می‌دهند که بخشی از آنچه برای مراقبت پزشکی برنامه‌ریزی شده حاصل نشود یا از ابتدا، برنامه‌ریزی نادرستی برای مراقبت‌ها صورت پذیرد. در حال حاضر این‌گونه خطاها سالانه باعث مرگ صدها هزار نفر و موجب میلیاردها تومان خسارت در جهان می‌شود. تحقیقات به‌روشنی نشان می‌دهد که بسیاری از خطاهای پزشکی قابل پیشگیری هستند. استانداردهای پروتکل‌های تشخیصی و درمانی و کنترل کیفی این خدمات از مهم‌ترین روش‌های پیشگیری از بسیاری از این خطاها هستند.

۱-۱ منابع خطا در تجزیه و تحلیل مطالعات پزشکی

عوامل مختلفی در برآورد بیشتر یا کمتر یک رابطه و یا اساساً مشاهده (غیرعلمی) تأثیرگذار هستند. این عوامل بر سه مؤلفه مهم تقسیم‌بندی می‌شوند که عبارت هستند از شانس، تورش و مخدوش شدن.

۱-۱-۱ تصادف یا شانس

فرض کنید مایل به تعیین فراوانی کمردرد در کارمندان یک اداره بزرگ هستیم. امکان دارد معاینه تمام کارمندان امکان‌پذیر نباشد و در عمل تنها قادر به تعیین فراوانی کمردرد در نمونه‌ای با حجم معین از این افراد هستیم. در این شرایط به‌صورت کاملاً تصادفی امکان دارد افراد نمونه، از کارمندان تقریباً سالم انتخاب شده باشد که این موضوع باعث کم برآورد شدن فراوانی تعداد افراد دچار کمردرد می‌شود. به‌عبارت‌دیگر خطای تصادفی باعث تحریف نتایج می‌شود.

هرقدر حجم نمونه انتخاب‌شده کمتر باشد، اثر تصادف در نتایج پررنگ‌تر خواهد بود. برای تعیین میزان تأثیرگذاری تصادف در نتایج مشاهده‌شده از نمونه‌گیری از روش‌های آماری استفاده می‌شود.

۱-۱-۲ تورش

یکی دیگر از عوامل مهم وجود خطای منظم در مرحله طراحی، هدایت و یا تجزیه و تحلیل اطلاعات هر مطالعه است. این عامل تورش نامیده می‌شود.

برای مثال در یک مطالعه مورد-شاهدی برای بررسی تأثیر مواجهه شغلی، با عامل یا شرایطی خاص (مثلاً

۱. Reading for Browsing, Reading for Information, Reading for Research

کار در صنایع خودروسازی کشور) اگر فرد معاینه کننده به نحوی نتواند موردها و شاهدها را به درستی شناسایی کند نتایج حاصل دارای تورش خواهند بود که این سوگرایی به تورش مشاهده گر معروف است.

۱-۱-۳ مخدوش شدن

عامل سوم در تحریف کردن روابط علمی، عامل مخدوش شدن یا مخدوش کنندگی است. این حالت به علت وابستگی چندگانه بین مواجهه، بیماری و عامل سومی که با هر دو وابسته است و به طور مجزا در ایجاد یا پیشرفت بیماری مؤثر است (عامل مخدوش کننده) رخ می دهد.

برای مثال در رابطه مشاهده شده بین آلودگی هوا و مرگومیر ناشی از بیماری های دستگاه تنفسی، وجود رابطه سببی بین آلودگی ناشی از ذرات ریز معلق در هوا و بیماری های دستگاه تنفسی فرضیه ای است که اکنون صحت آن مورد اشکال واقع شده زیرا عامل مخدوش کننده درجه حرارت محیط نیز در این پدیده مؤثر است. به طور کلی با سرد شدن هوا در زمستان، مرگومیر افزایش می یابد و هم چنین این موضوع با پدیده وارونه شدن جریان هوا^۱ در ارتباط است که باعث افزایش آلودگی هوا می شود. با در نظر گرفتن این عامل در تجزیه و تحلیل یافته ها، وابستگی بین آلودگی هوا و مرگومیر ناشی از بیماری دستگاه تنفسی ضعیف تر می شود.

استفاده از وسایل دقیق و قابل اعتماد، استفاده از آزمایش های ساده، دقیق و استاندارد، تکرار آزمایش ها و استفاده از میانگین چند سنجش یا اندازه گیری، انجام نمونه گیری کاملاً تصادفی، عدم استفاده از داوطلبان (حتی المقدور)، پوشش تمام منطقه مورد بررسی، استفاده از آزمون های آماری مناسب، انتخاب تعداد نمونه مناسب (نه کم و نه زیاد) و تدوین پروپوزال مناسب و بررسی و داوری درست از اقدامات عمومی پیشگیری کننده خطا در پژوهش محسوب می شود. در فصل های بعد به جزئیات بیشتر در شناسایی و کنترل خطاهای متداول در تحقیقات علوم پزشکی خواهیم پرداخت.

۱-۲ ساختار و ماهیت پژوهش های بالینی

هر پژوهش بالینی را می توان به دو رکن اصلی ساختار پژوهش و ماهیت پژوهش تقسیم بندی کرد. ساختار پژوهش شامل عناصر ملموس طرح مطالعه یعنی سؤال پژوهش، طرح مطالعه، اعضای مورد بررسی، نحوه اندازه گیری کمیت حجم نمونه است. یکی از اهداف محقق طراحی این عناصر به نحوی است که پروژه ارزان، سریع و آسان انجام شود.

ماهیت پژوهش شامل مکانیسم انجام تحقیق است. در هر تحقیق در ابتدا آنچه در نمونه رخ می دهد و در مرتبه بعد تعمیم آن به جامعه بزرگ تر که نمونه از آن گرفته شده است و استنباط های معتبر راجع به این وقایع به عنوان محصول و فایده هر پژوهش مدنظر است [۳].

۱. Inversion

۱-۲-۱ ساختار پژوهش

ساختار هر پژوهش در پروپوزال (پروتکل) هر تحقیق بیان می‌شود. پروپوزال قسمت مکتوب طرح پژوهش است. تنظیم صحیح پروپوزال به محققان کمک می‌کند تا مراحل تحلیل را به نحو منطقی طراحی کرده و به اندازه کافی در مسائل مهم و حیاتی متمرکز شوند. جدول ۱-۱ مؤلفه‌های اصلی هر پروپوزال را بیان می‌کند.

در این قسمت به شرح مختصر هر کدام از موارد ذکر شده در ساختار پروپوزال یک طرح پژوهش می‌پردازیم. تهیه و تنظیم فرم پروپوزال منظم و ساختاربندی شده علاوه بر اینکه نقش بسیار مهمی در پیدا کردن منابع مالی و ترغیب مسئولین در اجرای پژوهش دارند، به گروه محقق نیز کمک شایانی در فهم بهتر مراحل پژوهش و پیدا کردن کاستی‌ها و نواقص احتمالی مراحل مختلف پژوهش می‌کند. به همین علت، تهیه و تنظیم پروپوزال قبل از شروع پژوهش بسیار ضروری و راهگشا است. هرچند هر مؤسسه تحقیقاتی فرم پروپوزال ویژه‌ای دارد اما چارچوب کلی و محتوای این فرم‌ها تقریباً مشابه هستند.

ضمیمه الف، به نحوه تنظیم و تکمیل پروپوزال پژوهشی اختصاص دارد.

جدول ۱-۱: شمای کلی از پروپوزال طرح‌های تحقیقاتی

اجزا	هدف	عنصر
	پژوهش در ارتباط با کدام مشکل مطرح شده است؟ چرا این موضوع برای بررسی مهم است؟	سؤال پژوهش اهمیت موضوع (زمینه)
چارچوب زمانی روش اپیدمیولوژیکی معیارهای انتخاب طرح نمونه‌گیری متغیرهای وابسته	طرح مطالعه به چه صورت است؟ اعضای واجد شرایط برای بررسی کدام‌اند و چگونه انتخاب می‌شوند؟	طرح مسئله و بیان موضوع اعضای واجد شرایط
متغیرهای مخدوش‌کننده متغیر پاسخ فرضیه‌ها حجم نمونه روش تحلیل داده‌ها	اندازه کدام فاکتورها باید اندازه‌گیری و ثبت شود؟ وسعت مطالعه چقدر است و چگونه تحلیل خواهد شد؟	متغیرها روش‌های آماری

۱-۲-۱-۱ سؤال پژوهش

سؤال پژوهش در هر مطالعه، هدف کلیدی آن پژوهش است. این سؤال از یک نظریه غیر حتمی که پژوهشگر مایل به یافتن پاسخ آن است نشئت می‌گیرد. معمولاً سؤال اولیه ایجادشده در ذهن محققان به قدری گسترده است که دستیابی به آن یک‌باره امکان‌پذیر نیست و معمولاً به جزءهای کوچک‌تر و ساده‌تر، که عملی و امکان‌پذیر باشد، تقسیم می‌شود. به‌عنوان مثال ممکن است سؤال اولیه به این صورت باشد.

”آیا خانم‌های یائسه نیاز به مصرف هورمون دارند؟“

این سؤال برای عملی شدن ممکن است به جزءهای زیر تقسیم شود:

مصرف استروژن در خانم‌های یائسه تا چه اندازه شایع است؟
 آیا مصرف استروژن پس از یائسگی خطر پیشرفت بیماری عروق کرونر (CHD) را کاهش می‌دهد؟
 آیا فواید دیگری با استفاده از هورمون قابل تصور است؟

ویژگی‌های یک سؤال پژوهش مناسب

هر سؤال پژوهش باید دارای ۵ ویژگی زیر باشد:

- ♦ قابل اجرا باشد.^۵
- ♦ انجام آن برای محققان جذابیت داشته باشد.^۶
- ♦ جدید باشد.^۷
- ♦ از لحاظ اخلاقی قابل پذیرش باشد.^۸
- ♦ فواید و کاربردهای آن در مقایسه با هزینه‌ها، زمان و نیروی انسانی مصرف‌شده به صرفه باشد.^۹

۲-۱-۲-۲ قابل اجرا بودن مطالعه

یکی از موارد مهم در طراحی یک سؤال پژوهش آگاه بودن محققان به محدودیت‌ها و مشکلات موجود در اجرای مطالعه است. یکی از مسائل مهم در این زمینه امکان دسترسی به اعضا نمونه پیش‌بینی شده است. گام اول انجام برآوردهای اولیه از تعداد نمونه قابل دسترسی است. گام بعدی تعیین حجم نمونه لازم برای طراحی مطالعه است. در این مرحله باید امکان کاهش حجم مؤثر نمونه بر اثر گم‌شدن اعضای نمونه، کامل نبودن پیگیری نمونه‌ها و مواردی از این قبیل را نیز مدنظر قرارداد. حتی در طرح‌های مطالعاتی بسیار دقیق نیز غالباً پیش‌بینی‌های خوش‌بینانه‌ای در مورد اعضای قابل دسترسی در مطالعه صورت می‌پذیرد. گاهی اوقات برای اجتناب از این بی‌دقتی‌ها در برآورد اعضا مطالعه، انجام یک مطالعه راهنما بسیار مناسب است.

موضوع دیگر عملی بودن تحقیق از لحاظ فن و وجود مهارت‌های لازم برای اجرای پژوهش است. محققان باید از وجود افراد متخصص و ماهر و در دسترس بودن تجهیزات و لوازم مورد نیاز، تجربه و تبحر کافی برای ترغیب و تشویق اعضای نمونه برای شرکت در مطالعه و ادامه آن، جمع‌آوری متغیرهای تحقیق، توانایی مدیریت پروژه و تحلیل نتایج، اطمینان کسب نمایند. معمولاً بهترین راهبرد در این زمینه استفاده از منابع و روش‌های شناخته شده است، زیرا معمولاً فرایند ایجاد و یا توسعه روش‌های جدید و آزمون مهارت‌های نو وقت‌گیر، پرهزینه و نامطمئن هستند. هنگامی که به کار بردن روش‌های جدید، مانند استفاده از یک پرسشنامه جدید در مطالعه اجتناب‌ناپذیر است محققان باید توانایی طراحی و تکمیل موفقیت‌آمیز این روش‌ها را داشته باشند. مشاوره در مواردی که در حیطه تخصص محقق نیست معمولاً بسیار مفید است. استفاده از متخصصان در زمینه‌های مرتبط با موضوع پژوهش به عنوان همکار طرح اثر مطلوب‌تر خواهد داشت. به عنوان مثال غالباً لازم است از وجود یک

۱. Feasible

۲. Interesting

۳. Novel

۴. Ethical

۵. Relevant

متخصص آمار زیستی یا اپیدمیولوژی از ابتدای طراحی مطالعه تا مرحله تحلیل پردازش نتایج در گروه تحقیق بهره برد.

دقت و هزینه نیز جزء محدودیت‌هایی هستند که محققان باید در طراحی سؤال پژوهش به آن‌ها توجه داشته باشند. برای مطالعات پرهزینه، لازم است منابع مالی کافی قبل از طراحی مطالعه فراهم شوند. تعدیل و محدود کردن سؤال پژوهش نیز یک راهکار مناسب است زیرا پژوهش‌های کم‌هزینه سریع و آسان‌تر انجام می‌شوند. جدول ۱-۲ نواقص احتمالی سؤال پژوهش و طرح مطالعه و نحوه برطرف کردن آن‌ها را نمایش می‌دهد.

جدول ۱-۲: نواقص و راه بهبود سؤال پژوهش و طرح مطالعه

نقص بالقوه	راهکار
الف: سؤال پژوهش یکی از ۵ ویژگی ذکرشده را دارا نیست. ۱- سؤال پژوهش قابل اجرا نیست زیرا:	- محدود شدن به مجموعه متغیرهای کمتر - محدود کردن گستردگی سؤال پژوهش
- سؤال پژوهش بسیار گسترده است.	- گسترش معیارهای ورود به مطالعه - حذف، یا تعدیل معیارهای خروج از مطالعه - اضافه کردن منابع یا مراکز دیگر برای دستیابی به نمونه بیشتر - افزایش زمان مطالعه - به کار بردن راهکارهایی جهت کاهش حجم نمونه
- روش‌های موجود در مطالعه خارج از توان علمی و تجربه گروه تحقیق است.	- دعوت از متخصصان و افراد باتجربه برای همکاری در پژوهش - مشاوره و یا جستجو در متون تخصصی برای یافتن روش‌های جایگزین
- پژوهش پرهزینه است.	- توجه به امکان طراحی مطالعات کم‌هزینه‌تر - کاستن از اعضای نمونه و یا اندازه‌گیری‌ها - استفاده از ابزار و وسایل ارزان‌تر - کوتاه کردن پیگیری‌ها
۲- جذاب نبودن موضوع پژوهش، جدید نبودن یا کاربردی نبودن پژوهش.	- تغییر یا بهبود سؤال پژوهش - مشورت با افراد باتجربه
۳- عدم اطمینان از رعایت موازین اخلاقی.	- مشورت با کمیته اخلاقی - تغییر یا بهبود سؤال پژوهش
ب- برنامه‌ریزی مطالعه مبهم و غیردقیق است.	- بازبینی مجدد سؤال و اهداف پژوهش - بازنویسی طرح مطالعه بر اساس یک اسلوب استاندارد. - تعیین دقیق نحوه نمونه‌گیری، اندازه‌گیری متغیرها، انجام پیگیری‌ها و نحوه جمع‌آوری و تحلیل نتایج

۱-۲-۳ اهمیت موضوع

در این قسمت محققان باید اهمیت انجام چنین پژوهشی را برای مخاطب خود شرح دهند و مشخص نمایند چه

فوایدی از انجام این طرح متصور است و با انجام این پژوهش به کدام قسمت از مجهولات پاسخ داده می‌شود. محققان گاهی در این قسمت تنها سایر مطالعات مرتبط با این پژوهش که در گذشته انجام شده است را فهرست‌وار ذکر می‌نمایند.

۱-۲-۱-۴ طرح مطالعه

طرح مطالعه یکی از پیچیده‌ترین و فنی‌ترین قسمت‌های انجام هر مطالعه است. این موضوع به‌طور مفصل در فصل‌های دوم و سوم این کتاب مورد بررسی قرار خواهد گرفت. در جدول ۱-۳ یک دسته‌بندی کلی از انواع طرح‌های پژوهشی جهت انجام یک مطالعه فرضی برای ارزیابی اثر هورمون درمانی در خانم‌های یائسه ارائه شده است.

جدول ۱-۳: نحوه طراحی یک پژوهش بالینی برای مطالعه اثر هورمون درمانی پس از یائسگی برای جلوگیری از CHD

مطالعه	مشخصه اصلی	مثال‌ها
مطالعه هم‌گروهی	پیگیری یک گروه در طول زمان معین	محققان یک گروه از خانم‌های یائسه را برای چند سال از لحاظ بروز حمله قلبی، در دو گروه مصرف‌کننده هورمون و آن‌ها که هورمون مصرف نمی‌کنند به‌طور سالیانه مورد معاینه و بررسی قرار می‌دهند.
مطالعه مقطعی	بررسی یک گروه در یک دوره زمانی کوتاه‌مدت	در یک مقطع زمانی خانم‌های یائسه از لحاظ شیوع سکتة قلبی و مصرف یا عدم مصرف هورمون مورد بررسی و طبقه‌بندی قرار می‌گیرند.
مطالعه مورد-شاهدی	بررسی دو گروه بر مبنای وقوع یا عدم وقوع پیامد	محققان یک گروه از خانم‌های یائسه را که دچار حمله قلبی شده‌اند (موردها) با یک گروه سالم (بدون سکتة قلبی) از خانم‌های یائسه (شاهدها) از لحاظ مصرف یا عدم مصرف هورمون مقایسه می‌کنند.
۲- مطالعات مداخله‌ای:		
کارآزمایی تصادفی و کور شده	بررسی دو گروه که به تصادف هورمون یا دارونما مصرف می‌کنند و از نوع درمان اعمال شده بی‌اطلاع هستند.	محققان به‌طور تصادفی به افراد تحت مطالعه هورمون یا دارونما تجویز می‌کنند و پس از پیگیری هر دو گروه طی مدت چند سال میزان بروز حمله قلبی را بین دو گروه مقایسه می‌نمایند.

۱-۲-۱-۵ افراد تحت مطالعه

یکی از مهم‌ترین تصمیم‌گیری‌ها در هر مطالعه پژوهشی تصمیم‌گیری برای مشخص کردن معیارهای مورد نظر در انتخاب اعضای نمونه است. به‌طور کلی هر طرح باید دارای دو معیار زیر باشد: معیارهای ورود به مطالعه که جمعیت هدف را مشخص می‌کند و معیارهای خروج از مطالعه که مشخص می‌کند کدام‌یک از اعضای مورد مطالعه فاقد شرایط لازم برای پیگیری و بررسی بیشتر است و

باید از مطالعه کنار گذاشته شود.

معیارهای ورود، ویژگی‌های اصلی جامعه هدف را معین می‌کند (جدول ۴-۱). فرایند تعیین مشخصات بالینی غالباً شامل انجام قضاوت‌های دشواری درباره انتخاب عوامل مهم و تأثیرگذار در پژوهش و نحوه تعریف آن‌هاست. برای مثال محققان باید تصمیم بگیرند که «اصطلاح «وضعیت عمومی مناسب» در عمل چگونه باید تعریف شود؟» فرض کنید این اصطلاح باید برای تحقیقی راجع به پوکی استخوان تعریف شود. در این صورت منطقی است بیماری‌هایی که تأثیر چندانی بر این بیماری ندارند و یا کشندگی سریع ایجاد نمی‌کنند (نظیر فشارخون) را جز رده سالم قرار دهیم ولی بیماری‌هایی که باعث ایجاد پوکی استخوان می‌شوند مانند فلج هر دو پا و یا به علت کشندگی سریع پیگیری‌ها را مختل می‌کنند (نظیر متاستازها) را جز افراد بیمار طبقه‌بندی کنیم. در مورد مشخصات جمعیت‌شناختی (نظیر سن، جنس و نژاد) نیز همواره باید حالت متعادلی که قابلیت تعمیم نتایج و کارایی مطالعه را به همراه داشته باشد، در نظر گرفته شود. در این مقوله به‌طور مطلق هیچ تصمیم اشتباه یا درستی وجود ندارد. نکته مهم معنادار بودن تصمیم‌ها، ثابت بودن معیارها در طول مطالعه و قابل‌مقایسه بودن مطالعه با سایر تحقیقات است.

معیارهای خروج از مطالعه به مشخص کردن زیرگروه‌هایی از جامعه می‌پردازد که به‌طور بالقوه امکان حضور در مطالعه را بر اساس سؤال تحقیق دارا هستند، اما به علت مسائل و مشکلاتی نظیر پایین بودن کیفیت داده‌ها، قابل‌قبول نبودن انتخاب تصادفی نمونه‌ها و یا وجود عوارض جانبی باید از مطالعه کنار گذاشته شوند (جدول ۴-۱).

استفاده از معیارهای خروج با کاستن از میزان تعمیم‌پذیری نتایج باعث افزایش قابلیت اجرا و سرراست شدن تفسیر نتایج می‌شود. برای مثال بررسی افرادی که دارای محل سکونت مشخصی نیستند (کارتن خواب‌ها)، باعث افزایش تعمیم‌پذیری نتایج می‌شود و این امکان را به وجود می‌آورد که محققان وضعیت اقتصادی و اجتماعی را با اندازه شیوع در زیرگروه‌ها مقایسه نمایند. باین حال فواید ذکرشده با مشکلاتی از قبیل افزایش حجم نمونه، افزایش میزان گم‌شده و پیچیده شدن فرایند پیگیری همراه است. به همین علت امکان دارد محققان تصمیم بگیرند کارتن خواب‌ها از ابتدا به مطالعه وارد نشوند. در این صورت تفکیک افراد تحت مطالعه به افراد دارای محل سکونت مشخص و آن‌هایی که محل ثابتی برای زندگی ندارند، یکی دیگر از مسائل پیش روی محققان خواهد بود. در مورد کارآزمایی‌های بالینی معیارهای خروج می‌تواند تا حدی متفاوت باشند. برای مثال وجود عوارض جانبی ناشی از مصرف دارو، مسائل اخلاقی پیش‌آمده به علت مصرف نکردن دارو، مؤثر نبودن داروی مورد استفاده بر روی بیماران، احتمال گم‌شدن بیماران در پیگیری و واجد شرایط نبودن بیمار برای دریافت دارو از جمله مواردی هستند که می‌توانند معیارهای خروج از مطالعه باشند [۴].

ضمیمه ب، به بررسی دو روش مهم در جمع‌آوری اطلاعات یعنی پرسشنامه و مصاحبه می‌پردازد.

۱-۲-۱ روش‌های نمونه‌گیری

یکی دیگر از قسمت‌های مهم هر پژوهش تصمیم‌گیری در مورد نحوه جمع‌آوری واحدهای مطالعه

است. این واحدها می‌توانند انسان‌ها، بیمارستان‌ها، منازل مسکونی و داروخانه‌ها یا فروشگاه‌ها باشند. مهم‌ترین قسمت درباره نحوه جمع‌آوری واحدها تصمیم‌گیری راجع به استفاده از اطلاعات موجود در پرونده‌ها، داده‌های ثبتی و سوابق مکتوب یا انجام نمونه‌گیری و تکمیل اطلاعات مورد نیاز پژوهش به وسیله پرسشنامه و مصاحبه است. محققان باید در مرحله تدوین پروپوزال مسائلی نظیر هزینه‌های انجام نمونه‌گیری یا جمع‌آوری اطلاعات موجود، خطاهای موجود در نمونه‌گیری و آمارهای ثبتی، زمان انجام طرح، امکانات و کارکنان مجرب و آموزش‌دیده را مدنظر قرار داده و راجع به این موضوع تصمیم‌گیری نمایند.

جدول ۱-۴: تنظیم معیارهای انتخاب برای یک کارآزمایی بالینی که رابطه میزان ذخیره کلسیم و ابتلا به پوکی استخوان را بررسی می‌کند

ملاحظات	مثال‌ها
معیارهایی که مشخصات جامعه مورد بررسی را از لحاظ مرتبط بودن با سؤال پژوهش و کاراتر شدن مطالعه معین می‌کند:	یک کارآزمایی با دوره پیگیری پنج‌ساله در مورد بررسی رابطه بین ذخیره کلسیم و ابتلا به پوکی استخوان می‌تواند شامل موارد زیر باشد:
مشخصات جمعیت‌شناختی	زنان سفیدپوست ۶۰-۵۰ ساله
ویژگی‌های بالینی	دارای وضعیت عمومی مناسب
ویژگی‌های جغرافیایی	بیمارانی که به بیمارستان محل انجام تحقیق مراجعه می‌کنند
ویژگی‌های مختص به هر مطالعه	بیمارانی که طی ماه‌های اردیبهشت تا آخر بهمن‌ماه همان سال به بیمارستان مراجعه کنند
معیارهای خروج:	مشخص کردن معیارهایی که تعیین می‌کند کدام زیرگروه‌ها از جامعه در مطالعه مورد بررسی قرار نخواهند گرفت:
	مطالعه پوکی استخوان می‌تواند شامل معیارهای خروج از مطالعه زیر باشد:
	محتمل بودن گم‌شدن بیماران در پیگیری‌ها
	عدم توانایی در فراهم آوردن داده‌های باکیفیت
	در معرض عوارض جانبی قرار داشتن
	تصمیم بیمار برای ترک شهر محل اقامت
	مسلمت نبودن به زبان فارسی
	سارکوییدوز ^۱ به علت وجود کلسیم بالا در خون
غیراخلاقی بودن کارآزمایی روی اعضا	مصرف کورتیکواستروئید ^۲

به‌طور کلی دو روش نمونه‌گیری وجود دارد که به روش نمونه‌گیری احتمالاتی و روش نمونه‌گیری غیر احتمالاتی معروف‌اند. نمونه‌گیری احتمالاتی، اصطلاحی است که دمینگ^۳ در سال ۱۹۵۰ به کار برد. در نمونه‌گیری احتمالاتی، هر واحد نمونه با احتمالی مشخص از جامعه استخراج می‌شود. این شیوه از سال ۱۹۴۰ به بعد سریعاً گسترش یافته و تقریباً در همه کاربردها جانشین نمونه‌گیری غیر احتمالاتی شده است.

۱. Sarcoidosis

۲. Corticosteroids

۳. Deming

در شیوه غیر احتمالاتی، به واحدهای منتخب، احتمال‌هایی را نسبت نمی‌دهند و در نتیجه تعیین دقت نتایج نمونه‌گیری میسر نیست. نمونه‌گیری‌های غیر احتمالاتی که در دهه‌های ۱۹۲۰ و ۱۹۳۰ مورد استفاده بودند عمدتاً نمونه‌گیری مبتنی بر قضاوت یا قصدی و نمونه‌گیری سهمیه‌ای بودند.

نمونه‌گیری احتمالاتی شامل روش‌هایی نظیر نمونه‌گیری تصادفی، نمونه‌گیری با احتمال متغیر، نمونه‌گیری طبقه‌ای، نمونه‌گیری خوشه‌ای (یک مرحله‌ای، دومرحله‌ای و چندمرحله‌ای) و نمونه‌گیری منظم است. مباحث فنی مربوط به مفاهیم، نحوه انجام و محاسبه برآوردهای موردنظر از این روش‌ها خارج از حیطه این کتاب می‌باشند و خوانندگان برای مطالعه بیشتر می‌توانند منابعی نظیر [۵] و [۱۰] را مطالعه نمایند.

۱-۲-۱-۷ متغیرها

معمولاً هر طرح پژوهشی شامل اندازه‌گیری، مشاهده و ثبت برخی متغیرها است. نوع این متغیرها، نحوه اندازه‌گیری، میزان دقت و اعتبار نتایج موضوعی است که در فصل پنجم مورد بررسی قرار خواهد گرفت.

به‌طور کلی متغیرها دارای مقیاس کیفی و کمی هستند. متغیرهای کیفی به دو رده کلی متغیرهای اسمی و رتبه‌ای تقسیم‌بندی می‌شوند. متغیرهای اسمی، متغیرهایی هستند که فاقد ماهیت رتبه‌ای هستند. مثال‌هایی از این نوع متغیر عبارت‌اند از مذاهب مختلف که می‌تواند شامل رسته‌های: مسیحی، مسلمان و یهودی باشد، نحوه رفتن به محل کار با رسته‌های اتومبیل، دوچرخه، اتوبوس، مترو و پیاده. در متغیرهای اسمی، نحوه فهرست‌بندی دارای الگوی رتبه‌ای نیست.

بسیاری از متغیرها دارای مقیاس رتبه‌ای هستند. به این متغیرها، متغیرهای رتبه‌ای گفته می‌شود. مثال‌هایی از این نوع مقیاس عبارت‌اند از طبقه اجتماعی شامل: مرفه، متوسط و فقیر و وضعیت بیماری شامل: خوب، مناسب، وخیم و بحرانی. متغیرهای رتبه‌ای دارای ترتیب (رتبه‌بندی) هستند، اما فاصله بین هر کدام از رده‌ها معلوم نیست.

متغیرهای کمی، دارای مقیاس عددی هستند و فاصله بین هر کدام از مقادیر مشخص است. برای مثال فشارخون، طول عمر، میزان درآمد ماهانه مثال‌هایی از این قبیل هستند. گاهی اوقات به متغیرهای کمی، متغیرهای نسبتی هم گفته می‌شود در این نوع متغیرها، نسبت دو مقدار نیز معنی‌دار هستند.

نحوه اندازه‌گیری هر متغیر ماهیت آن را مشخص می‌کند. برای مثال اگر تحصیلات به‌صورت مدرسه دولتی و مدرسه غیرانتفاعی طبقه‌بندی شود، دارای مقیاس اسمی است و اگر بر اساس میزان تحصیل طبقه‌بندی شود، یعنی ابتدایی، راهنمایی، دبیرستان، فوق‌لیسانس و بالاتر دارای مقیاس رتبه‌ای و هنگامی که بر اساس میزان سال‌های تحصیل طبقه‌بندی می‌شود، یعنی... ۳، ۲، ۱، ۰ دارای مقیاس کمی است.

متغیرها بر اساس رابطه‌ی که با یکدیگر دارند به سه گروه تقسیم می‌شوند.

متغیر پاسخ (وابسته): این متغیرها پیامد یا پاسخ نهایی و یا معلول وقایع قبلی هستند. برای مثال در رابطه هوش و پیشرفت تحصیلی، پیشرفت تحصیلی متغیر پاسخ و در رابط سرطان پوست و اشعه مستقیم خورشید، سرطان پوست متغیر پاسخ است.

متغیر مستقل (متغیر مواجهه در مطالعات مشاهده‌ای و یا مداخله در مطالعات تجربی): این

متغیرها نقش علت و یا توجیه پیامدهای اتفاق افتاده را به عهده دارند و به نحوی (که می‌تواند علت و معلولی باشد) بر متغیرهای دیگر تأثیر می‌گذارند. در دو مثال بالا هوش و اشعه مستقیم خورشید متغیرهای مستقل هستند. **متغیر مخدوش‌کننده یا زمینه‌ای:** معمولاً در هنگام بررسی رابطه بین دو متغیر، متغیرهای دیگری وجود دارند که بر قدرت رابطه تأثیرگذار هستند. برای مثال برای بررسی یک داروی جدید در پایین آوردن فشارخون عواملی نظیر سن، جنس، نژاد، رژیم غذایی، بیماری‌های دیگری که فرد به آن‌ها مبتلا است، میزان فعالیت روزانه و غیره می‌توانند تأثیرگذار باشند. این متغیرها می‌توانند همراه یا واسطه‌ای یا مخدوش‌کننده باشند. در این حالت مشاهده رابطه واقعی بین فشارخون و مصرف دارو تنها با شیوه‌های ویژه آماری که در فصل‌های بعدی به آن‌ها اشاره می‌شود امکان‌پذیر است.

۱-۲-۱-۸ فرضیه پژوهش

یکی از ارکان هر مطالعه، سؤال پژوهش است. در اکثر مواقع لازم است سؤال پژوهش به نسخه‌ای فنی به نام فرضیه تبدیل شود تا به وسیله آن بتوان مبنای آزمون‌های آماری متناسب با پژوهش را طرح‌ریزی کرد. هر فرضیه خلاصه‌ای از ارکان مطالعه: نمونه، طرح مطالعه، متغیر پاسخ و متغیر مستقل را در بر دارد. ارائه فرضیه در مطالعات توصیفی که به بیان نحوه پخش ویژگی‌ها و صفات افراد تحت مطالعه در جامعه می‌پردازد (نظیر مطالعات تعیین شیوع) لازم نیست اما ساختن فرضیه در مطالعات تحلیلی که هدف مقایسه یافته‌ها در دو یا چند گروه متفاوت است (مانند مطالعات کارآزمایی بالینی) الزامی است. به علت آنکه در اکثر مطالعات مشاهده‌ای (نظیر مطالعات هم‌گروهی و مورد-شاهدی) هدف مقایسه گروه‌های مختلف با یکدیگر است حداقل به ساخت یک فرضیه نیاز داریم.

اگر هرکدام از عبارتهای زیر در سؤال پژوهش وجود داشته باشد "بزرگ‌تر از" (greater than)، "کوچک‌تر از" (less than)، "باعث می‌شود" (causes)، "منجر می‌شود" (lead to)، "در مقایسه با" (compared with)، "شباهت بیشتری به" (more likely than)، "وابسته با" (associated with)، "مرتبط با" (related to)، "مشابه با" (similar to) و یا "همبسته با" (correlated with)، مطالعه، دیگر یک مطالعه توصیفی ساده نخواهد بود و باید برای آن فرضیه تدوین شود.

ویژگی‌های فرضیه خوب

هر فرضیه باید ساده و مشخص بیان شده باشد.

ساده یا مرکب: در فرضیه ساده یک متغیر وابسته و یک متغیر مستقل وجود دارد (شیوه زندگی کم‌تحرک با افزایش خطر وجود پروتئین در ادرار^۴ در بیماران دیابتی همراه است). در فرضیه مرکب بیش از یک متغیر مستقل (شیوه زندگی کم‌تحرک و مصرف الکل با افزایش خطر وجود پروتئین در ادرار در بیماران دیابتی وابسته است) یا متغیر وابسته (مصرف الکل با افزایش خطر وجود پروتئین در ادرار و نوروپاتی در بیماران دیابتی وابسته

۱. Proteinuria

است) وجود دارد. فرضیات مرکبی نظیر موارد بالا برای بررسی به‌وسیله یک آزمون آماری ساده مناسب نیستند و به‌سادگی قابل تفکیک به دو یا چند فرضیه ساده هستند.

با این حال برخی اوقات از متغیرهای مستقل یا وابسته مرکب نیز استفاده می‌شود (استعمال سیگار، قلیان یا پیپ با افزایش خطر وجود پروتئین در ادرار در بیماران دیابتی وابسته است).

مشخص یا مبهم: هر فرضیه مشخص هیچ نکته گنگ و دو پهلوئی درباره موضوع، متغیرها و نوع آزمون آماری باقی نمی‌گذارد (سابقه مصرف داروهای ضدافسردگی تری‌سایکلید که با مراجعه به پرونده دارویی بیماران به دست می‌آید، در بیماران بستری شده به علت سکت قلبی بستری شده در بیمارستان شهید مدرس دانشگاه علوم پزشکی شهید بهشتی در سال گذشته شایع‌تر از گروه کنترلی است که شامل بیماران بستری شده در همین بیمارستان در طی این دوره به علت پنومونی است).

معمولاً به‌وسیله فرضیه تحقیق نوع متغیر از نظر اسمی، پیوسته یا رتبه‌ای بودن معین می‌شود. در مواردی که این موضوع واضح نباشد، نوع متغیر باید مشخص شود (مصرف الکل بر حسب (mg/day). با افزایش خطر وجود پروتئین در ادرار ($> 30\text{mg/dl}$) در بیماران دیابتی وابسته است).

اگر هر کدام از تعاریف و اصلاحات، واحدها و مقیاس‌ها در جای دیگر از پروپوزال تعریف شده باشد، آوردن همه آن‌ها در فرضیه که باعث طولانی شدن جمله‌های فرضیه می‌شود الزامی نیست.

فرضیه‌های آماری

برای انجام آزمون‌های آماری، فرضیه‌ها باید به نحوی بیان شوند که تفاوت‌های منتظره بین گروه‌های مختلف مطالعه قابل فرمول‌بندی شود.

فرضیه صفر و فرضیه مقابل: فرضیه صفر بر این مبنا تدوین می‌شود که هیچ ارتباطی بین متغیر مستقل و متغیر پاسخ در جامعه وجود ندارد (هیچ ارتباطی بین فراوانی مصرف آب سالم و وجود زخم معده وجود ندارد). فرضیه صفر مبنای اساسی شکل‌گیری آزمون‌های آماری است. با فرض عدم وابستگی بین گروه‌های جامعه، آزمون‌های آماری احتمال وجود رابطه‌ای شانس بین متغیر پاسخ و مستقل را محاسبه می‌کنند. جمله‌ای که وجود ارتباط بین متغیرهای پاسخ و مستقل را بیان می‌کند فرضیه مقابل است (بین فراوانی مصرف آب سالم و بهداشتی و وجود زخم معده رابطه‌ای معکوس وجود دارد).

فرضیه‌های مقابل یک‌طرفه و دوطرفه: در فرضیه یک‌طرفه جهت ارتباط (مستقیم یا معکوس) بین متغیر پاسخ و مستقل مشخص می‌شود. به‌عنوان مثال هنگامی که بیان می‌کنیم "فراوانی مصرف آب سالم و بهداشتی در اشخاصی که مبتلا به زخم معده شده‌اند در مقایسه با افراد سالم کمتر است" یک فرضیه یک‌طرفه که در آن رابطه معکوسی عنوان شده، ساخته‌ایم. در فرضیه دوطرفه تنها مسئله وجود ارتباط (نه جهت آن) منظور می‌شود. به‌عنوان مثال هنگامی که بیان می‌کنیم "فراوانی مصرف آب سالم و بهداشتی در اشخاص مبتلا به زخم معده در مقایسه با افراد سالم، متفاوت است" یک فرضیه دوطرفه ساخته‌ایم.

فرضیه یک‌طرفه در شرایط خاصی نظیر مواقعی که ارتباط تنها از یک جهت مهم باشد یا از نظر بیولوژیکی

معنی‌دار باشد استفاده می‌شود. برای مثال در میزان اثربخشی یک داروی جدید در برابر دارونما امکان تأثیر بیشتر داروی جدید از دارونما در بهبود بیماری همواره مطلب درستی نیست؛ اما موضوع حائز اهمیت کارایی داروی جدید است. بنابراین فرضیه یک‌طرفه انتخاب مناسبی است. هنگامی که بر اساس مطالعات قبلی مدارک و قراینی مبنی بر تأیید رابطه در یک‌جهت وجود دارد (مانند بررسی ارتباط بین استعمال دخانیات و افزایش خطر سرطان مغز) نیز استفاده از فرضیه یک‌طرفه مناسب است. باین‌حال پژوهشگران باید به این نکته توجه کنند که بسیاری از فرضیاتی که دارای پشتیبان‌های محکم بوده‌اند (نظیر درمان با بتاکارتن در کاهش خطر ابتلا به سرطان ریه) قویاً در کارآزمایی‌های تصادفی رد شده‌اند. انتخاب فرضیه یک‌طرفه یا فرضیه دوطرفه در تعیین حجم نمونه مهم هستند.

۱-۲-۱-۹ محاسبه حجم نمونه

پس از معین شدن موضوع تحقیق، نوع مطالعه و روش نمونه‌گیری، گام بعدی تصمیم‌گیری درباره تعداد اعضای نمونه است. حتی در طرح‌های بسیار ساده و واضح از لحاظ محتوی و روش اجرا هنگامی که حجم نمونه کافی نباشد قادر به بررسی علمی و دقیق سؤال پژوهش نیستیم. در عوض وقتی حجم نمونه بیش از اندازه زیاد باشد، در وقت و هزینه و امکانات اسراف می‌شود. هدف در تعیین حجم نمونه، برآورد بهینه تعداد اعضای لازم در مطالعه است.

هرچند موضوع تعیین حجم نمونه ابزار مفیدی در پژوهش بشمار می‌آید، هدف مهم دیگر در محاسبه حجم نمونه دستیابی به یک دورنمای روشن از میزان مقبولیت آزمون‌های آماری به‌کاررفته در مطالعه است. اعتبار حجم نمونه به‌دست‌آمده مستقیماً به میزان اعتبار اطلاعات و داده‌هایی که محاسبه بر مبنای آن‌ها صورت گرفته مربوط می‌شود (که معمولاً در حد حدس و گمان باقی می‌ماند) محاسبه حجم نمونه، قابل انجام بودن طرح مطالعه را معین می‌کند. محاسبه حجم نمونه در مرحله اولیه طراحی مطالعه، هنگامی که هنوز امکان ایجاد تغییرات اساسی در طرح مطالعه وجود دارد به محققان کمک می‌کند در صورت زیاد بودن بیش‌ازحد تعداد نمونه‌های موردنیاز، به تغییر روش‌ها، کنترل متغیرهای دیگر و یا محدود کردن مطالعه اقدام کنند. بنابراین، محاسبه حجم نمونه باید در مراحل اولیه مطالعه و در هنگام طراحی پروپوزال که هنوز امکان تغییرات اساسی وجود دارد، انجام شود.

در این قسمت به بررسی برخی مفاهیم آماری پایه جهت محاسبه حجم نمونه می‌پردازیم.

حداقل اختلاف مهم از نظر بالینی

شانس این‌که یک مطالعه مبتنی بر نمونه‌گیری، رابطه‌ای میان متغیر پاسخ و متغیر مستقل را نمایان سازد به مقدار بسیار زیادی به قدرت رابطه در جامعه هدف بستگی دارد. برای مثال اگر میانگین قند خون ناشتا در زنان دیابتی که ورزش می‌کنند 20mg/dl پایین‌تر از زنانی که ورزش نمی‌کنند باشد نمایان ساختن این اختلاف (و در نتیجه استنباط وجود رابطه میان ورزش و کنترل دیابت) آسان‌تر از زمانی است که این اختلاف فرضاً 2mg/dl باشد.

متأسفانه، محققان معمولاً از بزرگ یا کوچک بودن اختلاف واقعی در جامعه بی‌اطلاع هستند. درواقع یکی از

اهداف مطالعه معمولاً برآورد اندازه اثر^۵ است. به همین علت معمولاً در مطالعات، از حداقل میزان اختلاف مهم از نظر بالینی و یا حداقل میزان اختلافی که محقق مایل است در مطالعه نمایان شود استفاده می‌شود. این کمیت در پژوهش‌های بالینی اندازه اثر خوانده می‌شود. انتخاب اندازه اثر مناسب، مشکل‌ترین قسمت برآورد حجم نمونه است. به این منظور ابتدا باید اطلاعات حاصل شده از مطالعات قبلی بررسی شوند تا حدس‌های قابل‌قبولی از اندازه اثر به دست آید. وقتی هیچ اطلاع قابل‌قبولی در دست نباشد احتمالاً انجام یک مطالعه راهنما الزامی است.

خطاهای نوع اول، دوم و توان آزمون

گاهی تنها بر اساس شانس، نمونه انتخاب‌شده، نماینده مناسبی از جامعه مطالعه نیست. در این حالت نتایج حاصله از نمونه، واقعیت‌های موجود در جامعه را منعکس نمی‌کند و استنباط‌های صورت گرفته نادرست است. خطای نوع اول، هنگامی رخ می‌دهد که محقق فرضیه صفر را رد می‌کند (بر اساس نمونه در دست) اما در جامعه هدف، فرضیه صفر واقعاً درست است. خطای نوع دوم هنگامی رخ می‌دهد که محقق فرضیه صفر را رد نمی‌کند (بر اساس نمونه در دست) در صورتی که در جامعه هدف فرضیه صفر واقعاً درست نیست. در این صورت با چهار حالت مواجه هستیم. جدول ۱-۵ این حالت‌ها را بررسی می‌کند.

جدول ۱-۵: حالت‌های مختلف قضاوت‌های آماری و واقعیت‌های جامعه

نتایج به‌دست‌آمده از نمونه	واقعیت‌های جامعه	
	وجود رابطه بین متغیرهای پاسخ و مستقل	عدم وجود رابطه بین متغیرهای پاسخ و مستقل
رد کردن فرضیه صفر	صحیح	خطای نوع I
رد نکردن فرضیه صفر	خطای نوع II	صحیح

تعیین کردن بیشترین مقدار مجاز برای وقوع هر کدام از خطاهای نوع اول و دوم یکی از مهم‌ترین قسمت‌های هر مطالعه است. احتمال وقوع خطای نوع اول را α یا سطح معنی‌داری می‌نامند.

برای مثال در مطالعه اثر ورزش در کنترل دیابت، اگر سطح α معادل ۰/۰۵ در نظر گرفته شود به این معنی است که محقق حداکثر شانس معادل ۵٪ را برای رد نادرست فرضیه صفر (یعنی استنباط این موضوع که ورزش و میزان قند ناشتای خون با هم مرتبط هستند در صورتی که واقعاً این دو با هم رابطه نداشته باشند) قبول کرده است.

احتمال وقوع خطای نوع دوم را β می‌نامند و بر این اساس توان آزمون به صورت $1-\beta$ تعریف می‌شود. اگر β را برابر ۰/۱ در نظر بگیریم در این صورت پذیرفته‌ایم در ۱۰٪ مواقع با اندازه اثری که از قبل معین شده است، رابطه‌ای که حقیقتاً وجود دارد را نادیده بگیریم. این موضوع معادل توان آزمون ۰/۹ که به معنی انجام استنباط واقعاً درست در ۹۰٪ مواقع است، است. برای مثال فرض کنید که ورزش کردن واقعاً ۲۰ mg/dl قند خون ناشتا را در افراد دیابتی پایین می‌آورد. حال اگر محقق نمونه‌های متعددی از یک جامعه بگیرد و هر مرتبه همان رویه قبلی را برای تمام اعضای نمونه‌ها تکرار کند در ۹۰ مطالعه از مجموع ۱۰۰ مطالعه، محقق قادر به رد فرضیه

۱. Effect size

صفر مبنی بر عدم وجود رابطه می‌شود. با این حال این موضوع به معنی آن نیست که محقق قادر است اندازه اثرات کوچک‌تر مثلاً کاهش 10mg/dl را نیز با همین حجم نمونه با شرایط بالا آشکار سازد. در حالت ایده‌آل α و β باید صفر باشند، اما در عمل باید سعی شود این دو خطا تا حد امکان کوچک شوند. این موضوع معمولاً با بزرگ‌تر کردن حجم نمونه امکان‌پذیر است. هدف اصلی در محاسبه حجم نمونه کنترل α و β در سطحی قابل قبول برای محقق است، بدون آنکه مطالعه بیش‌ازحد لزوم مشکل، وقت‌گیر و پرهزینه شود.

تغییرات متغیر پاسخ

نکته مهم دیگری که در تعیین حجم نمونه به آن توجه می‌شود اندازه تغییرات (انحراف معیار) متغیر پاسخ است. هر قدر که انحراف معیار متغیر پاسخ در افراد تحت مطالعه بیشتر باشد، بیشتر امکان دارد که در مقادیر پاسخ گروه‌ها همپوشانی ایجاد شود. در این صورت آشکار شدن تفاوت‌ها در گروه‌ها مشکل‌تر است و طبیعتاً به حجم نمونه بیشتری نیاز است.

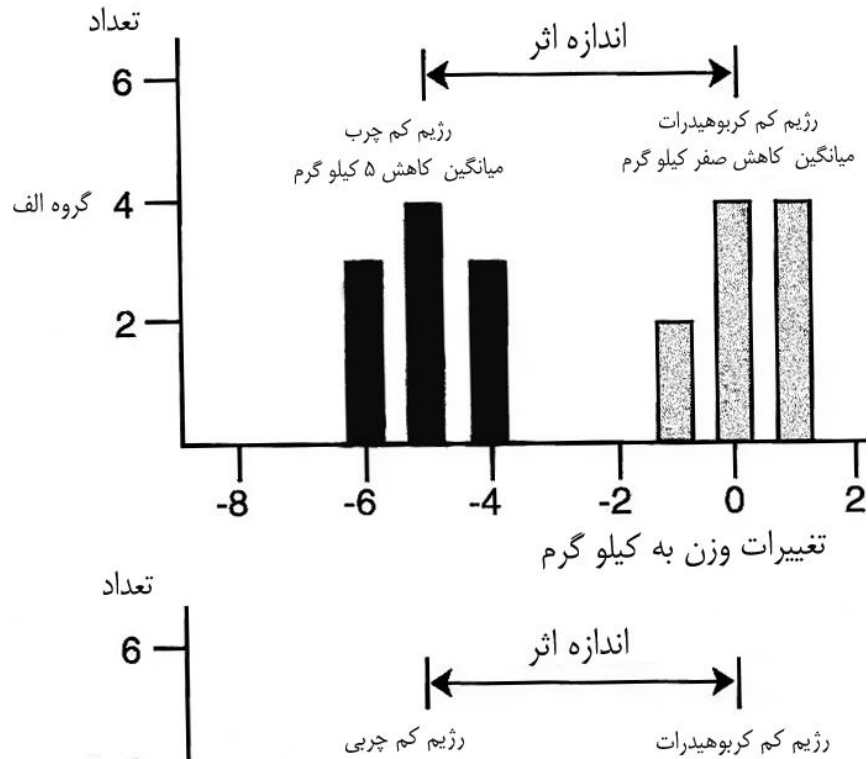
به عنوان مثال مطالعه‌ای را برای بررسی دو رژیم غذایی برای کاهش وزن در ۲۰ داوطلب در نظر بگیرید. این رژیم‌ها عبارت‌اند از رژیم دارای چربی کم و رژیم دارای هیدروکربن کم. اگر تمام داوطلب‌های با رژیم چربی پایین به طور میانگین ۵ کیلوگرم وزن کم کنند و داوطلب‌های با رژیم هیدروکربن پایین وزن کم نکنند (اندازه اثر برابر ۵ کیلوگرم) می‌توان استنباط کرد رژیم با چربی کم بهتر است. (نمودار ۱-۱ الف)؛ اما اگر کم شدن وزن در گروه رژیم چربی کم همان ۵ کیلوگرم باشد و گروه هیدروکربن کم بدون تغییر باقی بماند این امکان وجود دارد که نظیر نمودار ۱-۱ ب توزیع پراکندگی کاهش دو گروه وزن همپوشانی وجود داشته باشد. در چنین موقعیت‌هایی استنباط راجع به وجود اختلاف بین دو رژیم به حجم نمونه بزرگ‌تری نیاز دارد [۷].

ضمیمه ج، به بررسی روش‌های محاسبه حجم نمونه برای برخی از شاخص‌های آماری اختصاص داده شده است.

مقدار احتمال

پس از جمع‌آوری و تحلیل داده‌ها برای رد کردن فرضیه صفر به معیار استاندارد نیاز داریم. این معیار در تحلیل‌های آماری مقدار احتمال خوانده می‌شود. مقدار احتمال، عددی است بین صفر و یک که مشخص می‌کند اندازه اثری که محقق در نظر گرفته یا هر اندازه اثری بزرگ‌تر از این مقدار چقدر احتمال دارد که تصادفی پذیرفته شود. هنگامی فرضیه صفر رد می‌شود و فرضیه مقابل آن (وجود ارتباط) پذیرفته می‌شود که مقدار احتمال عددی کمتر از α (که قبل از آزمون مقدار آن مشخص شده است مثلاً $0/05$ یا $0/01$) باشد. وقتی که نتایج از لحاظ آماری معنی‌دار نمی‌شود یعنی وقتی که مقدار احتمال بزرگ‌تر از α می‌شود تنها معنی آن این است که نتایج مشاهده‌شده در نمونه از آنچه می‌توانست تنها بر اساس تصادف (مقدار α) اتفاق بیافتد ضعیف است. برای مثال ممکن است پژوهشگری در تحقیق خود به این نتیجه برسد که افرادی که فشارخون بیش‌ازحد طبیعی دارند دو برابر افراد نرمال دچار سرطان پروستات می‌شوند. تنها به این علت که تعداد افراد مبتلا به سرطان در مطالعه نسبتاً کم بوده و به‌طور تصادفی در گروه دارای فشارخون بالا قرار گرفته‌اند و مقدار احتمال برابر $0/18$ شده است. این موضوع نشان می‌دهد که اگر فشارخون و سرطان پروستات هیچ رابطه‌ای با هم نداشته باشند نیز در 18% مواقع تنها از

روی شانس امکان مشاهده یک رابطه معنی‌دار از لحاظ آماری وجود دارد.



نمودار ۱-۱: در نمودار الف کاهش وزن به وسیله انجام دو نوع رژیم غذایی نمایش داده شده است. اعضای گروه دارای رژیم کم چربی همگی بین ۴ تا ۶ کیلوگرم کاهش وزن داشته‌اند در حالی که کاهش وزن در گروه رژیم کم کربوهیدرات از ۱- تا ۱ کیلوگرم در نوسان بوده است. بنابراین اندازه اثر این دو رژیم غذایی ۵ کیلوگرم است. به علت اینکه بین دو گروه همپوشانی وجود ندارد استنباط وجود برتری در استفاده از رژیم کم چربی نسبت به رژیم کم کربوهیدرات منطقی به نظر می‌رسد. در نمودار ب نیز کاهش وزن حاصل از دو رژیم غذایی نمایش داده شده است. در این نمودار همپوشانی گسترده‌ای در دو گروه مشاهده می‌شود؛ و برخی اعضا در هر کدام از گروه‌ها افزایش وزن داشته‌اند. در این حالت نیز اندازه اثر ۵ کیلوگرم و به نفع رژیم کم چربی است. باین‌حال نتایج مطالعه شواهد محکمی در بهتر بودن رژیم کم چربی نسبت به رژیم کم کربوهیدرات ارائه نمی‌کند.

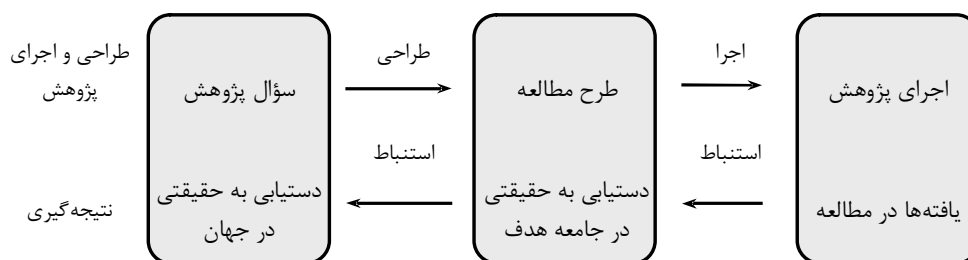
۱-۲-۱-۱ نحوه تحلیل آماری نتایج

تنظیم طرح تحقیق، تعیین حجم نمونه و توان آماری آزمون‌ها، تعیین نحوه انجام هر کدام از استنباط‌ها و روش‌های آماری مناسب برای رسیدن به اهداف موردنظر در هر پژوهش، اقداماتی هستند که محققان پیش از اجرای طرح باید تدابیر لازم برای آن‌ها را در نظر گرفته و در پروپوزال پژوهش به آن‌ها اشاره کنند.

۱-۲-۲ ماهیت پژوهش

هدف هر پژوهش بالینی، انجام استنباط‌هایی از نتایج مطالعه درباره واقعه‌ای در جهان است. برای دستیابی به این موضوع باید دو مرحله مجزا طراحی و اجرا شود. گام اول طراحی یک مطالعه بالینی شامل: انتخاب اعضای نمونه،

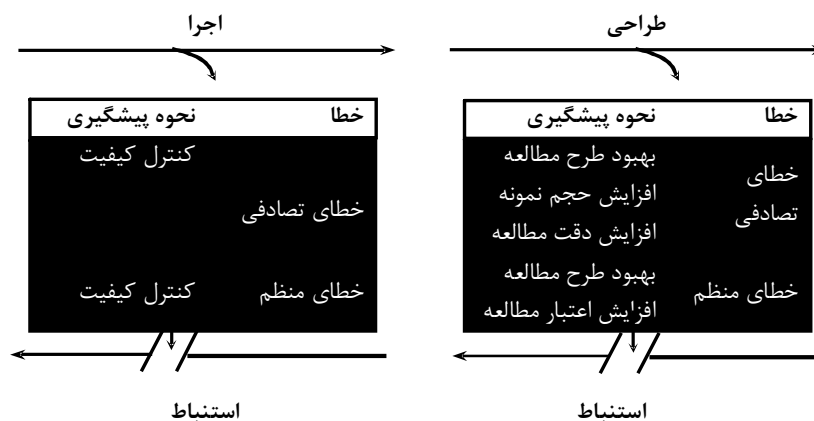
اندازه‌گیری متغیرهای موردنیاز، جمع‌آوری داده‌ها و تحلیل نتایج برای دستیابی به پاسخ مناسب برای سؤال پژوهش و تعمیم آن به جامعه هدف است. گام بعدی شامل نتیجه‌گیری صحیح از آنچه واقعاً در مطالعه رخ داده، است. به عبارت دیگر بررسی دقیق فرایند تحقیق برای دستیابی به نتایجی است که تا حد ممکن به واقعیت نزدیک باشند. این فرایند در نمودار ۲-۱ بیان شده است.



نمودار ۲-۱: فرایند طراحی و اجرای پژوهش و مرحله انجام استنباطها و نتیجه‌گیری کلی

۳-۱ منابع خطا در طراحی و اجرای مطالعه

مسئله هیچ تحقیقی خالی از خطا نیست. به همین علت، استنباطهای صورت گرفته بر این مبنا نیز هیچ‌گاه به طور کامل صحیح و معتبر نیستند. به طور کلی یکی از اهداف طراحی هر مطالعه حداکثر کردن اعتبار نتایج مطالعه است. بهترین راهبرد برای کاهش خطا در مطالعه، تأمل و دقت کافی در مرحله طراحی تحقیق است. پیشگیری از وقوع خطا عملی‌تر و به صرفه‌تر از هر اقدام دیگری است، هرچند باید توجه داشت که برخی از خطاها در مرحله تجزیه و تحلیل اطلاعات مهارنشده هستند. خطاهای تحقیق به دودسته خطاهای تصادفی و خطاهای منظم تقسیم‌بندی می‌شوند. نمودار ۳-۱ روش‌های کنترل این خطاها را نمایش می‌دهد.



نمودار ۳-۱: نحوه کنترل خطاهای تحقیق

خطاهای تصادفی، نتیجه اشتباهات حاصل شده به علت شانس هستند. منابع ناشناخته که باعث تحریف نتایج به دست آمده از نمونه در هر دو جهت (افزایش خطر یا کاهش آن) می شود اصطلاحاً خطاهای تصادفی خوانده می شود. به عنوان مثال اگر شیوع واقعی مصرف هورمون استروژن در خانم‌های ۵۰ تا ۶۹ ساله ۲۰٪ باشد، یک طرح نمونه‌گیری مناسب با حجم ۱۰۰ نفر از جامعه مربوطه می تواند دقیقاً شامل ۲۰ مورد با این ویژگی باشد هر چند امکان وجود ۱۸، ۱۹، ۲۰ یا ۲۲ مورد نیز کاملاً محتمل است. شانس یا خطای تصادفی بندرت حتی می تواند باعث تولید نتایج کاملاً متفاوت از واقعیت نظیر ۹ یا ۳۵ مورد شود. از میان روش‌های مختلف پیشنهاد شده برای کاهش اثر خطای تصادفی، مشهودترین و ساده‌ترین آن‌ها افزایش حجم نمونه است. استفاده از نمونه‌های بیشتر احتمال مشاهده نتایج اشتباه را با افزایش دقت مطالعه کاهش می دهد.

خطاهای منظم، نتیجه اشتباهات حاصل شده به علت تورش هستند. در خطای منظم نتایج به صورت جهت‌دار تحریف می شوند. به عنوان مثال انتخاب اعضای نمونه از خانم‌های مراجعه‌کننده به درمانگاه برای برآورد شیوع مصرف استروژن به احتمال زیاد باعث برآورد بیش از اندازه واقعی می شود. افزایش حجم نمونه هیچ تأثیری در کنترل یا بهبود خطای منظم ایفا نمی کند. تنها راه کاهش خطاهای منظم افزایش اعتبار تحقیق است. این موضوع معادل شناخت منابع تورش در مطالعه و حذف یا تعدیل اثر آن‌ها است.

۱-۳-۱ طراحی مطالعه

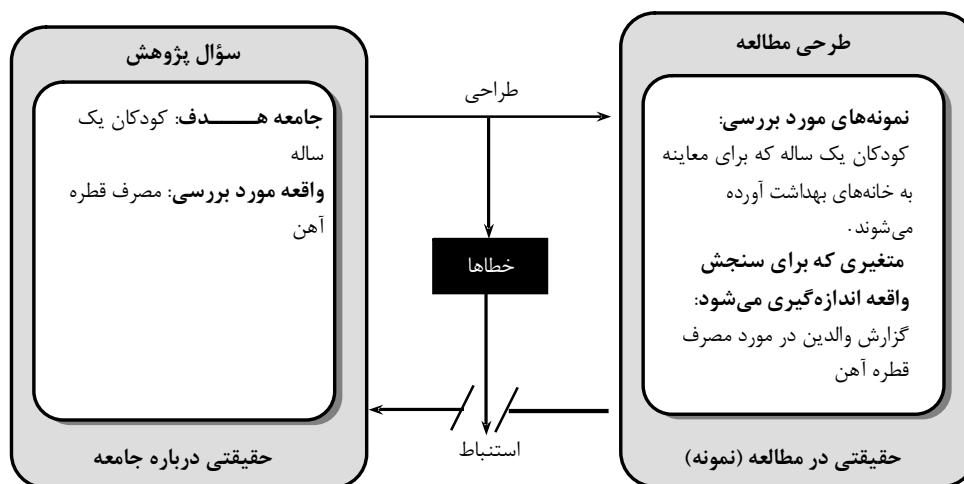
سؤال پژوهش همان‌طور که قبلاً راجع به آن توضیح داده شد، موضوعی است که پژوهشگران مایل به یافتن پاسخ‌هایی درباره آن هستند. برای ساده‌تر بودن مطالب، موضوع را با یک مثال ساده دنبال می کنیم. فرض کنید مایل هستیم بدانیم شیوع استفاده از قطره آهن در کودکان یک‌ساله چقدر است؟

جواب دادن به این سؤال با دقت صد درصد ناممکن است زیرا عملاً بررسی تمام کودکان زیر یک سال امکان پذیر نیست. به همین علت پژوهشگران باید به طراحی سؤال پژوهش دیگری بپردازند که قابل بررسی باشد. این سؤال می تواند این‌گونه مطرح شود: «شیوع استفاده از قطره آهن در بین کودکان یک‌ساله که جهت معاینه به خانه بهداشت آورده می شوند چقدر است؟»

یکی از مؤلفه‌های کلیدی در تبدیل سؤال فراگیر پژوهش به سؤالی قابل پژوهش، انتخاب نمونه‌ای است که جانشین مناسبی از جامعه باشند. برای مثال محدود شدن مطالعه به کودکانی که به خانه‌های بهداشت آورده می شوند تصمیمی قابل بحث است که تنها به خاطر موانع عملی موجود در انتخاب افراد تحت مطالعه اتخاذ شده است.

نکته بعدی انتخاب متغیرهایی است که با اندازه‌گیری آن‌ها قادر به بررسی وضعیت متغیر موردنظر باشیم. استفاده از یک پرسشنامه برای اندازه‌گیری شیوع مصرف قطره آهن یکی از روش‌های سریع و ارزان برای جمع‌آوری اطلاعات لازم است؛ اما استفاده از این روش به‌طور کامل دقیق نخواهد بود زیرا برخی والدین به کودکان قطره می‌دهند اما اکنون فراموش کرده‌اند و یا برخی ممکن است به این سؤال اشتباه پاسخ دهند زیرا واژه «قطره آهن» در پرسشنامه را متوجه نشده‌اند و یا نمی‌دانند چه قطره‌ای به کودکان خود داده‌اند.

به‌طور کلی تفاوت ایجاد شده بین سؤال فراگیر پژوهش و طرح اصلی مطالعه یکی از منابع خطاست که می‌تواند به یافته‌های غلط منجر شود. [۷-۹] نمودار ۱-۴ متداول بودن بروز این خطاها در طراحی هر مطالعه را نمایش می‌دهد.

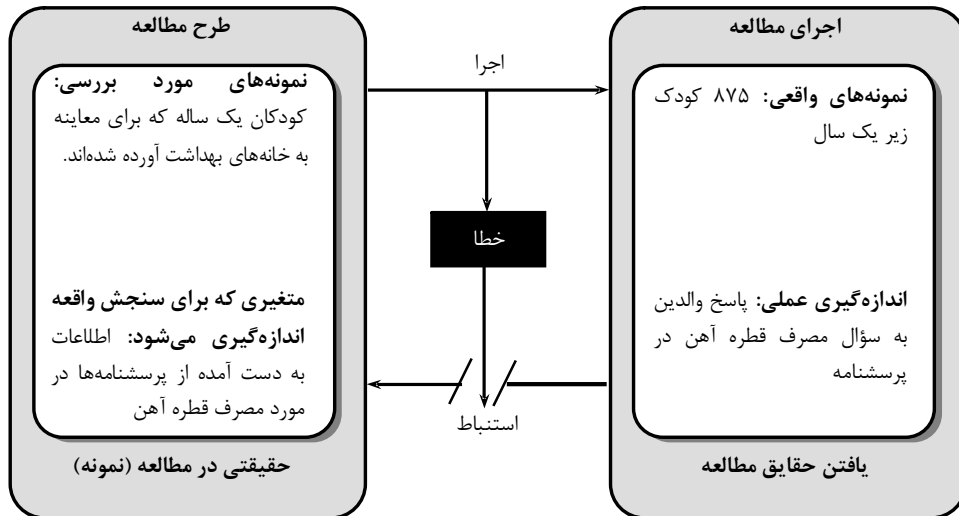


نمودار ۱-۴: منابع خطا در طراحی یک مطالعه فرضی

۱-۳-۲ اجرای مطالعه

اعضای واقعی تحت مطالعه در هر پژوهش، تقریباً همیشه با نمونه‌ای که واقعاً باید جمع‌آوری می‌شد متفاوت است. طرح نمونه‌گیری از کودکان یک‌ساله که به خانه‌های بهداشت آورده می‌شوند را در نظر بگیرید. ممکن است از ۱۰۰۰ نمونه که از قبل برای تعیین شیوع انتخاب شده‌اند تنها ۸۷۵ نفر قابل دسترسی باشند. اشخاصی که در دسترس نبوده‌اند و یا از همکاری سرباز زده‌اند ممکن است از جهاتی با بقیه متفاوت باشند و این موضوع باعث تغییر شیوع استفاده از قطره آهن در بین آنها شود. به‌جز این موضوع مواردی نظیر پیچیده بودن عبارات پرسشنامه و ناخوانا بودن جملات نیز عواملی هستند که می‌توانند باعث جواب ندادن به یک سؤال و یا پاسخ اشتباه به سؤالات شوند. این تفاوت‌های آشکار از طرح واقعی مطالعه و نحوه اجرای آن باعث تغییر و یا تحریف پاسخ‌ها می‌شوند. نمودار ۱-۵ نقش منابع دیگر خطا که در مرحله اجرا امکان بروز دارند را نمایش می‌دهند.

توضیح مفصل و مثال‌های بیشتری راجع به خطاهای تصادفی و منظم، نحوه افزایش دقت و اعتبار مطالعه و روش‌های ارزیابی دقت و اعتبار نتایج تحقیق، مباحثی است که در فصل چهارم بررسی می‌شوند. **ضمیمه الف**، مطالب بیشتر راجع به تنظیم پروپوزال طرح‌های پژوهشی را بیان می‌کند.

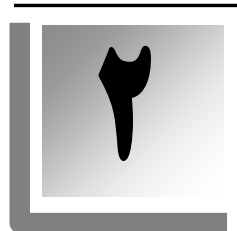


نمودار ۱-۵: منابع خطا در اجرای یک مطالعه فرضی

منابع

1. Greenhalgh T. How to read a paper. London, BMJ Publishing group, 1997.
2. Hall GM. How to write a paper: (2nd ed.). London, BMJ Publishing group, 1998.
3. Kahn TS. The structure of scientific revolutions. Chicago, University of Chicago Press, 1962.
4. Design of the woman's health initiative clinical trial and observational study. The Woman's Health Initiative Study Group. *Controlled Clinical Trials* 1998, 19:61-109.
5. Cochran WG. *Sampling Techniques: (3rd. ed.)*. New York, John Wiley 1977.
۶. عمیدی ع. نظریه نمونه‌گیری و کاربردهای آن، جلد اول و دوم، تهران: مرکز نشر دانشگاهی، ۱۳۷۸.
7. Hulley SB, et al. *Designing clinical research: An epidemiologic approach (2nd ed.)*. Philadelphia Lippincott Williams & Wilkins, 2001.
۸. سوری ح. بررسی و نقد مقالات اپیدمیولوژی، مجله دانشکده پزشکی دانشگاه علوم پزشکی گیلان، ۱۳۸۱؛ دوره ۱۱، بهار، شماره ۴۱: صفحات ۶۹-۶۴.
۹. گروه مترجمان دانشگاه علوم پزشکی شهید بهشتی، به همت و ویراستاری سوری ح. مدیریت و اداره تحقیق، انتشارات دانشگاه علوم پزشکی شهید بهشتی. تهران: نشر صبا، ۱۳۸۶.
۱۰. چانگ چاو ش و همکاران. حجم نمونه در پژوهش‌های بالینی. ترجمه: علوی مجد ح. دانشگاه علوم پزشکی شهید بهشتی. زمستان ۱۳۸۵.

راهبردهای طراحی مطالعات در تحقیقات علوم پزشکی



به‌طور کلی مطالعات در علوم پزشکی را می‌توان به دو رده عمده مشاهده‌ای و مداخله‌ای تقسیم‌بندی کرد که هر کدام از این دو رده بر اساس صفاتی نظیر واحد جمع‌آوری داده‌ها، جهت‌گیری زمانی، نحوه جمع‌آوری اطلاعات یا نمونه‌ها، نقش محققان در تماس افراد تحت مطالعه با مواجهه‌های موردبررسی به زیرگروه‌های کوچک‌تری تقسیم‌بندی می‌شوند. این فصل به بررسی دقیق این تقسیم‌بندی‌ها و راهکارهای ممکن در طراحی مطالعات پزشکی اختصاص دارد.

۲-۱ مسائل اساسی در انتخاب روش مطالعه

انتخاب روش مطالعه و طراحی آن به عوامل مختلفی بستگی دارد. معمولاً محققان علاقه‌مند به انتخاب روش مطالعه‌ای معتبر با حداقل خطا در تحقیق هستند، اما محدودیت‌ها و موانع عملی باعث می‌شود مطالعه‌ای با وسعت کمتر، اهداف محدودتر و توان علمی کمتر طراحی و اجرا شود. در انتخاب روش مطالعه در درجه اول می‌بایست به آنچه هدف نهایی مطالعه نامیده می‌شود اشراف کامل داشت و به مزایا و محدودیت‌های هر یک از انواع طرح‌های مطالعه، بضاعت و امکانات، همین‌طور قابلیت انجام‌پذیری مطالعه توجه کرد. برخی از عوامل مؤثر در انتخاب روش مطالعه به‌قرار زیر است:

- ♦ هدف مطالعه
- ♦ اندازه‌گیری رخدادها
- ♦ تعیین ارتباط بین عوامل خطرزا و بیماری
- ♦ مطالعه درباره علل ایجادکننده بیماری
- ♦ قابلیت انجام مطالعه
- ♦ قابلیت دسترسی به افراد مورد مطالعه
- ♦ ابزار و مواد اولیه
- ♦ کارکنان، زمان و بودجه تحقیق
- ♦ اعتبار مطالعه
- ♦ منابع خطا

۲-۲ انواع مطالعات پزشکی

به‌طور کلی مطالعات پزشکی به دودسته قابل تقسیم‌بندی هستند.

دسته اول مطالعات مبتنی بر مشاهدات و دسته دوم مطالعاتی هستند که با انجام مداخلاتی بر روی افراد تحت مطالعه صورت می‌گیرند. بر این اساس دو گروه اصلی مطالعات پزشکی عبارت‌اند از:

- ♦ مطالعات مشاهده‌ای
- ♦ مطالعات مداخله‌ای

از جهت دیگر مطالعات را می‌توان به دو رده کلی مطالعه‌های توصیفی و مطالعه‌های تحلیلی یا مطالعات کمی و کیفی نیز دسته‌بندی کرد. بیشتر مطالعات علوم پزشکی از نوع کمی هستند. مطالعات کیفی در علوم اجتماعی ریشه دارند. گزارش یک مطالعه کیفی مملو از متن است تا بیان اعداد و محقق مطالعه کیفی می‌پذیرد که کل مطالعه در یک فضای با تورش ذهنی می‌تواند انجام پذیرد. از روش‌های جمع‌آوری داده‌ها در مطالعات کیفی می‌توان از روش دلفی و روش بحث متمرکز گروهی نام برد.

مطالعات توصیفی بر توصیف توزیع بیماری یا رخداد موردبررسی متمرکز است. به‌عنوان مثال توجه به موارد ذیل که در چه جمعیت‌ها یا زیرگروه‌هایی از جمعیت بیماری ایجاد می‌شود و یا نمی‌شود، در چه محل‌های جغرافیایی بیشترین یا کمترین میزان شیوع بیماری وجود دارد و یا اینکه فراوانی بیماری در طول زمان چگونه تغییر می‌یابد، پرسش‌هایی هستند که اطلاعات آن‌ها با استفاده از مطالعات توصیفی فراهم می‌شود. اطلاعات در مورد هر یک از این مشخصات می‌تواند سرنخ‌هایی برای تنظیم فرضیه‌ای که با دانش موجود در مورد بروز بیماری سازگار باشد به دست دهد.

مطالعات تحلیلی بر عوامل تعیین‌کننده یک بیماری با فرضیه‌ای که در مطالعات توصیفی تنظیم شده است متمرکز است. هدف این مطالعات امکان تشخیص عوامل خطر بیماری است. دسته دیگر از مطالعات تحلیلی، مطالعات مداخله‌ای هستند، در این مطالعات دو یا چند روش درمانی، رژیم غذایی، شیوه مدیریتی و امثال این‌ها که تحت عنوان مداخله از آن‌ها نام برده می‌شود (مداخله‌ها معمولاً تصادفی به زیرگروه‌های تحت مطالعه منتسب می‌شوند)، با یکدیگر مقایسه می‌شوند.

طرح هر مطالعه توصیفی یا تحلیلی نقاط قوت و محدودیت‌های منحصر به خود را دارد. از آنجاکه این موضوع نیاز به شناخت کلی از انواع مطالعه‌ها و ارتباط‌های داخلی آن‌ها برای فراهم کردن اساس نظری و علمی دارد، در این فصل نگاه مختصری بر انواع طرح‌های مطالعاتی خواهیم داشت. جدول (۱-۲) شمای کلی از دسته‌بندی طرح‌های مطالعات پزشکی را نمایش می‌دهد. مباحث بیشتر در این زمینه در فصل سوم ارائه می‌شود.

جدول ۲-۱: دسته‌بندی طرح‌های مطالعات در پزشکی

(Observational studies)	♦ مطالعات مشاهده‌ای
(Descriptive studies)	الف: مطالعات توصیفی
(Case study)	- مطالعه مورد
(Case-series study)	- مرور موارد
(Normative study)	- مطالعه هنجاریابی
(Ecologic study)	- مطالعه همبستگی (بوم‌شناختی)*
(Longitudinal study)	- مطالعه طولی
(Historical study)	- مطالعه تاریخی
(KAP study)	- مطالعه KAP**
(Research based on existing data)	ب: مطالعات مبتنی بر داده‌های موجود
(Secondary data analysis)	- تحلیل مجدد داده‌ها
(Ancillary study)	- مطالعه فرعی
(Systematic review and meta analysis)	- مرور منظم و متاآنالیز
(Cross-sectional studies)	ج: مطالعات مقطعی (**)
(Case- control studies)	د: مطالعات مورد-شاهدی
(Cohort studies)	ه: مطالعات هم‌گروهی
(Prospective cohort study)	- مطالعه هم‌گروهی آینده‌نگر
(Retrospective cohort study)	- مطالعه هم‌گروهی تاریخی
(Interventional studies)	♦ مطالعات مداخله‌ای
(Experimental studies)	الف) مطالعات تجربی
(Parallel or concurrent controls experiment)	- مطالعه تجربی با شاهد‌های هم‌زمان یا موازی
(Cross-over study)	- مطالعه متقاطع
(Field study)	مطالعه مبتنی بر روش‌های میدانی
(Quasi- Experimental study)	ب) مطالعات نیمه تجربی
(External controls experiments)	- مطالعه با شاهد‌های خارجی
(Sequential controls experiment)	- مطالعه تجربی با شاهد‌های متوالی
(Interventional studies with no controls)	ج) مطالعات مداخله‌ای بدون گروه شاهد

* در برخی متون مطالعه همبستگی جزء مطالعات تحلیلی طبقه‌بندی شده است.

** مطالعه KAP ممکن است به تعبیری به لحاظ ساختاری جزء انواع مطالعات طبقه‌بندی نشوند.

۲-۲-۱ مطالعات توصیفی

مطالعه توصیفی همان‌طور که از نام آن برمی‌آید به توصیف خصوصیات و مشخصات عمومی توزیع یک بیماری در ارتباط با شخص، مکان و زمان می‌پردازد. مطالعات توصیفی از لحاظ مقایسه فراوانی‌های بیماری بین گروه‌های مختلف به سه دسته مطالعات انفرادی و مطالعات همبستگی و مقطعی تقسیم‌بندی می‌شوند. برخی از انواع مطالعات توصیفی عبارت‌اند از:

- مطالعات همبستگی

مطالعه همبستگی، از این جهت با سایر مطالعات متفاوت است که بجای اندازه‌گیری ویژگی‌های موردنظر برای تک‌تک افراد جامعه، مقادیر این ویژگی‌ها برای گروه‌های بزرگی از جوامع، اندازه‌گیری و گزارش می‌شود. این گروه‌ها ممکن است کلاس‌های یک مدرسه و یا مدارس یک شهر، کارخانه‌ها، شهرها و کشورها، باشند. تنها شرط انجام این مطالعات در دسترس بودن چنین اطلاعاتی از جمعیت مورد مطالعه است. شیوع بیماری‌ها و مرگ‌ومیر، به‌طور گسترده‌ای در این مطالعات برای معین کردن وقوع بیماری در گروه‌ها به کار گرفته شده و عوامل در معرض، با یکسری شاخص‌های کلی اندازه‌گیری و بیان می‌شوند. برای مثال فراهم بودن اطلاعات مصرف سالیانه گوشت قرمز و تعداد مرگ‌ومیرهای ثبت‌شده از سرطان روده بزرگ، در شهر یا کشورهای مختلف، انجام یک مطالعه همبستگی را برای بررسی اولیه وجود ارتباط بین میزان مصرف گوشت قرمز و ابتلا به سرطان روده بزرگ عملی می‌کند. از آنجاکه در مطالعات همبستگی، اندازه‌ها بر اساس مقدار میانگین برای هر شخص اندازه‌گیری می‌شود، اندازه ارتباط محاسبه‌شده بین مواجهه و بیماری، قابل انتساب به هر فرد جامعه نیست.

- مطالعات هنجاریابی (هنجاری)

در این مطالعات خصوصیات یا صفات کمی در یک جمعیت تعیین می‌گردد و معلوم می‌شود که این خصوصیات یا اندازه‌ها تا چه حد با یکدیگر تفاوت دارند. انجام مطالعات هنجاریابی در مورد خصوصیات یا اندازه‌هایی که به دلایل مختلف (شیوه زندگی، خصوصیات جمعیتی، ویژگی‌های اقلیمی و غیره) ممکن است از جامعه‌ای به جامعه دیگر متفاوت باشد در هر جامعه ضروری است. برای مثال اندازه‌های نرمال وزن کودکان در موقع تولد را به دلایل تفاوت‌های مختلف مرتبط با این صفت در جمعیت‌های مختلف در هر جامعه باید برای همان جمعیت تعیین کرد و نمی‌توان الگوی نرمال یکسانی را برای جمعیت‌های مختلف در نظر گرفت.

- مطالعات طولی

این مطالعات برای تعیین الگوی یک رخداد در طول زمان استفاده می‌شود. این نوع مطالعه را ممکن است جزء مطالعات تحلیلی طبقه‌بندی کرد اما زمانی که هدف تحقیق تعیین ارتباط سببی نیست مطالعه طولی یک مطالعه توصیفی محسوب می‌شود.

- مطالعات تاریخی

مطالعه تاریخی برای توصیف و تشریح مسئله‌ای است که در گذشته رخ داده است. این بررسی کمک می‌کند تا وضعیت مسئله موردپژوهش در حال حاضر مشخص‌تر و آشکارتر شود. داده‌های مربوط به این مطالعات از طریق

بررسی اسناد، مدارک و یا مصاحبه با افراد مطلع به دست می‌آید.

– مطالعات KAP

این مطالعات برای بررسی دانش یا معلومات (Knowledge)، نگرش (Attitude) و رفتار یا عملکرد (Practice) افراد مورد مطالعه انجام می‌شوند. این نوع بررسی‌ها بیشتر در آموزش پزشکی و علوم بهداشتی کاربرد دارد. برای مثال سنجش میزان معلومات یا اطلاعات دانش‌آموزان در زمینه پیشگیری از ایدز، دیدگاه آنان در مورد این بیماری و تمهیداتی که برای محافظت از خود در مقابل این بیماری انجام می‌دهند یک مطالعه KAP محسوب می‌شود. ممکن است مطالعه KAP به لحاظ ساختاری تنها موضوع پژوهش به حساب آید و جزء انواع مطالعات طبقه‌بندی نشود.

– مطالعه با استفاده از داده‌های موجود

بسیاری از سؤالات پژوهشی به‌طور سریع، کارا و اقتصادی با استفاده از اطلاعاتی که قبلاً جمع‌آوری شده‌اند قابل پاسخگویی هستند. سه روش عمومی برای طراحی این نوع مطالعات موجود است. تحلیل مجدد داده‌ها روشی است که در آن محققان با استفاده از اطلاعات جمع‌آوری شده در یک مطالعه، که اساساً برای پاسخگویی به یک سؤال پژوهشی مشخص طراحی و اجرا شده، به بررسی و آزمون یک سؤال پژوهشی متفاوت می‌پردازند. مطالعه فرعی، طراحی است که با افزودن برخی سؤالات به پرسشنامه و یا اندازه‌گیری متغیرهای اضافه در فرآیند جمع‌آوری اطلاعات از افراد تحت مطالعه، محققان قادر به پاسخگویی به یک سؤال پژوهشی مجزا می‌شوند. بازبینی منظم مطالعه‌ای مرکب از تمام مطالعات قبلی مرتبط با سؤال پژوهش است. در این روش برآوردی از عوامل مؤثر تلخیص شده از تمام مطالعات واجد شرایط محاسبه می‌شود. استفاده خلاقانه از روش‌های مبتنی بر داده‌های موجود، در مواقعی که با محدودیت زمان و منابع موردنیاز روبرو هستیم، یکی از راه‌های مؤثر برای پاسخگویی به برخی سؤالات مهم پژوهشی است.

– متآنالیز یا فراتحلیل

در علوم پزشکی درباره یک موضوع، مطالعات متعددی در نقاط مختلف دنیا انجام می‌شود که اغلب نتایج متناقضی به همراه دارد. برای رسیدن به نتیجه واحد یکی از روش‌ها، ادغام نتایج این مطالعات و تحلیل نهایی آن‌هاست که فراتحلیل نامیده می‌شود. اصطلاح متآنالیز را اولین بار گلاس، رئیس انجمن تحقیقات آموزشی آمریکا، در سال ۱۹۷۶ بکار برد. گلاس گفت: فراتحلیل عبارت است از ادغام یافته‌های مجموعه بزرگی از مطالعات تکی منتشر شده و تحلیل آماری کلی آن‌ها. فراتحلیل، تحلیل تحلیل‌ها در مطالعات مختلف است. روش‌های فراتحلیل، مجموعه‌ای از تکنیک‌ها برای ترکیب آماری نتایج چند مطالعه مستقل به‌منظور یافتن پاسخ کلی است. هدف از این روش‌ها، فراهم کردن آزمونی با توان بالاتر از توان تک‌تک مطالعات مستقل موردبررسی است. در یک فراتحلیل، محقق برای ادغام یافته‌های یک مجموعه از مطالعات روش‌های آماری را بکار می‌برد و نتایج را با استفاده از برآوردهای **اندازه اثر** عددی توصیف می‌کند.

۲-۲-۲ مطالعات تحلیلی

اکثر طرح‌های مطالعاتی، به‌طور تلویحی برای مقایسه عوامل خطر و شناخت وضعیت بیماری طراحی می‌شوند. برای مثال در گزارش یک مورد، هنگامی که تنها یک‌شکل خاص از بیماری مشاهده می‌شود فرضیه‌ای بر اساس

مقایسه ضمنی^۶ با تجربه متداول تنظیم می‌شود. در مطالعات تحلیلی، مقایسه به‌طور صریح صورت می‌گیرد زیرا محقق، گروه‌هایی از افراد را برای هدف خاص به‌طور منظم جمع‌آوری می‌کند تا تفاوت خطر بیماری برای افرادی که مواجهه ویژه‌ای دارند با افرادی که آن مواجهه را ندارند را مشخص کند. راه‌های مختلفی برای طراحی مطالعات تحلیلی وجود دارد که بر اساس شرایط و امکانات موجود نوع خاصی از آنها به خدمت گرفته می‌شود. دودسته کلی از مطالعات تحلیلی مطالعات مشاهده‌ای و مطالعات مداخله‌ای هستند.

تفاوت اصلی بین این دو راهبرد در نقشی است که توسط محقق ایفا می‌شود. در مطالعات مشاهده‌ای، محقق به‌طور ساده جریان اتفاقات طبیعی را با توجه به اینکه چه کسی در معرض مواجهه قرار دارد و یا اینکه چه کسانی از افراد سالم به بیماری مبتلا شده‌اند را مشاهده می‌کند. در مطالعات مداخله‌ای، پژوهشگران مواجهه گروه‌های تحت مطالعه را انتخاب می‌کنند و سپس افراد را از نظر بروز بیماری یا پیامد پیگیری می‌کنند. مشکل اساسی در مطالعات مشاهده‌ای این است که معمولاً گروه‌های مشاهده‌شده علاوه بر عامل خاص تحت مطالعه، از نظر پاره‌ای از خصوصیات دیگر هم با یکدیگر متفاوت هستند. فرضاً گروه‌های شغلی مختلف نه تنها از نظر مواجهه با خطرهای شغلی با یکدیگر اختلاف دارند بلکه از لحاظ رویدادهای گذشته زندگی نیز متفاوت‌اند، زیرا افراد تقریباً مشابهی در شغل‌های همسان استخدام می‌شوند. این تشابه می‌تواند به علت علایق شخصی، تحصیلات و یا مناسب بودن برای شغل موردنظر، حاصل شده باشد. به دلیل این عوامل مخدوش‌کننده که اغلب اندازه‌گیری آنها با دشواری‌های زیادی همراه هستند، نمایش نقش عامل تحت بررسی مشکل‌تر می‌شود. هرچند مطالعات تجربی به‌مراتب بهتر از مطالعات مشاهده‌ای می‌تواند رابطه سببی یک عامل با یک بیماری را نشان دهند اما بیشتر معلومات ما درباره بیماری‌ها از مطالعات مشاهده‌ای سرچشمه می‌گیرد.

– مطالعات مشاهده‌ای

دو نوع اصلی بررسی‌های مشاهده‌ای تحلیلی مطالعات مورد-شاهدی و مطالعات هم‌گروهی هستند. از لحاظ نظری این امکان وجود دارد که یک فرضیه با استفاده از هر کدام از این دو روش مورد بررسی قرار گیرد. هر کدام از این دو نوع مطالعه دارای مزایا و معایبی هستند که در فصل سوم درباره آنها بحث خواهد شد. به‌طور کلی تصمیم‌گیری برای استفاده از یک راهبرد به عواملی چون خصوصیات عامل خطر و بیماری در مرحله کنونی، میزان آگاهی علمی و ملاحظات منطقی نظیر منابع و زمان بستگی دارد. در مطالعات مورد-شاهدی، وجود یا فقدان بیماری موردنظر مشخص بوده و پژوهشگر با بررسی سابقه بیمار به کشف علت یا عامل‌های خطر احتمالی مطرح شده در مطالعه می‌پردازد. موردها، افراد بیمار یا مبتلا به پیامد موردنظر و شاهد‌ها افراد سالم یا فاقد بیماری موردنظر هستند. پژوهشگران خصوصیت یا عامل خطری را که تنها در بیماران و نه در افراد سالم وجود دارد شناسایی کرده و شرح حال وقایع قبلی زندگی بیماران و افراد شاهد را بررسی می‌کنند.

در مطالعات هم‌گروهی پژوهشگران در شروع مطالعه، افراد را انتخاب کرده و سپس وجود عامل خطر یا میزان تماس با آن را تعیین می‌کنند. آنها طی دوره زمانی مشخص، افراد در معرض مواجهه یا بدون مواجهه را با در نظر گرفتن عامل خطر و جهت‌گیری زمانی رو به آینده مطالعه می‌کنند.

نوع دیگری از مطالعات هم‌گروهی، مطالعات هم‌گروه تاریخی هستند. این مطالعات با استفاده از اطلاعات

۱. Implicit

جمع‌آوری شده از گذشته (موجود در پرونده‌ها) انجام می‌شود. با این وجود، جهت‌گیری زمانی مطالعه، همچنان روبه آینده است. یعنی شروع مطالعه مبتنی بر اطلاعاتی در مورد عامل‌های خطر بوده و پیامدهای عوامل خطر در آینده تعیین می‌شود.

– مطالعات مداخله‌ای

تجربه‌ها محکم‌ترین منابع برای آزمودن هر فرضیه‌ای هستند؛ اما به‌کارگیری آن‌ها در جوامع بشری به‌ندرت امکان‌پذیر است. در پزشکی آن دسته از مطالعات تجربی که بر روی انسان‌ها انجام می‌گیرد (تصادفی‌سازی در مورد افراد بیمار انجام می‌شود)، کارآزمایی‌های بالینی نامیده می‌شوند. مداخلات به دو گروه شاهددار و بدون شاهد (تنها گروه درمان) تقسیم‌بندی می‌شوند. کارآزمایی‌های شاهددار از نظر اختصاص مداخله (درمان)، به گروه‌های انواع کارآزمایی‌های تصادفی شده، کارآزمایی‌های تصادفی نشده و حالتی بینابین این دو تقسیم می‌شوند.

آن دسته از مطالعات تجربی که اختصاص مداخله در آن‌ها به تصادف صورت می‌گیرد در صورت بهره بردن از حجم نمونه کافی و لحاظ نمودن برخی راهکارها (نظیر کور کردن مطالعه، کنترل شرایط محیطی) محکم‌ترین شواهد را در بررسی مطالعات بین عوامل خطر و بیماری ارائه می‌دهند. این امر مربوط به توان منحصر به فرد تصادفی کردن به‌عنوان ابزاری برای تعیین و کنترل مداخله است. وقتی شرکت‌کنندگان برای دریافت یک دارو یا دارونما، قرار گرفتن در معرض یک مواجهه یا عدم آن، به تصادف انتخاب شوند به‌طور متوسط و در درازمدت همه عوامل دیگری که ممکن است روی خطر بیماری اثر بگذارد مهار می‌شوند. در عین حال اگر چنین متغیرهایی برای محقق از قبل شناخته شده باشد هنگام طراحی یا تجزیه و تحلیل قابل مهار شدن هستند. خوانندگان محترم برای مطالعه بیشتر می‌توانند به منابع [۱] تا [۱۲] مراجعه نمایند.

منابع

1. Armenian HK, Shapiro S. Epidemiology and health services. New York, Oxford University Press, 1998.
2. Albert DA, Munson R, Resnik MD. Reasoning in medicine. Baltimore, John Hopkins University Press, 1998.
3. Thibault. GE. The appropriate degree of diagnostic certainty. N Engl J Med. 1994; 331: 1216-20.
4. Armenian HK, Gordis L. Future perspectives. Epidemiol Rev. 1994; 16:163-40.
5. Greenland S, Watson E, Neutra RR. The case-control method in medical care evaluation. Med. Care. 1981; 19; 872-8.
6. Selby JV. Case-control evaluation of treatment and program efficacy. Epidemiol Rev. 1994; 16:90-101.
7. Canadian Task Force on the Periodic Health Examination. Can Med Assoc J 1979; 121:1193-1254.
8. Benenson AS. Control of communicable diseases in man, (32nd ed.). Washington D.C, American Public Health Association, 1995.
9. Blair AJ, Brug J, Foran, et al. Guidelines for application of meta-analysis in environmental epidemiology. Regul Toxicol Pharmacol. 1996; 22:189-97.
10. Friedenreich C. Methods for pooled analysis of epidemiologic studies. Epidemiology 1993; 4:295-302.
۱۱. سوری ح، اساسی ن. روش‌های مطالعه توصیفی در: روش‌شناسی پژوهش‌های کاربردی در علوم پزشکی، گروه مؤلفین. معاونت تحقیقات و فناوری وزارت بهداشت، درمان و آموزش پزشکی، انتشارات علوم پزشکی تهران ۱۳۸۳.
۱۲. سوری ح. کلیات و اصول اپیدمیولوژی، در: کلیات بهداشت عمومی (۳ جلد)، تألیف حاتمی ح و همکاران (گروه مؤلفان)، انتشارات ارجمند ۱۳۸۵.

روش‌های طراحی مطالعات علوم پزشکی



در فصل قبل راجع به راهکارهای متفاوت در طراحی مطالعات علوم پزشکی با توجه به امکانات و محدودیت‌ها و همچنین نحوه تقسیم‌بندی این مطالعات بحث و بررسی شد. در این فصل جزییات مربوط به نحوه طراحی و اهداف هر کدام از این مطالعات بررسی می‌شود.

۳-۱ روش‌های مطالعات توصیفی

مطالعات توصیفی، اطلاعات مربوط به فراوانی رخداد یک وضعیت خاص و یا الگوهای رخداد را بر اساس عوامل مربوط به شخص، زمان و مکان ارائه می‌دهند. آمارهای جمع‌آوری شده متداول، از قبیل داده‌های مربوط به گواهی‌های فوت، پرونده‌های پزشکی، برنامه‌های پایش و مراقبت از بیماری‌ها و داده‌های حاصل از نمونه‌گیری و یا سرشماری می‌توانند در مطالعات توصیفی استفاده شوند.

مطالعات توصیفی جزء مطالعات غیرتجربی طبقه‌بندی می‌شوند. در این مطالعات مشخصات عمومی توزیع یک بیماری از قبیل؛ خصوصیات فردی (مانند سن، جنس، شغل، نژاد، وضع تأهل)، ویژگی‌های مرتبط با شیوه زندگی (نظیر عادات غذایی، استفاده از داروهای مختلف)، خصوصیات اقلیمی، توزیع جغرافیایی عوامل و دوره‌های زمانی معین، با توجه به شاخص‌های سلامتی، بیماری یا مرگ در افراد مورد مطالعه، توصیف می‌شوند.

۳-۲-۱ گزارش مورد یا مطالعه مورد

گزارش یا مطالعه مورد، توجه و بررسی دقیق یک مورد غیرمعمول است که چگونگی و مکانیسم احتمالی رخداد موردنظر را بررسی می‌کند. این مطالعات ارزان، راحت و معمولاً بدون مشکلات ناشی از ملاحظات اخلاقی در طراحی مطالعه هستند. نتیجه‌گیری قطعی از علیت رخداد در مطالعه مورد امکان‌پذیر نیست و هدف اصلی این مطالعات توصیف مورد نادر بیماری است.

۳-۲-۲ مرور موارد

مرور موارد، گسترش گزارش مورد و یا گزارش چند مورد بجای یک مورد است. در مرور موارد، معمولاً خصوصیات تعدادی از بیماران در یک مقطع یا دوره زمانی مطالعه می‌شود و توزیع رخدادها در زیرگروه‌های جمعیتی مشخص می‌شود.

۳-۲-۳ مطالعات همبستگی (بوم‌شناختی)

در مطالعات همبستگی بجای افراد، تأکید اصلی بر مقایسه گروه‌ها قرار دارد. دلیل انجام این نوع مطالعات فقدان اطلاعات، در سطح تک‌تک افراد مورد مطالعه است. امکان دارد حداقل برای توزیع توأم دو متغیر و یا حتی تمام متغیرها، اطلاعات دقیق و جامعی موجود نباشد. به همین علت، در بسیاری از زمینه‌ها، در علوم پزشکی انجام این مطالعات موجه به نظر می‌رسد. باین‌حال، تمایز بین اندازه‌های فردی و اندازه‌های کلان (اکولوژیک) و نحوه انجام استنباط‌ها در این زمینه، امر بسیار دشواری است.

مطالعات همبستگی بر اساس محتوای اندازه‌گیری‌ها به سه گروه تقسیم می‌شوند:

۱- **اندازه‌های خلاصه:** در این حالت تنها خلاصه‌ای (مثلاً میانگین یا نسبت) از مشاهدات، که مرتبط با افراد تحت مطالعه است، در دست است. (مثلاً نسبت افراد سیگاری و میانه درآمد خانوار)

۲- **اندازه‌های محیطی:** اندازه‌گیری مشخصه‌های فیزیکی در محلی که گروه تحت مطالعه در آن زندگی یا کار می‌کنند، مانند میزان آلودگی هوا و ساعت تابش خورشید، اندازه‌های محیطی را تشکیل می‌دهند. نکته قابل‌توجه در این نوع اندازه‌ها این است که برای هر کدام از افراد گروه مطالعه، اندازه‌های مجزا، قابل‌اندازه‌گیری است که این مواجهه‌های انفرادی (و یا دوز مصرف)، معمولاً در زیرگروه‌های مورد بررسی متفاوت هستند. این امکان وجود دارد که برخی افراد کلاً بدون مواجهه باشند.

۳- **اندازه‌های کلان:** این اندازه‌ها، مربوط به ویژگی‌هایی از گروه‌ها، سازمان‌ها یا مکان‌ها هستند که هیچ مقدار مشخصی برای مقدار این ویژگی‌ها در سطح فردی قابل‌محاسبه نیست؛ مانند تراکم جمعیت، وجود یکسری قوانین مشخص، نوع نظام‌های خدمات درمانی و نحوه پوشش بیمه.

در تحلیل هر مطالعه، یکی از نکات مهم واحد جمع‌آوری داده‌ها است. معمولاً در تحلیل اطلاعات، کوچک‌ترین سطح مشترکی که داده‌ها به‌وسیله آن قابل‌بیان هستند، به‌عنوان واحد و مبنای تحلیل به‌کاربرده می‌شود. هنگامی که واحد اندازه‌گیری، تک‌تک افراد مورد مطالعه باشد به ازای هر عضو مطالعه یک اندازه‌گیری وجود دارد؛ اما هنگامی که با اندازه‌های اکولوژیک مواجه هستیم، هیچ اندازه مجزایی برای تک‌تک افراد جامعه در دست نیست. مثلاً هنگامی که موقعیت اجتماعی افراد بر اساس محل سکونت آن‌ها مشخص گردد، تمام افرادی که در یک ناحیه خاص زندگی می‌کنند دارای رده اجتماعی مشترک خواهند بود. گاهی اوقات که تمام متغیرهای مورد بررسی (مواجهه، بیماری و متغیرهای مخدوش‌کننده) از نوع اکولوژیک هستند مطالعه را مطالعه همبستگی کامل می‌خوانند. در این هنگام واحد تحلیل نتایج نیز گروه‌ها هستند. برای مثال، این واحدها می‌توانند نواحی مختلف یک استان یا کشور، محیط‌های کاری، مدارس و یا فواصل زمانی باشند. در این صورت از توزیع توأم متغیرها در سطح انفرادی، اطلاعی در دسترس نیست (برای مثال فراوانی‌های موردهای در معرض، موردهایی که در معرض نیستند، شاهد‌های در معرض و شاهد‌هایی که در معرض نیستند، قابل‌اندازه‌گیری نیست). در برخی دیگر از مطالعات، دستیابی به اطلاعات اضافی در مورد برخی از توزیع‌های توأم امکان‌پذیر است، اما هنوز اطلاع کاملی از تمام توزیع‌های توأم در دست نیست. چنین مطالعاتی، مطالعات نیمه همبستگی خوانده می‌شوند. برای مثال در یک مطالعه همبستگی برای بررسی بروز سرطان در استان‌های مختلف توزیع توأم سن (یک متغیر

همراه) و بروز یا عدم بروز بیماری در هر استان، از روی دفاتر ثبت سرطان‌ها در بیمارستان‌ها و یا گواهی‌های فوت قابل محاسبه است [۱] و [۲].

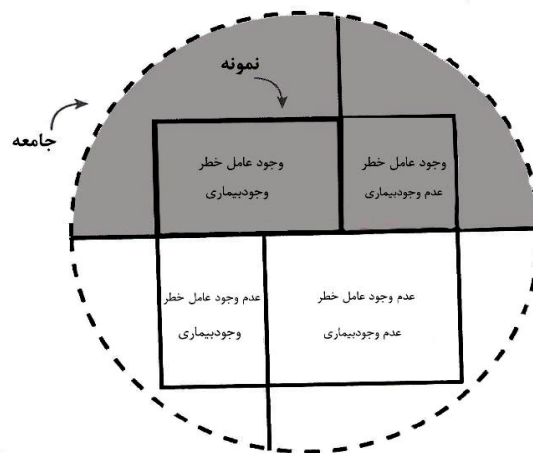
مزایا و معایب مطالعات همبستگی: دلایل مختلفی وجود دارد که علیرغم محدودیت‌های ساختاری و متدولوژیک مطالعات همبستگی، استفاده از آن‌ها جذاب و منطقی باشد.

- ♦ **سهولت انجام و کم‌هزینه بودن این مطالعات:** اگر منابع ثانویه برای جمع‌آوری داده‌های موردنیاز در دسترس باشد، جمع‌آوری اطلاعات تلخیصی برای شاخص‌های موردنظر در مطالعه به‌آسانی امکان‌پذیر است. این موضوع باعث می‌شود مطالعات همبستگی کم‌هزینه و در زمان کوتاه انجام‌پذیر باشند. منابع اطلاعاتی این مطالعات می‌توانند آمارهای ثبتی، سرشماری‌ها، بانک‌های اطلاعاتی و مطالعات میدانی گسترده باشند.
 - ♦ **وجود محدودیت‌های مختلف در اندازه‌گیری‌های انفرادی:** در مطالعات مرتبط با اثر مواجهه‌های طبیعی و بسیاری دیگر از مطالعات، اندازه‌گیری‌های انفرادی سطوح مواجهه یا دوز مصرف در مقیاس وسیع، عملی، قابل‌اعتماد و معتبر نبوده یا حداقل بدون صرف زمان و بودجه هنگفت امکان‌پذیر نیست. در این مواقع احتمالاً تنها روش عملی، اندازه‌گیری به‌صورت کلان است [۳] تا [۵].
 - ♦ **وجود محدودیت در طراحی مطالعات انفرادی:** هنگامی که مواجهه تحت بررسی، تغییرات کمی در گروه‌های موردبررسی داشته باشد، طراحی مطالعه‌ای بر مبنای اندازه‌گیری مواجهه برای تک‌تک افراد، احتمالاً عملی نیست. در چنین مواقعی با استفاده از مطالعات همبستگی امکان تحت پوشش قرار دادن نواحی جغرافیایی گسترده وجود دارد.
 - ♦ **هدف مطالعه:** گاهی اوقات یافتن یک رابطه اکولوژیکی در جوامع مختلف، هدف اصلی مطالعه است. مثلاً برای بررسی اثرات مداخله‌های جمعیتی، نظیر سیاست‌های جدید بهداشتی مطالعات همبستگی بسیار مناسب هستند [۶] تا [۸].
- باوجود مزایای عملی مطالعات همبستگی، مشکلات و دشواری‌های متدولوژیکی در طراحی و تفسیر نتایج این مطالعات وجود دارد که انجام استنباط‌های سببی از این مطالعات را به‌شدت محدود ساخته است. برخی از این مشکلات عبارت‌اند از:
- ♦ **تورش اکولوژیک:** مهم‌ترین محدودیت این مطالعه، وجود تورش اکولوژیک یا تجمعی است. تورش اکولوژیک ممکن است به علت وجود ارتباطی که میان متغیرها در سطح کلان و میان گروه‌ها مشاهده می‌شود رخ دهد، درحالی‌که لزومی ندارد در سطح فردی این رابطه برای تک‌تک افراد صادق باشد.
 - ♦ **تورش سوء طبقه‌بندی:** به علت ساختار این مطالعه که مستلزم جمع‌آوری داده‌ها در مقیاس کلان است، اثر سوء طبقه‌بندی در این مطالعات بسیار جدی است.
 - ♦ **کیفیت داده‌ها:** عدم وجود اطلاعات موردنیاز و یا پایین بودن کیفیت ثبت و جمع‌آوری داده‌های اکولوژیک یکی از مشکلات عملی انجام این مطالعات است [۹] و [۱۰].

۳-۲-۴ مطالعه مقطعی

مطالعات مقطعی، بررسی‌هایی توصیفی هستند که بجای استفاده از اطلاعات موجود یا منابع در دسترس، اندازه‌گیری عامل مواجهه و معلول (بیماری) در یک‌زمان صورت می‌گیرد. اندازه‌گیری شیوع بیماری‌ها، رفتارها و دیگر ویژگی‌های افراد یک جمعیت از اهداف معمول مطالعات مقطعی هستند؛ اما در برخی مواقع با استفاده از مطالعه مقطعی می‌توان ارتباط بین بیماری‌ها و سایر متغیرهای موردنظر را در شرایط موجود در یک جامعه معین و در یک‌زمان خاص تعیین کرد. این ارتباط می‌تواند به‌صورت مقایسه شیوع یک متغیر در بیماران در مقایسه با غیر بیماران باشد.

مطالعه مقطعی مانند یک عکس از وضعیت جامعه، درباره الگوهای رفتار، عادات، عوامل خطر و شیوع بیماری‌ها است. شکل ۳-۱ شمایی از این مطالعات را نمایش می‌دهد.



شکل ۳-۱: در مطالعات مقطعی، محققان ابتدا نمونه‌ای از جامعه گرفته و سپس متغیرهای مربوط به بیماری و عوامل خطر را اندازه‌گیری می‌کنند.

یکی از موارد کاربرد مطالعات مقطعی، ارزیابی وابستگی و همراهی بین متغیرهاست، اما وابستگی‌های محاسبه‌شده الزاماً به معنی رابطه سببی نیست و اساساً نام‌گذاری متغیرهای پاسخ و مستقل در این نوع مطالعات ظاهری بوده و تنها به فرضیات اولیه پژوهش بستگی دارد. علت اساسی این است که در این نوع مطالعات روند منطقی حرکت از علت «مواجهه» به سمت معلول «بیماری» رعایت نشده است و اساساً هیچ‌گونه پیگیری در مطالعه مقطعی انجام نمی‌شود. البته در متغیرهایی نظیر سن، جنس، قومیت، عوامل ژنتیکی که معمولاً به‌وسیله عوامل دیگر تحت تأثیر قرار نمی‌گیرد تعیین متغیر پاسخ و مستقل آسان است، اما همیشه تشخیص متغیر پاسخ و مستقل به همین سادگی نیست. به‌عنوان مثال وقتی که در یک مطالعه مقطعی رابطه بین چاقی و تماشای تلویزیون در کودکان مشاهده می‌شود، تعیین این که کدام یک علت دیگری بوده است تقریباً ناممکن است.

مزایا و معایب مطالعات مقطعی:

- ♦ یکی از مهم‌ترین مزایای مطالعات مقطعی نسبت به مطالعات هم‌گروهی و مداخله‌ای این است که برای وقوع پیشامد نهایی لازم نیست منتظر بمانیم. این موضوع باعث سریع و کم‌هزینه شدن مطالعه می‌شود.
- ♦ مشکل گم‌شدن موردها در پیگیری وجود ندارد.
- ♦ مطالعات مقطعی معمولاً به‌عنوان نقطه شروع مطالعه تحلیلی (به‌منظور تعیین گروه‌های مواجهه و بدون مواجهه و هم‌چنین افراد سالم و بیمار) در روند پژوهش قرار دارند. مطالعه مقطعی مطالعه‌ای است که در آن، قادر به محاسبه شیوع بیماری یا عامل خطر هستیم.
- ♦ از نقاط ضعف مطالعات مقطعی، چالش‌ها و دشواری‌های موجود در بیان یک رابطه سببی برای آن چیزی است که محقق معلول و عامل خطر می‌نامد.
- ♦ هنگامی که پیامد تحت مطالعه نادر باشد، مطالعه مقطعی عملی نیست. بخصوص هنگامی که هدف، نمونه‌گیری از جمعیت عادی باشد. (به‌عنوان مثال در یک مطالعه مقطعی در تعیین شیوع سرطان معده در افراد ۴۵-۵۹ ساله جامعه حدوداً ۱۰,۰۰۰ نمونه لازم است تا یک مورد سرطان پیدا شود).

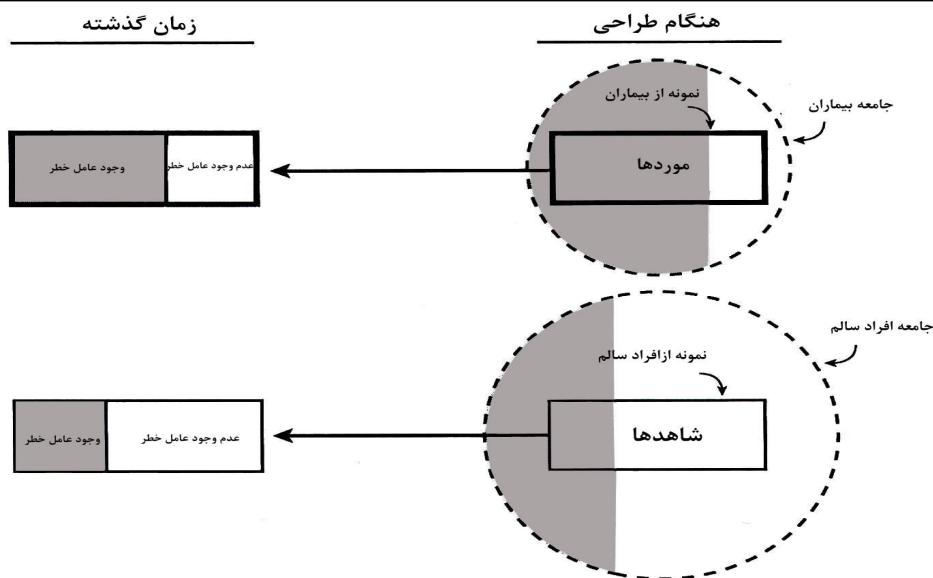
۳-۳ مطالعات مشاهده‌ای

مطالعات مشاهده‌ای شامل دو گروه کلی مطالعات مورد-شاهدی و مطالعات هم‌گروهی هستند. هرکدام از این مطالعات در این بخش مورد بررسی قرار می‌گیرند

۳-۳-۱ مطالعات مورد-شاهدی

برای بررسی علل بیماری، در اکثر مواقع مطالعات هم‌گروهی و مقطعی بر مبنای نمونه‌گیری از جمعیت افراد سالم، گران و هزینه‌بر هستند و لازم است هزاران نمونه برای تعیین عوامل خطر (مثلاً در مورد بیماری نادری همچون سرطان معده) بررسی شود. گاهی اوقات به‌وسیله یک مجموعه از بیماران مبتلا به‌نوعی بیماری خاص، می‌توان عوامل خطر واضحی را نمایان ساخت (مثلاً تزریق مواد مخدر و بیماری ایدز). در این صورت از اطلاعات قبلی خود درباره شیوع عامل خطر در جمعیت مبنا استفاده کرده‌ایم؛ اما در اکثر مواقع، لازم است یک گروه مرجع مشخص شود تا اندازه عامل خطر در افراد مبتلا به بیماری (موردها) قابل‌مقایسه با اندازه عامل خطر در افراد عادی (شاهدها) شود. شکل ۳-۲ نحوه عملکرد این طرح مطالعه را نمایش می‌دهد.

طراحی مطالعات مورد-شاهدی به علت امکان مواجهه با موارد مختلف تورش بسیار چالش‌زا است؛ اما مثال‌های زیادی از مطالعات مورد-شاهدی با طراحی مناسب که باعث حصول نتایج مهم شده است وجود دارد. در مطالعات مورد-شاهدی قادر به برآورد بروز یا شیوع بیماری نیستیم، زیرا نسبت افراد دارای بیماری و سالم در مطالعه به‌وسیله محققان و نه به‌وسیله نمونه‌گیری تصادفی از جامعه تعیین می‌شوند. آنچه از یک مطالعه مورد-شاهدی به دست می‌آید، برآوردی به‌صورت نسبت شانس است.



شکل ۳-۲: در مطالعات مورد-شاهدی محققان، ابتدا نمونه‌ای از جامعه بیماران (موردها) و سپس نمونه‌ای از جامعه افراد در معرض خطر اما سالم (شاهدها) می‌گیرند و عوامل خطر بیماری را اندازه‌گیری می‌کنند.

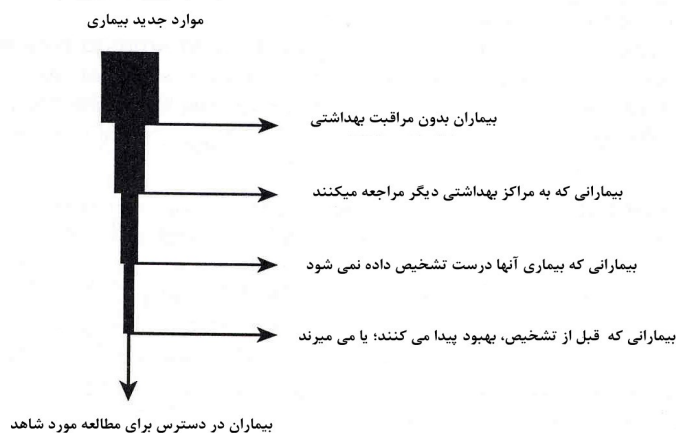
مزایا و معایب مطالعات مورد-شاهدی: مهم‌ترین مزایای مطالعات مورد-شاهدی عبارت‌اند از:

- ♦ مطالعات مورد-شاهدی ارزان‌تر و سریع‌تر از مطالعاتی هستند که مبتنی بر پیگیری هستند. زیرا هیچ زمان انتظاری برای ایجاد بیماری صرف نمی‌شود. این موضوع در مطالعه بیماری‌هایی که زمان زیادی برای بروز آنها نیاز است بسیار مطلوب است.
- ♦ عوامل خطر زیادی به‌طور هم‌زمان در این مطالعه قابل بررسی است. این موضوع باعث بررسی فرضیات زیادی در یک مطالعه مورد-شاهدی می‌شود.
- ♦ این مطالعات روش نسبتاً مناسبی برای بررسی عوامل خطر در بیماری‌های نادر هستند.
- ♦ مطالعات مورد-شاهدی نیاز به حجم نمونه بسیار کمتری نسبت به مطالعات هم‌گروهی مشابه دارند. برای مثال در مقایسه با مطالعات هم‌گروهی مشابه، در مواقعی که کل حجم نمونه کم باشد، ارزیابی متغیرهای مخدوش‌کننده و اثر متقابل عوامل خطر با توان آماری بالاتری انجام می‌شود.
- ♦ به علت رعایت نشدن توالی زمانی در مطالعات مورد-شاهدی، در این مطالعات قادر به ارزیابی دقیق روابط سببی نیستیم.
- ♦ در مطالعات مورد-شاهدی قادر به برآورد خطر نیستیم و تنها در موارد خاصی تقریبی از خطر نسبی قابل محاسبه است (با این‌ها نسبت شانس شاخصی است که محاسبه آن به‌اندازه کافی از اعتبار و جذابیت علمی برخوردار است).
- ♦ در مطالعات مورد-شاهدی منابع تورش بالقوه بسیار وسیع هستند. در اکثر مواقع منبع تورش، نحوه

انتخاب شاهد‌ها است.

۳-۳-۱- انتخاب موردها در مطالعات مورد-شاهدی

نمونه‌گیری در مطالعات مورد-شاهدی با جمع‌آوری موردها آغاز می‌شود. شناخت این موضوع که افراد گروه مورد نماینده کدام بخش از بیماران تحت مطالعه هستند بسیار مهم است. در حالت ایده‌آل، نمونه‌گیری از موردها، شامل نمونه‌ای تصادفی از تمام افرادی است که مبتلا به بیماری موردنظر هستند. با این ایده اولین سؤالی که پیش می‌آید اطلاع از این موضوع است که چه کسانی بیماری موردنظر را دارند و چه کسانی این بیماری را ندارند. در مطالعات مقطعی و هم‌گروهی محققان به‌طور منظم وجود بیماری را در تمام افراد تحت مطالعه بررسی می‌کنند؛ اما در مطالعه مورد-شاهدی موردها از جامعه‌ای انتخاب می‌شوند که وجود بیماری در آن‌ها از قبل تشخیص و گزارش شده است. این نمونه، مجموعه مناسب و نماینده بی‌طرفانه و کاملاً تصادفی از تمام بیماران مبتلا نیست، زیرا موارد تشخیص داده نشده، اشتباه تشخیص داده شده، گزارش و ارجاع نشده و فوت شده در انتخاب حذف شده‌اند. به همین علت انتخاب افراد تحت مطالعه در مطالعات مورد-شاهدی دارای اهمیت دوچندانی است. شکل ۳-۳ که به کوه یخ بیماری معروف است به بیان این موضوع می‌پردازد [۱۱]-[۱۳].



شکل ۳-۳: امکان دارد مطالعه مورد-شاهدی نماینده مناسبی از جامعه بیماران نباشد

قبل از انتخاب موردها، تعریف دقیقی از بیماری تحت بررسی باید ارائه شود. در صورت واضح و مشخص نبودن تعاریف، خطر سو طبقه‌بندی وجود دارد. اگر تعریف موردها محدوده وسیعی را شامل شود، مطالعه مورد-شاهدی بی‌حاصل خواهد بود. برای مثال بیماری روانی واژه‌ای کلی است که بیماری‌های مختلف و متفاوتی را شامل می‌شود.

دو منبع عمده برای انتخاب افراد گروه مورد وجود دارد:

۱- بیمارستان

۲- جمعیت

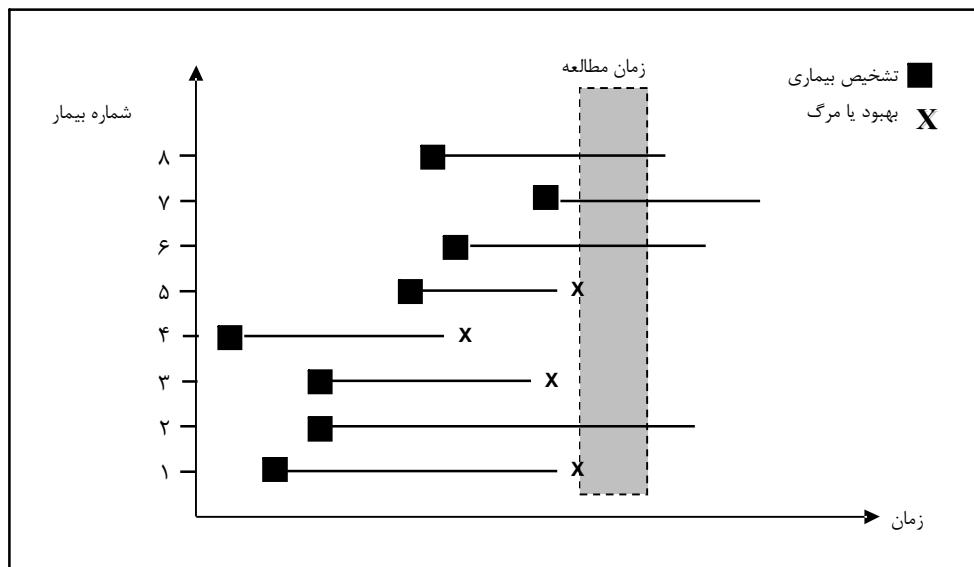
مزیت مشخص موردهای بیمارستانی دسترسی ساده به افراد و آسان بودن انجام مطالعه از لحاظ لجستیکی است؛ اما مشکل اصلی این منبع وجود تورش انتخاب در به کارگیری این روش است. انتخاب موردها از جامعه امکان ایجاد سوگرایی انتخاب را به حداقل می‌رساند. در این روش نیاز به نظام‌های ثبت و گزارش دقیق در مورد بیماری تحت مطالعه بسیار حیاتی است.

نکته دیگر در انتخاب گروه موردها شیوه انتخاب گروه مورد است (موارد جدید یا موارد موجود). شکل ۳-۴ (الف و ب)، نحوه انتخاب موردها در هر کدام از این دو حالت را نشان می‌دهد. شکل ۳-۴-الف مربوط به نحوه انتخاب موردهای موجود است. در این مطالعه تنها موردهای ۲، ۶، ۷ و ۸ که در زمان انجام مطالعه بیمار و زنده هستند بررسی می‌شوند. شکل ۳-۴-ب مربوط به انتخاب موردهای جدید است. در این مطالعه تمام موارد تشخیص داده شده در طول انجام مطالعه (صرف‌نظر از این که چه مدت پس از تشخیص، بیمار یا زنده هستند) بررسی می‌شوند.

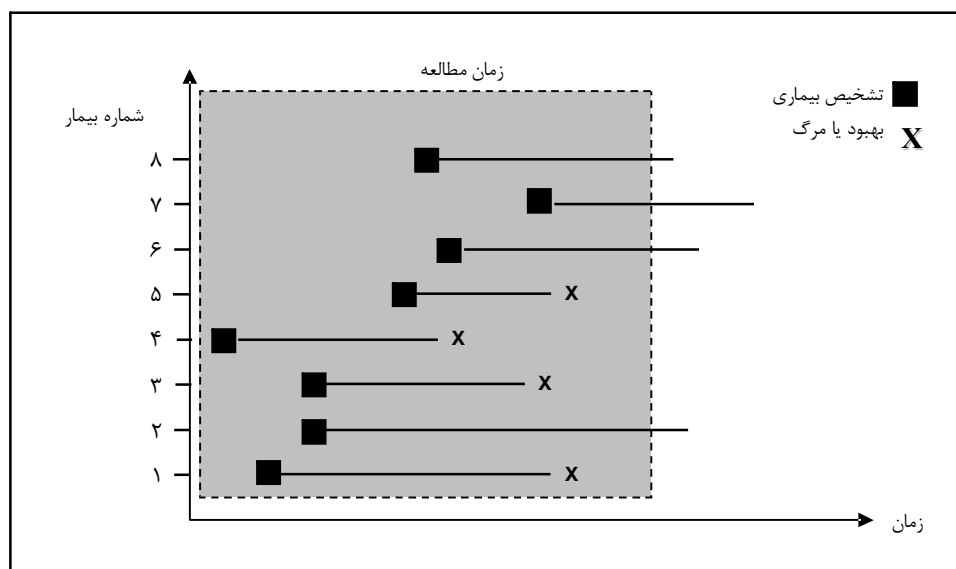
نکته مهم در انتخاب موارد جدید این است که چون هدف مطالعه بررسی عوامل مؤثر بر ایجاد بیماری است، نیازی به پیگیری وجود ندارد.

معیارهای ورود و خروج: گاهی اوقات تنها گروه‌هایی از بیماری در مطالعه مدنظر هستند که واجد یا فاقد ویژگی‌های خاص باشند. به عنوان مثال حذف گروه‌هایی که دارای یک یا چند بیماری همراه دیگر هستند یا مواردی استثنایی که عوامل خطر در آنها فوق‌العاده بیشتر از سایر افراد است، در بالا رفتن اعتبار داخلی مطالعه بسیار مؤثر است. شاید در شرایطی لازم باشد جمع‌آوری موردها را محدود به دوره زمانی مشخص و مکانی معین کنیم. به همین علت در بسیاری از مطالعات، موردها از یک منبع محدود جمع‌آوری می‌شوند (از نظر زمانی و مکانی). در مقابل برای برخی مطالعات، لازم است تعدادی از موردها را از مجموعه حذف کرد، زیرا شانس اندکی برای مواجهه با عوامل خطر برای آنها وجود داشته است.

منابع: موردها معمولاً بیماران تشخیص داده شده در سیستم پزشکی هستند. پرونده‌های بیمارستانی، دفاتر اتاق عمل و گزارش‌های پاتولوژی منابع معمول در این زمینه هستند. گاهی اوقات نیز ترکیبی از این منابع برای انتخاب موارد استفاده می‌شود.



شکل ۳-۴-الف: شیوه انتخاب موردهای موجود در مطالعه مورد-شاهدی. در این روش تنها بیماران شماره ۲، ۶، ۷ و ۸ به مطالعه وارد می‌شوند.



شکل ۳-۴-ب: شیوه انتخاب موردهای جدید. در این روش همه بیمارانی که در فاصله زمانی انجام مطالعه، بیماری آن‌ها تشخیص داده شده در مطالعه بررسی می‌شوند.

شکل ۳-۴: روش‌های انتخاب موردها در مطالعه مورد-شاهدی

انتخاب شاهد‌ها (گروه کنترل) در مطالعات مورد-شاهدی:

در عمل، منابع بیماران که موردها را تشکیل می‌دهند محدود هستند. به همین علت انتخاب آنها در مطالعه نسبتاً ساده است. قسمت مشکل تصمیم‌گیری در هر مطالعه مورد-شاهدی، انتخاب شاهد‌های مناسب است. شرایط انتخاب گروه شاهد به‌جز ملاک‌های مربوط به تشخیص بیماری باید منطبق بر شرایط انتخاب گروه مورد باشد. مهم‌ترین و بحث‌انگیزترین مقوله در مطالعات مورد-شاهدی نحوه انتخاب شاهد‌ها و قابل‌مقایسه بودن آنها با موردها است [۱۴].

اصول زیر در انتخاب گروه شاهد باید رعایت شود:

- ♦ گروه شاهد باید از بین افرادی انتخاب شود که عاری از بیماری تحت مطالعه باشد. معمولاً گروه شاهد متشکل از افرادی هستند که هرچند هم‌اکنون بدون بیماری تحت مطالعه هستند، اما انتظار داریم در آینده مبتلا شوند.
 - ♦ این گروه باید از همان جامعه مبتلا که موردها از آن انتخاب شده‌اند، انتخاب شود. این موضوع برای اجتناب از تحریف اثرات مواجهه، وجود عوامل خطر ناشناخته یا اندازه‌گیری نشده و عوامل مخدوش‌کننده بالقوه صورت می‌گیرد. معیارهای ورود و خروج از مطالعه (به‌جز بیماری تحت مطالعه) باید در دو گروه مشابه باشد.
 - ♦ گروه شاهد‌ها باید به‌طور بالقوه شانس ابتلا به بیماری را داشته باشند. برای مثال زنانی که رحم آنها برداشته شده است نباید به‌عنوان شواهد در مطالعه‌ای مربوط به سرطان مخاط رحم انتخاب شوند.
 - ♦ گروه شاهد باید از همان جایی انتخاب شود که افراد گروه مورد انتخاب می‌شوند. به‌عبارت‌دیگر هنگامی که منبع انتخاب موردها جمعیت است، باید شاهد‌ها نیز باید از بیمارستان انتخاب شوند. راهبردهای زیر در انتخاب شاهد‌ها متداول است.
 - ♦ گروه شاهد بیمارستانی
 - ♦ همسان کردن
 - ♦ نمونه‌گیری بر مبنای افراد سالم جامعه
 - ♦ استفاده از گروه‌های شاهد چندگانه
- اکنون به بحث و بررسی منابع ذکر شده برای انتخاب شاهد‌های مطالعه مورد-شاهدی می‌پردازیم.

- شاهد‌های بیمارستانی

یکی از روش‌های موجود برای برطرف کردن تورش انتخاب در این مطالعات، استفاده از گروه شاهدی است که در همان بیمارستان یا کلینیک بستری شده‌اند. عوامل تأثیرگذار زیادی وجود دارد که بیماران، کلینیک یا بیمارستان خاصی را جهت معالجه انتخاب می‌کنند. پاره‌ای از این عوامل عبارت‌اند از وضع مالی، منطقه مسکونی، قومیت و وابستگی مذهبی مشابه با موردها. این امر اثر عوامل تأثیرگذار را تا حدود زیادی خنثی می‌کند. به‌عنوان مثال در یک مطالعه مورد-شاهدی درباره همبستگی بین سیگار کشیدن و سکت قلبی در زنان، افراد

مورد را می‌توان از کسانی که در واحد مراقبت قلبی (C.C.U) بیمارستان پذیرش شده‌اند انتخاب کرد و افراد شاهد را از بین کسانی که در بخش‌های جراحی، ارتوپدی و خدمات پزشکی همان مراکز پزشکی پذیرش شده‌اند و شامل بیماران عضلانی- استخوانی و داخلی و انواع بیماری‌های به‌جز بیماری‌های قلبی هستند انتخاب کرد. چند مزیت مهم عملی و علمی در استفاده از گروه شاهد بیمارستانی وجود دارد. اول اینکه شناسایی این افراد آسان بوده و به‌راحتی و به تعداد کافی در دسترس هستند، بنابراین مخارج مطالعه و کارهای عملی که باید انجام گیرد به حداقل می‌رسد.

دوم اینکه چون این افراد در بیمارستان بستری هستند بیشتر از افراد سالم تمایل دارند که از سابقه مواجهه یا پیشامدهای گذشته با اطلاع باشند. به‌این ترتیب تشابه بین دو گروه برای گزارش اطلاعات باعث کاهش تورش یادآوری می‌شود.

سوم آنکه استفاده از بیماران بستری در بیمارستان که به علت‌های دیگر تحت معالجه هستند، به این معنی است که بین این گروه و موردهای تحت مطالعه تشابهاتی وجود داشته که باعث شده به یک کلینیک یا بیمارستان خاص مراجعه کنند.

چهارم آنکه گروه شاهد بیمارستانی همانند گروه موردها تمایل بیشتری نسبت به افراد سالم برای همکاری نشان می‌دهند، بنابراین تورش ناشی از پاسخ ندادن و همکاری نکردن کاهش می‌یابد.

مهم‌ترین عیب استفاده از گروه شاهد بیمارستانی این است که بر طبق تعریف، این افراد بیمار هستند و لذا با افراد معمولی متفاوت هستند. بنابراین وضعیت این بیماران ممکن است نماینده مناسبی برای بررسی شیوع عوامل خطر در جمعیتی که گروه مورد از آن گرفته شده است نباشد.

مطالعه‌های انجام شده روی گروه‌های شاهد بیمارستانی و غیر بیمارستانی ثابت کرده که برای مثال بیماران بستری در بیمارستان بیشتر سیگار می‌کشیدند، بیشتر قرص‌های ضدبارداری استفاده می‌کردند و یا در کشورهای غربی مشروبات الکلی بیشتری نسبت به افراد غیر بستری مصرف می‌کردند به همین علت اگر فرضاً یک مطالعه مورد-شاهدی با استفاده از شاهد‌های بیمارستانی در مورد استفاده از قهوه و ابتلا به سرطان مثانه انجام شود ممکن است باعث به دست آمدن ارتباط معنی‌دار بین این دو متغیر شود، که اگر از گروه شاهد دیگری استفاده می‌شد این نتیجه به دست نمی‌آمد.

این پرسش که چه رده یا رده‌هایی از بیماران باید در گروه شاهد بیمارستانی گنجانده شود مسئله مهمی است. موضوعی که باید موردتوجه قرار گیرد این است که نوع بیماری که گروه شاهد به خاطر آن بستری شده‌اند، ممکن است با عوامل خطر مورد مطالعه ارتباط داشته باشد. بنابراین در عمل بیمارانی که همبستگی مثبت یا منفی با عوامل خطر مورد نظر دارند باید از گروه شاهد حذف شوند. به‌عنوان مثال سیگار کشیدن یک عامل خطر مهم برای بسیاری از بیماری‌های تنفسی مانند برونشیت و ذات‌الریه است. قرار دادن چنین بیمارانی در گروه مقایسه یک مطالعه مورد-شاهدی درباره سیگار کشیدن و سرطان ریه، باعث می‌شود که نسبت بالایی از افراد سیگاری در گروه شاهد قرار بگیرند در نتیجه این گروه نماینده مناسبی برای افراد سالم جامعه نیست. یک راهبرد مناسب در انتخاب گروه شاهد بیمارستانی انتخاب بیماران از بخش‌های مختلف بیمارستان که با بیماری مورد مطالعه ایجاد شباهت تشخیصی ندارند یا در ارتباط با عوامل خطر بیماری مورد مطالعه نیستند، است.

- همسان کردن

همسان‌سازی یکی از شیوه‌های ساده و مؤثر برای قابل‌مقایسه کردن گروه موردها و شاهدها است. در این روش عوامل مهم مرتبط با بیماری، که محقق مایل به بررسی اثر آن بر بیماری نیست، همسان می‌شوند. برای مثال عوامل و بیماری‌های زیادی با سن و جنس مرتبط هستند. به همین علت، مقایسه دو گروه درحالی‌که از نظر سن و جنس قابل‌مقایسه نیستند، چندان منطقی به نظر نمی‌رسد. یکی از روش‌ها برای اجتناب از این مسئله، همسان کردن شاهدها و موردها از لحاظ سن و جنس است. در ضمن بسیاری از عوامل خطر که اندازه‌گیری آنها دشوار است (نظیر عوامل محیطی، ژنتیکی و هورمونی) که با سن و جنس و سوابق خانوادگی در ارتباط هستند به‌وسیله همسان کردن در مطالعه تعدیل می‌شوند.

همسان‌سازی به دو روش همسان‌سازی فردی و گروهی انجام می‌شود. در همسان‌سازی فردی برای هر یک از افراد گروه مورد در مطالعه یک شاهد مشابه (از نظر متغیرهای موردنظر) انتخاب می‌شود. در همسان‌سازی گروهی افراد گروه شاهد به طریقی انتخاب می‌شوند که نسبت متغیرهای معینی از آنها مشابه نسبت آن متغیرها در افراد گروه مورد باشد. برای این مثال اگر ۳۰ درصد از افراد گروه مورد کارگر هستند، ۳۰ درصد از افراد گروه شاهد نیز از کارگرها انتخاب می‌شوند. معمولاً در این روش باید ابتدا تمام افراد گروه مورد انتخاب‌شده و افراد گروه شاهد متناسب با آنها انتخاب شوند.

- نمونه‌گیری بر مبنای افراد سالم جامعه

انتخاب گروه شاهد از جمعیت عادی، از راه‌های گوناگونی انجام می‌گیرد. از جمله نمونه‌گیری از مناطق هدف در مطالعه، شماره تلفن‌های اتفاقی یا تعیین افراد از طریق نقشه‌های سرشماری است. به‌هرحال، انتخاب افراد سالم از جمعیت مشکلاتی را در بر دارد. تعیین و بررسی افراد شاهد از جمعیت عمومی، به‌طورمعمول گران‌تر و وقت‌گیرتر از شاهدهای بیمارستانی است. فهرست افراد جامعه همیشه در دسترس نیست و به خاطر گرفتاری‌های شغلی و با توجه به وضعیت اوقات فراقت آنها، تماس گرفتن و برقرار کردن ارتباط با آنان، به‌طورمعمول دشوار است. علاوه بر این کیفیت اطلاعات بین دو گروه مورد و شاهد ممکن است متفاوت باشد، زیرا کسانی که در جامعه معمولی زندگی می‌کنند، ممکن است به‌اندازه کسانی که دچار بیماری شده‌اند، عوامل خطر را به خاطر نیاورند. همچنین افرادی که اخیراً مشکل بهداشتی خاصی نداشته‌اند، انگیزه کمتری برای مشارکت در مطالعه دارند. بنابراین تعداد افرادی که در گروه شاهد جامعه از مشارکت در برنامه خودداری می‌کنند، تقریباً همیشه بیش از گروه مورد و گروه شاهد بیمارستانی است.

منبع دیگر انتخاب شاهدها، شامل گروه‌های خاصی نظیر دوستان، وابستگان، همراهان یا همسایگان موردها است. این گروه‌ها برخی مزایای گروه شاهد جامعه را دارا هستند، زیرا سالم‌اند و تمایل بیشتری به همکاری نسبت به جمعیت عادی دارند. مزیت این روش این است که برخی عوامل مخدوش‌کننده اصلی وابسته به نژاد، وضعیت اقتصادی و اجتماعی و محیطی مهار می‌شود.

- استفاده از گروه‌های شاهد چندگانه

گاهی اوقات استفاده از یک گروه شاهد، نواقص خاصی دارد که با استفاده از گروه دیگر قابل جبران است.

به‌عنوان مثال در یک مطالعه مورد-شاهدی درباره رابطه مصرف قهوه و سرطان لوزالمعده، بیماران باید با گروهی از افراد شاهد که به دلایل دیگری در بیمارستان بستری شده‌اند مقایسه شوند. درحالی‌که استفاده از گروه شاهد بیمارستانی به دلیل کم‌هزینه بودن و دسترسی آسان مناسب است؛ اما موضوع مهم آن است که شاهد‌های بیمارستانی از لحاظ مصرف قهوه با افراد سالم جامعه تفاوت دارند. برای اصلاح این وضعیت، گروه شاهد دیگری که شامل افراد عادی جامعه است باید انتخاب شود. مقایسه همبستگی مشاهده‌شده در مورد مصرف قهوه و سرطان لوزالمعده در هر یک از گروه‌های شاهد، اطلاعاتی در مورد اثر حقیقی این عامل و هم‌چنین کافی نبودن استفاده از گروه شاهد بیمارستانی برای مطالعه چنین فرضیه‌ای را اثبات می‌کند.

پس از مشخص شدن منبع و تعداد گروه‌های شاهد لازم است در مورد نسبت شاهد‌ها به موردها تصمیم‌گیری شود. بهترین نسبت شاهد به بیمار یک‌به‌یک است، اما زمانی که حجم نمونه بیماران محدود است و یا هزینه به دست آوردن اطلاعات در گروه مورد زیاد است نسبت شاهد به بیمار می‌تواند تا حدی تغییر یابد؛ اما به‌طور معمول توصیه نمی‌شود که این نسبت از ۴ به ۱ بیشتر شود.

پس از تعریف خصوصیت و منابع انتخاب گروه‌های مورد و شاهد، باید اطلاعات درباره بیماری و مواجهه به دست آید. نکته مهم دیگر دقت و صحت مقایسه و قابل‌مقایسه بودن منابع اطلاعات است. اطلاعات لازم درباره وضعیت بیماری را می‌توان از چندین منبع از جمله بازبینی گواهی‌های مرگ، ثبت موردها در برنامه‌های پایش، پرونده‌های مطب پزشکان، پرونده‌های پذیرش و ترخیص بیمارستان‌ها، دفاتر اتاق عمل و بخش آسیب‌شناسی به دست آورد. اطلاعات مربوط به مواجهه را می‌توان از آزمودنی‌های مطالعه شامل پرسش‌گری و اطلاعات درج‌شده در پرونده‌های پزشکی به دست آورد. رویکردهای استفاده شده برای به دست آوردن اطلاعات در دو گروه شاهد و مورد باید تا آنجا که امکان دارد مشابه باشد. به‌عنوان مثال، باید محل و شرایط مصاحبه در هر دو گروه یکسان باشد و در موقعیت‌هایی که امکان کور کردن پرسشگر یا خلاصه‌کننده پرونده‌های پزشکی درباره وضعیت موردها و شاهد‌ها وجود داشته باشد، این موضوع حتماً باید رعایت گردد. از این گذشته، افرادی که مسئول جمع‌آوری اطلاعات هستند باید حتی‌المقدور درباره فرضیه‌های مطالعه بی‌اطلاع نگه داشته شوند. این کار برای به حداقل رساندن تورش مشاهده‌گر بسیار ضروری است [۱۴].

۳-۳-۲ مطالعات هم‌گروهی

مطالعات هم‌گروهی، مطالعاتی هستند که در آن‌ها افراد تحت مطالعه در طول زمان پیگیری می‌شوند. دو هدف اصلی در انجام مطالعات هم‌گروهی وجود دارد:

- ♦ توصیف: برای توصیف موارد بروز پیشامدی خاص در طول زمان
- ♦ تحلیل: برای ارزیابی همراهی و وابستگی بین متغیرهای پاسخ و مستقل

به‌طور کلی مطالعات هم‌گروهی به دو صورت کلی آینده‌نگر و گذشته‌نگر تقسیم می‌شوند. در مطالعات هم‌گروهی آینده‌نگر، اعضای نمونه، متغیرهای زمینه‌ای و متغیرهای مواجهه قبل از وقوع پیشامدها در ابتدای مطالعه توسط محققان معین و اندازه‌گیری می‌شوند. افراد تحت مطالعه به گروه‌های دارای مواجهه و بدون مواجهه تقسیم شده و تا وقوع پیشامد موردنظر در یک دوره زمانی معین پیگیری می‌شوند. در مطالعات هم‌گروهی

گذشته‌نگر محققان اعضای نمونه، متغیرهای زمینه‌ای و مواجهه را هنگامی تعیین کرده و مطالعه را آغاز می‌کنند که پیشامدها در افراد مورد مطالعه رخ داده است. در این نوع مطالعه با بررسی سوابق افراد تحت مطالعه یک نوع پیگیری تصنعی به صورت گذشته‌نگر بازسازی می‌شود.

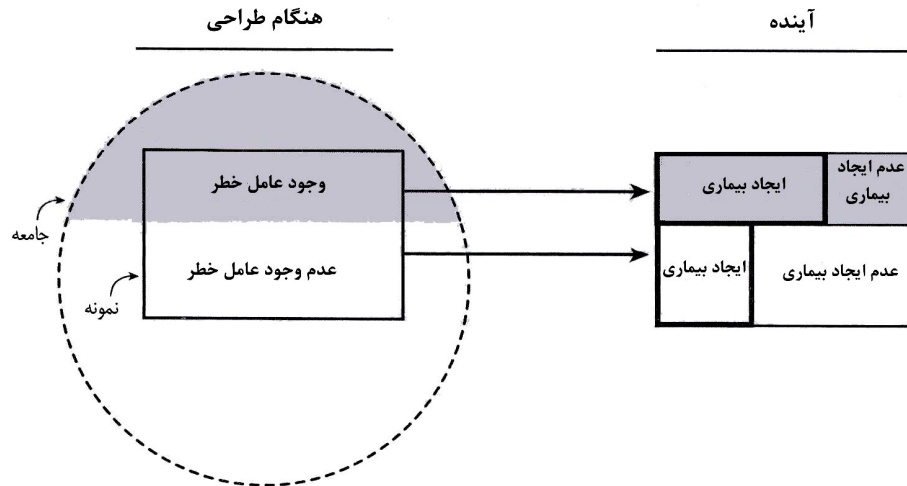
۳-۲-۱ مطالعات هم‌گروهی آینده‌نگر

در مطالعات هم‌گروهی آینده‌نگر، محققان نمونه مورد نظر را انتخاب یا تعریف می‌کنند و سپس مشخصاتی نظیر شیوه زندگی، الگوی فعالیت، محل سکونت، اشتغال و هر متغیر دیگر که احتمال دارد میزان بروز پیشامد مورد نظر را پیش‌بینی نماید را اندازه‌گیری می‌کنند. این اعضا در مرحله پیگیری به‌طور مستمر مورد معاینه قرار می‌گیرد و مشخصات مهم آنها دوباره اندازه‌گیری می‌شود و بروز یا عدم بروز پیشامد مورد نظر در آنها بررسی می‌شوند. در شکل ۳-۵ نحوه انجام این مطالعه نمایش داده شده است.

مطالعه هم‌گروهی در ساده‌ترین حالت آن، شامل انتخاب دو گروه از افراد جامعه (جمعیت مبنا) است که یک گروه افراد دارای مواجهه هستند. مواجهه، ویژگی خاصی است که باعث می‌شود اعضای دارای این ویژگی جزء گروه‌های پرخطر برای ابتلا به پیشامد مورد نظر باشند. (مثلاً سیگار کشیدن در ابتلا به سرطان ریه) و گروه دیگر افرادی هستند که دارای مواجهه نیستند. دو گروه برای مدت معینی پیگیری می‌شوند این پیگیری می‌تواند تا مرگ آخرین افراد هر گروه ادامه یابد و بروز پیشامد مورد نظر در دو گروه مقایسه می‌شود. برای مثال، در مطالعه‌ای که در مورد خطرات مواجهه با زغال‌سنگ انجام می‌شود معدنچیان شاغل در معادن زغال‌سنگ و سایر کارگران صنایع سنگین می‌توانند دو گروه مورد نظر (جمعیت مبنا) را تشکیل دهد. این دو گروه برای مدت معینی (مثلاً ده سال) پیگیری می‌شوند تا بروز برونشیت بین این دو گروه مقایسه شود.

گروه بدون مواجهه (کارگران شاغل در صنایع سنگین) به این علت وارد مطالعه می‌شوند که به‌عنوان گروه کنترل برای مقایسه موارد بیماری مرتبط با عامل مواجهه (زغال‌سنگ) مورد بررسی قرار گیرند. در حالت ایده‌آل فقط کسانی مجاز به ورود در مطالعه هستند که در ابتدا عاری از بیماری یا پیامد مورد نظر باشند. بنابراین در مثال قبل تنها افرادی که بدون نشانگان برونشیت هستند باید به‌عنوان جمعیت مبنا تلقی شوند.

مطالعه هم‌گروهی آینده‌نگر، یک راهبرد نیرومند در توصیف بروز و بررسی علت‌های احتمالی یک پیشامد در شرایط گوناگون است. زیرا عوامل بالقوه ایجادکننده بیماری قبل از وقوع پیامد اندازه‌گیری می‌شوند. توالی صحیح زمانی، باعث تقویت این نظریه که فاکتورهای اندازه‌گیری شده علت ایجاد پیامد هستند می‌شود. مطالعه آینده‌نگر به محققان فرصت اندازه‌گیری کامل و دقیق تمام فاکتورهای مهم را ارائه می‌کند. این موضوع در برخی متغیرها، نظیر رژیم غذایی روزانه، که اندازه‌گیری آنها به علت سلیقه‌ای بودن پاسخها و خطای فراموشی، بسیار مهم است. اندازه‌گیری برخی از متغیرها، پس از ایجاد بیماری به علت خطا و تورش ناشی از اطلاع از بیمار بودن یا مداخله بیولوژیکی بی‌اعتبار و بی‌دقت می‌شوند. این مشکل در مطالعات هم‌گروهی آینده‌نگر از بین می‌رود.



شکل ۳-۵: در مطالعات هم‌گروهی آینده‌نگر محققان نمونه‌ای از جامعه می‌گیرند و متغیرهای مربوط به پیش‌آگهی و عوامل خطر را اندازه‌گیری می‌نمایند و در طول پیگیری این جمعیت بروز بیماری اندازه‌گیری می‌شود.

مزایا و معایب مطالعات هم‌گروهی

- ♦ مطالعات هم‌گروهی آینده‌نگر معمولاً بسیار گران و وقت‌گیر هستند. علت اصلی برنامه‌ریزی برای پیگیری مجموعه بزرگی از افراد در یک دوره طولانی است.
- ♦ این مطالعات برای بررسی بیماری‌هایی با دوره طولانی مناسب نیستند، زیرا دوره زمانی مطالعه به‌طور غیرقابل‌قبولی طولانی خواهد شد. برای مثال اگر جوان‌های سالم سیگاری و غیر سیگاری تا زمان ایجاد سرطان ریه پیگیری شوند، حداقل به یک برنامه بیست‌ساله برای پیگیری نیازمندیم تا تومور، ایجاد شده و به مرحله تشخیص برسند.
- ♦ مطالعات آینده‌نگر هم‌گروهی برای بررسی بیماری‌های نادر مناسب نیستند، زیرا در این صورت به جمعیت مبتدای بسیار بزرگ با پیگیری بسیار طولانی نیازمندیم که هر دوی این موارد می‌توانند غیرقابل‌قبول باشند. برای مثال وقتی بروز یک بیماری، سالانه ۵ مورد در ۱۰۰,۰۰۰ نفر باشد و مایل باشیم حدوداً ۱۰۰ مورد را برای بررسی یک رابطه معنی‌دار بین مواجهه و بیماری مطالعه کنیم، باید یک جمعیت اولیه هم‌گروهی ۲۰۰,۰۰۰ نفری را ده سال پیگیری کنیم. بررسی یک جمعیت اولیه هم‌گروهی ۴۰۰,۰۰۰ نفری با پنج سال پیگیری نیز نتیجه‌ای مشابه دارد.
- ♦ اثر مطالعه عامل مهمی است که ممکن است برخی ویژگی‌های قابل‌تغییر در افراد تحت مطالعه را تحت تأثیر قرار دهد. برای مثال ممکن است اعضای مورد مطالعه، پس‌ازاین که تحت مطالعه و پیگیری مستمر قرار گرفتند، به‌طور آگاهانه الگوی زندگی، رژیم غذایی و میزان فعالیت روزانه خود را تغییر دهند و یا عامل

- مواجهه را حذف کنند. فرضاً سیگار را ترک کنند و یا خود را در معرض اشعه مستقیم آفتاب قرار ندهند. این موضوع به شدت صحت استنباطها را در مورد گروه‌های مواجهه و عدم مواجهه تحت تأثیر قرار می‌دهد.
- ♦ عامل مواجهه، بخصوص در مطالعاتی که زمان پیگیری طولانی است بدون ارتباط با اثر تحقیق ممکن است حذف یا کاهش یابد، اضافه یا تشدید شود. برای مثال یک شخص سیگاری، ممکن است سیگار را ترک کند یا تعداد نخ‌های مصرفی در روز را افزایش دهد. این موضوع علاوه بر ایجاد مشکل در اندازه‌گیری مقدار و مدت‌زمان مواجهه، تحلیل نتایج را نیز پیچیده‌تر می‌کند.
 - ♦ خروج افراد از مطالعه، عاملی است که می‌تواند نتایج را تحت تأثیر قرار دهد. خروج از مطالعه می‌تواند به دلایلی غیر مرتبط با بیماری صورت گیرد، که در این صورت باید از روش‌های خاص آماری در تعدیل آنها استفاده شود و یا می‌تواند مرتبط با بیماری و به صورت منظم صورت پذیرد که باعث ایجاد تورش در نتایج می‌شود در مثال مواجهه با زغال‌سنگ امکان دارد افراد مبتلا به برونشیت به یک منطقه با آب‌وهوای گرم مهاجرت کنند و یا افراد سیگاری به علت مشکلات ریوی از شهر به روستا مهاجرت کنند.
 - ♦ مطالعه اصلی تنها در مورد یک عامل مواجهه و یک پیشامد یا بیماری طراحی می‌شود. وقتی عوامل خطر متفاوت و پیشامدهای مختلف مدنظر هستند، روش‌های پیچیده‌تری موردنیاز است.

۳-۲-۲-۲ مطالعه هم‌گروهی گذشته‌نگر (تاریخی)

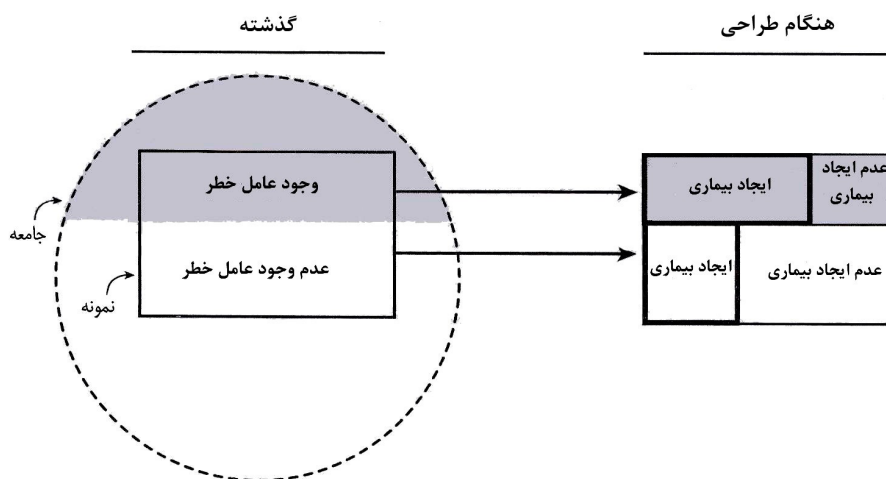
طرح مطالعه هم‌گروهی تاریخی، اساساً مشابه مطالعه هم‌گروهی آینده‌نگر است. گروهی از افراد جامعه در طول زمان پیگیری می‌شوند و متغیرهای عامل خطر بالقوه در آنها اندازه‌گیری می‌شود و از ایجاد بیماری یا ضایعه در آنها اطمینان حاصل می‌شود. تنها تفاوت، تاریخ وقوع وقایع (اندازه‌گیری عوامل خطر، پیگیری و تأیید ایجاد بیماری) است، که تماماً در گذشته صورت می‌گیرد. این نوع مطالعه تنها در صورتی در دسترس بودن داده‌های مناسب و باکیفیت قابل انجام است. امکان دارد این افراد در یک مطالعه هم‌گروهی به‌منظور دیگری تحت پیگیری قرار گرفته باشند. شکل ۳-۶ الگوی نحوه انجام این مطالعه را نشان می‌دهد.

مطالعات هم‌گروهی تاریخی در تعیین عوامل خطر و پیش‌بینی هستند. این مطالعات در مقایسه با مطالعه هم‌گروهی آینده‌نگر سریع‌تر و به‌صرفه‌تر است.

به علت اینکه مطالعه هم‌گروهی تاریخی در گذشته صورت پذیرفته و محقق نقشی در اندازه‌گیری‌های اولیه، پیگیری‌ها و تشخیص بیماری ندارد، ممکن است کیفیت داده‌ها مطلوب و مناسب نباشد. معمولاً داده‌های موجود، بسیاری از اطلاعات موردنیاز برای ارزیابی سؤال‌های پژوهش را در بر ندارد.

در مطالعات تاریخی داده‌های گم‌شده، اندازه‌گیری متغیرها با روشی غیر از آنچه مدنظر محققان است و ناکامل بودن فرم‌های اطلاعاتی از مشکلات جدی تحقیق هستند.

انتخاب جمعیت دارای مواجهه: افرادی که برای بررسی به‌عنوان جمعیت دارای مواجهه در یک مطالعه هم‌گروهی انتخاب می‌شوند، ممکن است از منابع گوناگونی باشند. انتخاب یک گروه خاص، وابسته به



شکل ۳-۶: در مطالعات هم‌گروهی گذشته‌نگر، محقق جمعیت هم‌گروهی را که در گذشته وجود داشته بازسازی می‌کند و اطلاعات مربوط به عوامل خطر و پیش‌آگهی بیماری که در گذشته اندازه‌گیری شده جمع‌آوری می‌کند. سپس بروز بیماری را اندازه‌گیری می‌کند.

ملاحظات علمی و اجرایی زیادی دارد که برخی از آن‌ها عبارت‌اند از: فراوانی مواجهه مورد مطالعه، نیاز به وجود اطلاعات کامل و دقیق برای مواجهه، امکان دستیابی به افراد مورد مطالعه، قابل‌پیگیری بودن افراد تحت مطالعه و ماهیت پرسش‌هایی که در پژوهش ارزیابی می‌شوند [۱۴].

در صورتی که مواجهه نسبتاً شایع باشد (مانند سیگار کشیدن یا مواجهه با اشعه مستقیم خورشید) به تعداد کافی افراد دارای مواجهه را می‌توان در بین جمعیت‌های هدف شناسایی کرد. ولی در مواجهه‌های کمیاب مانند مواجهه‌های مرتبط با یک حرفه خاص، عامل‌های زیست‌محیطی، شرایط اقلیمی ویژه انتخاب یک جمعیت ویژه کارایی بیشتری دارد. گروه‌های در معرض خطر ویژه می‌توانند شامل افرادی از یک حرفه معین مانند کارگران شاغل در کارخانه تولید لاستیک، معدن‌کاران اورانیوم، سازندگان کشتی، بیماران که به علت ابتلا به سرطان تحت پرتودرمانی قرار دارند یا معاینه‌های فلویوروسکوپی مکرر برای تشخیص بیماری سل داشته‌اند، باشند. افرادی که نزدیک یک محیط زیان‌بار زندگی می‌کنند، مانند زمین محل آزمایش‌های هسته‌ای یا محل انبار کردن زباله‌های سمی، افرادی که رژیم غذایی روزمره آن‌ها با افراد عادی متفاوت است یا شیوه زندگی خاص دارند، افرادی که در یک پیشامد خاص با مواجهه ویژه دیگری روبه‌رو شده‌اند (پرتاب بمب اتمی در هیروشیما یا حادثه‌های صنعتی در کارخانه‌های بزرگ)، مثال‌هایی از این قبیل هستند. مزیت‌گزینه‌های جمعیت دارای یک مواجهه خاص آن است که امکان پیگیری افراد دارای مواجهه در دوره زمانی منطقی و به تعداد کافی فراهم شود.

استفاده از این هم‌گروه‌ها می‌تواند باعث شناسایی عامل بیماری در شرایط ویژه تحت بررسی شود. افزون بر این، هم‌گروه‌ها می‌توانند وسیله کارایی برای شناسایی عوامل خطری که در جمعیت مرجع وجود دارند باشند. به این ترتیب برای ارزیابی ارتباط بین فعالیت بدنی و بیماری سرخرگ تاجی قلب، تعدادی مطالعه هم‌گروهی در بین گروه‌های دارای مواجهه خاص مانند رانندگان اتوبوس و باربران لنگرگاه‌ها در لندن اجرا شد. با اطلاعات

به دست آمده از این پژوهش‌ها با توجه به مزیت‌های آشکار آنها، از نظر کم بودن حجم نمونه و توانایی آشکار نمودن مواجهه و پیگیری مناسب شرکت‌کنندگان برای تعیین پیامد مورد نظر، نتایج مطالعه برای ارزیابی نقش فعالیت بدنی در سبب‌شناسی بیماری سرخرگ تاجی قلب به جمعیت عام تعمیم داده شد [۱۵].

افزون بر این با استفاده از یک هم‌گروه دارای مواجهه خاص امکان ارزیابی پیامدهای نادر فراهم می‌شود. حتی اگر پیامدی در جمعیت عام بسیار کمیاب باشد، ممکن است در یک گروه دارای مواجهه خاص به اندازه‌ای شایع باشد که امکان گردآوری تعداد کافی از موردها فراهم شود. به عنوان مثال میزان بروز سالیانه مزوتلیوما در بین جمعیت عام نزدیک به ۸/۲۴ در یک میلیون است. بنابراین در یک مطالعه هم‌گروهی که شامل ۲۰،۰۰۰ نفر باشد نیز شناسایی هر مورد از این بیماری در مدت ۵ سال پیگیری نامحتمل است؛ اما چون بروز مزوتلیوما در بین کسانی که با پنبه نسوز سروکار دارند نسبتاً شایع است، در یک مطالعه هم‌گروهی شامل ۲۰،۰۰۰ نفر از کارگرانی که با این ماده کار می‌کنند، ممکن است تعداد کافی از موردهای بیماری بروز کند که برای جستجوی ارتباط‌های مواجهه-بیماری کافی باشد. به این ترتیب هر چند مطالعه هم‌گروهی به طور کلی برای ارزیابی بیماری‌های نادر مناسب نیست اما اگر پیامد مورد نظر در گروه مواجهه کمیاب نباشد، می‌توان از یک طرح هم‌گروهی با کارایی مناسب استفاده کرد.

نخستین الزام معتبر بودن مطالعه هم‌گروهی توانایی دستیابی کامل و دقیق به اطلاعات مربوط به شرکت‌کنندگان، به ویژه از نظر اثبات داده‌های مربوط به مواجهه و پیامد مورد نظر است، به همین دلیل بیشتر اوقات مطالعه‌های هم‌گروهی در بین گروه‌هایی که به طور خاص برگزیده شده‌اند اجرا می‌شوند. گروه‌های متفاوتی که برای مطالعه‌های هم‌گروهی به عنوان هدف هستند، می‌توانند شامل: کارگران یک کارخانه، افراد یک اتحادیه، مستمری‌بگیران، دانشجویان یک دانشکده خاص یا شرکت‌کنندگان در برنامه‌های مراقبت پزشکی باشند. هر یک از این گروه‌ها مزیت‌های خاصی برای پژوهشگر دارند، که این مزایا از در دسترس بودن نشانی‌های به هنگام سالیانه تا سازوکارهای مربوط به پیگیری دوره‌ای، تدارک گزارش‌های کامل پزشکی و شغلی را شامل می‌شوند. علت انتخاب این گروه‌ها بالای مواجهه نیست بلکه امکان پیگیری و مطالعه بهتر، منجر به این انتخاب‌ها می‌شود.

به این ترتیب انتخاب یک گروه ویژه به عنوان جمعیت هم‌گروه به فرضیه مورد بررسی و جنبه‌های ویژه طرح مطالعه مربوط می‌شود. به عنوان مثال برای ارزیابی رابطه احتمالی بین مواجهه با حلال و خطر ایجاد سرطان، یک مطالعه هم‌گروهی روی تعدادی از افراد شاغل در کارخانه‌هایی که حلال به کار می‌برند و گزارش‌های موجود درباره مواجهه شغلی آنها در دسترس است، قابل طراحی است. در واقع چنین مطالعه‌ای به روی کارگران پلاستیک، لینولیم و لاستیک به روی شش کارخانه عمده لاستیک‌سازی در ایالات متحده انجام گرفت. در این مطالعه پیامدهای مربوط به مواجهه صنعتی و خطر سرطان بررسی شد [۱۴]. برای آنکه این نوع پژوهش درباره عامل‌های خطر مهم در بیماری‌های نسبتاً مزمن و شایع کارایی داشته باشد، بهترین گزینش برای جمعیت تحت مطالعه یک هم‌گروه عمومی است که بتوان آن را مورد بررسی قرار داد و وضعیت مواجهه ابتدایی این افراد با عامل‌های خطر را اثبات کرد و سپس به طور متناوب معاینه کرد تا پیامدهای بعدی در آنها تشخیص داده شود. یک طرح کلاسیک از این نوع مطالعه، مطالعه بیماری‌های قلبی در شهر فرامینگهام ایالات متحده است. در این

مطالعه یک هم‌گروه آینده‌نگر با جمعیت نزدیک به ۵،۱۰۰ نفر ساکنان بالاتر از ۳۰ سال ایالت ماساچوست معاینه شدند. شهر فرامینگهام به خاطر آن که جمعیت آن، ویژگی خاص در رابطه با مواجهه یا پیامد داشته باشند که امکان بررسی بیشتری را فراهم کنند به‌عنوان محل مطالعه انتخاب نشد، بلکه از آن‌رو که برای پژوهشگران امکان شناسایی و پیگیری شرکت‌کنندگان به مدت چند سال در این شهر فراهم بود، انتخاب شد. جمعیت این شهر ثابت و صنایع آن مشخص بودند. یک بیمارستان عمده و بزرگ در این شهر بخش بسیار بزرگی از جمعیت را تحت پوشش قرار می‌داد و فهرست‌هایی که پیگیری شرکت‌کنندگان را آسان سازد به‌طور سالانه و به‌نگام فراهم بود. این‌ها مواردی هستند که در اجرای موفقیت‌آمیز یک مطالعه هم‌گروهی آینده‌نگر با جمعیت عمومی که امکان ارزیابی اثرهای طیف‌گوناگونی از عوامل که برای بسیاری از بیماری‌ها عامل خطر به‌حساب می‌آیند. در این مطالعه امکان بررسی بیماری سرخرگ تاجی قلب، بیماری روماتیسم قلبی، نارسایی قلب، آنژین صدری و برخی بیماری‌های چشم فراهم شد [۱۶] و [۱۷].

انتخاب گروه مورد مقایسه (جمعیت بدون مواجهه)

پس از تعیین منابع، انتخاب آزمودنی‌های دارای مواجهه، گام مهم بعدی انتخاب یک گروه مقایسه‌ای مناسب از افراد بدون مواجهه است. انتخاب گروه مقایسه در مطالعه هم‌گروهی به‌اندازه انتخاب گروه شاهد در مطالعه مورد-شاهدی دشوار و مهم است. اصل عمده زیربنای این تصمیم‌گیری آن است که باید گروه‌های مقایسه هر چه بیشتر از نظر همه عامل‌هایی که می‌توانند با بیماری رابطه داشته باشند، مشابه با گروه‌های دارای مواجهه باشند (به‌جز از نظر عامل مواجهه موردبررسی)، به‌طوری که اگر واقعاً همبستگی بین مواجهه و بیماری وجود نداشته باشد، شاخص‌های بیماری در دو جمعیت مشابه شوند. نکته مهم دیگر، مطمئن بودن از این موضوع است که اطلاعات به‌دست‌آمده از گروه بدون مواجهه برای مقایسه با جمعیت دارای مواجهه کافی است.

در مطالعه‌های هم‌گروهی که در آن یک هم‌گروه عمومی شرکت داده شده و سپس اعضای آن برحسب رده‌های گوناگون مواجهه طبقه‌بندی می‌شوند می‌توان از گروه مقایسه‌ای درون هم‌گروه بهره جست. یعنی تجربه با یک مواجهه خاص در یک زیرگروه از اعضای هم‌گروه طبقه‌بندی و با تجربه زیرگروه دیگری از اعضای همان گروه که بدون مواجهه بوده یا مواجهه‌ای با شدت متفاوت داشته‌اند مقایسه می‌شود. به‌عنوان مثال در مطالعه هم‌گروهی در پزشکان انگلیسی که در سال ۱۹۵۰ توسط دال^۷ و هیل^۸ انجام شد، میرایی از سرطان ریه در بین کسانی که هرگز سیگار نکشیده بودند با میرایی این بیماری در بین همه سیگاری‌ها و هم‌چنین کسانی که تعداد متفاوتی سیگار می‌کشیدند مقایسه شد. پژوهشگران دریافتند که میزان‌های مرگ از سرطان ریه در بین سیگاری‌ها در مقایسه با غیرسیگاری‌ها افزایش یافته و هم‌چنین افزایش چشمگیر و مداوم از میرایی سرطان ریه با افزایش تعداد سیگارهای کشیده شده مشاهده می‌شود [۱۸]. به همین ترتیب در مطالعه فرامینگهام سطح پایه سرم کلسترول خون و فشارخون سیستولیک و دیاستولیک تعیین شد و شرکت‌کنندگان برحسب صدک‌های

۱. Doll

۲. Hill

مواجهه تقسیم‌بندی شدند. در ادامه مطالعه میزان‌های بیماری سرخرگ تاجی قلب در هر یک از سطح‌ها محاسبه و مقایسه شد. توجه به این نکته اهمیت دارد که وقتی چند عامل خطر به‌طور هم‌زمان در نظر گرفته شوند باید گروه بدون مواجهه شامل اشخاصی باشند که هیچ‌یک از عامل‌های خطر مورد ارزیابی را نداشته باشند [۱۹].

شناسایی آن بخش از هم‌گروه که می‌تواند با اطمینان بدون مواجهه فرض شوند و برای مقایسه به کار روند در مطالعه‌های هم‌گروهی که برای بررسی یک گروه مواجهه خاص مانند مجموعه‌های کارگری یا شرایط اقلیمی ویژه طراحی می‌شود، معمولاً بسیار مشکل است. در این حالت از گروه‌های مقایسه‌ای بیرونی استفاده می‌شود. یکی از گروه‌های مقایسه‌ای بیرونی، جمعیت عام منطقه‌ای است که هم‌گروه در آنجا زندگی می‌کنند. نسبت بیماری مشاهده‌شده در گروه مطالعه با برآورد نسبت بیماری در جمعیت عام در طول پیگیری هم‌گروه مقایسه می‌شود. به‌عنوان مثال برای ارزیابی خطرهای بالقوه مربوط به کار کردن در کارخانه لاستیک‌سازی میرایی مشاهده‌شده در کارگران کارخانه لاستیک‌سازی را با میرایی جمعیت عام آن منطقه که از لحاظ سن و جنس و سایر عوامل خطر بالقوه مهم تعدیل شده‌اند، مقایسه شد [۲۰].

البته مقایسه با میزان‌های جمعیتی فقط برای پیامدهایی امکان‌پذیر است که میزان‌های مربوط به جمعیت در دسترس باشند. میزان بروز سرطان، میرایی از سرطان یا موقعیت‌های مشابه که بتوان داده‌های مربوط به پیامد را به دست آورد، مثال‌هایی از این قبیل هستند. گذشته از این به‌کارگیری جمعیت عام به‌عنوان شاخصی مطمئن از تجربه گروه بدون مواجهه با این فرض است که فقط بخش کوچکی از آن جمعیت در عمل و از طریق هر نوع منبع عامل خطر تحت بررسی در معرض مواجهه است. در واقع بسته به دامنه آن بخش از جمعیت عام که دارای مواجهه است، همبستگی حقیقی بین مواجهه و بیماری کم برآورد می‌شود.

عیب بزرگ به‌کارگیری جمعیت عام به‌عنوان گروه مقایسه آن است که ممکن است افراد جمعیت عام به‌طور مستقیم قابل‌مقایسه با افراد هم‌گروه نباشند، زیرا اولاً کسانی که شغلی دارند به‌طور متوسط سالم‌تر از کسانی هستند که شغلی ندارند و چون جمعیت عام، افرادی که به علت بیماری نمی‌توانند کار بکنند و افراد شاغل را شامل می‌شود، میزان‌های بیماری و مرگ در بین جمعیت عام تقریباً همیشه بیش از میزان‌های بیماری و مرگ در اشخاص مشمول قانون کار است. به این پدیده، اثر کارگر سالم گفته می‌شود.

امکان دارد جمعیت مرجع و جمعیت هم‌گروه از لحاظ برخی عوامل نظیر خصوصیت‌های جمعیت‌شناسی و جغرافیایی مشابه باشند اما از نظر سایر عامل‌های خطر بیماری تفاوت کلی داشته باشند (از جمله سیگار کشیدن، رژیم غذایی، مصرف نوشابه‌های الکلی). چون این اطلاعات درباره افراد جمعیت عام در دسترس نیست، احتمال دارد تفاوت‌های مشاهده‌شده ناشی از مهار نشدن عوامل مخدوش‌کننده باشد.

یک راه دیگر برای استفاده از میزان‌های بیماری در جمعیت عام به‌عنوان گروه مقایسه، مقایسه تجربه با مواجهه خاص هم‌گروه با یک هم‌گروه مشابه از نظر خصوصیت‌های جمعیت‌شناختی ولی بدون مواجهه است. در مطالعه‌های مربوط به مواجهه‌های شغلی چنین گروهی شامل کارگران انواع شغل‌های متفاوت است که در محل‌های متفاوت یک مرکز صنعتی انجام‌وظیفه می‌کنند. به‌عنوان مثال، میزان‌های بیماری در گروه بدون مواجهه کارگران بخش اداری یک کارخانه را می‌توان با میزان‌های بیماری گروه کارگران بخش تولید در همان کارخانه مقایسه کرد. یکی از این مطالعات هم‌گروهی برای بررسی مواجهه خلبانان پرواز بمب‌های شیمیایی در

جنگ‌های ویتنام با گروهی از کارکنان نیروی هوایی شاغل در بخش حمل‌ونقل هوایی در همان منطقه از ویتنام و در همان دوره زمانی است [۲۱]. هم‌چنین برای ارزیابی خطر مواجهه درازمدت با مقدارهای کم‌پرتوی یون‌ساز در رادیولوژیست‌ها، این اشخاص با پزشکان متخصص بیماری‌های داخلی مقایسه شدند [۲۲]. چون متخصصان بیماری‌های داخلی از نظر خصوصیت‌های جمعیت‌شناسی، توجه به سلامتی و بهره‌گیری از مراقبت‌های پزشکی همانند رادیولوژیست‌ها هستند، این مقایسه بسیار موجه به نظر می‌رسد. در مطالعه‌ای درباره مواجهه با پنبه نسوز و ارتباط آن با سرطان ریه، هم‌گروه از کارگران دارای مواجهه با پنبه نسوز که در صنعت ریسندگی شاغل بودند با کارگران معمولی ریسندگی با پنبه مقایسه شدند. این دو گروه وظایف شغلی مشابهی دارند، خصوصیت‌های اقتصادی اجتماعی مشترک دارند ولی از نظر شغلی گروه دوم در معرض مواجهه با پنبه نسوز نیستند [۲۳]. مزیت استفاده از هم‌گروه در مقایسه با استفاده از گروه عام آن است که می‌توان گروه را سازگارتر و مشابه با گروه دارای مواجهه انتخاب کرد. از این گذشته اطلاعات مربوط به عامل‌های مخدوش‌کننده بالقوه را می‌توان در مورد افراد مورد مطالعه به دست آورد و باقیمانده تفاوت‌ها را هنگام تجزیه و تحلیل مهار کرد.

در بسیاری از مطالعات هم‌گروهی، به‌ویژه هنگامی که به‌ظاهر هیچ گروهی به‌تنهایی شباهت کافی با اشخاص دارای مواجهه ندارد، اطمینانی برای معتبر بودن مقایسه وجود ندارد. در این موارد داشتن چند گروه مقایسه‌ای سودمند است. در چنین شرایطی اگر همبستگی مشابهی از گروه‌های مختلف حاصل شود، نتیجه مطالعه موجه‌تر است. به‌عنوان مثال در مطالعه مربوط به پنبه نسوز که پیش‌از این شرح داده شد، علاوه بر مقایسه کارگران پنبه نسوز با کارگران کارخانه‌های ریسندگی پنبه معمولی، میزان‌های میرایی با جمعیت مردان سفیدپوست ایالات متحده هم مقایسه شد. کارگران شاغل در بخش‌های ریسندگی پنبه نسوز دارای میزان‌های میرایی بالاتری از همه علت‌ها، سرطان ریه، بیماری‌های تنفسی دیگر و هم‌چنین بیماری پرفشاری خون نسبت به کارگران ریسندگی پنبه معمولی و نسبت به جمعیت عام بودند. به‌این ترتیب میزان میرایی کارگران پنبه نسوز نه تنها بیش از جمعیت عام بود بلکه از گروه مردانی هم که کار مشابه داشتند ولی پنبه نسوز در محیط کار آن‌ها نبود، به‌طور چشم‌گیری بیشتر بود. سازگار بودن نتیجه این دو مقایسه باعث تقویت شدن این نظریه که کارگران کارخانه‌های تولید پنبه نسوز در خطر ایجاد بیماری در رابطه با شغل هستند، شد. هم‌چنین در ارزیابی پیامدهای نامطلوب مرتبط با مصرف قرص‌های خوراکی جلوگیری از بارداری (OCP) انتخاب یک گروه مقایسه متناسب اهمیت فراوانی دارد. برای مقایسه گروهی باید انتخاب شود که نوعی وسیله جلوگیری از بارداری به‌جز قرص خوراکی به کار می‌برند. زنانی که از هیچ روش جلوگیری از بارداری بهره نمی‌گیرند ممکن است با استفاده‌کنندگان هر نوع وسایل جلوگیری از بارداری، از نظر توانایی باردارشدن، تمایل به بارداری و ماهیت رفتار جنسی متفاوت باشند. از سوی دیگر زنانی که از اشکال گوناگون روش‌های جلوگیری از بارداری استفاده می‌کنند احتمالاً از زنانی که مصرف‌کننده قرص‌های خوراکی هستند از نظر مذهب، وضعیت اقتصادی اجتماعی، گروه نژادی و سایر عوامل مؤثر بر سلامتی و شیوه زندگی متفاوت هستند. به‌این ترتیب ممکن است هیچ نوع گروه مقایسه‌ای به‌وضوح برتر نباشد. اطلاعات مربوط به نقش قرص‌های خوراکی جلوگیری از بارداری و بیماری را می‌توان با مقایسه نتایج به‌دست‌آمده از مطالعه‌های هم‌گروهی دارای هم‌گروه‌هایی از گروه‌های مختلف به نحو مطلوبی بررسی کرد [۲۴].

منابع داده‌ها: هنگام طراحی هر نوع مطالعه هم‌گروهی، توجه اصلی بر در دسترس بودن اطلاعات دقیق و کامل است که امکان طبقه‌بندی تمام افراد عضو هم‌گروه (برحسب داشتن مواجهه با عامل‌های خطر موردتحقیق یا ایجاد هر نوع پیامد موردنظر) فراهم شود. اطلاعات درباره مواجهه را می‌توان از چند منبع از جمله پرونده‌هایی که به‌طور مستقل از مطالعه گردآوری شده (مانند پرونده‌های پزشکی یا اشتغال)، اطلاعات به‌دست‌آمده از خود مطالعه (از طریق پرسشگری یا پرسش‌نامه)، داده‌های به‌دست‌آمده از معاینه پزشکی یا دیگر آزمایش‌هایی که برای شرکت‌کنندگان انجام شده و اندازه‌گیری‌های مستقیم در محل زندگی یا کار افراد هم‌گروه به دست آورد. اطلاعات درباره پیامد را هم می‌توان از پرونده‌های از پیش موجود از جمله گواهی‌های مرگ یا پرونده‌های پزشکی، پرسشنامه یا معاینه‌های بدنی جمع‌آوری کرد. هر یک از این منابع داده‌ها معایب و مزایای خاص خود را دارا است که به هنگام طراحی یا تفسیر مطالعه باید آن‌ها را در نظر گرفت.

۳-۴ انواع مطالعات مداخله‌ای

در نوع مطالعه هم‌گروهی، گروه‌های مواجهه‌دار با گروه‌های بدون مواجهه با یکدیگر مقایسه می‌شوند ولی عامل مواجهه در مطالعات کارآزمایی بالینی اغلب مداخله است؛ اما در مطالعات هم‌گروهی (مشاهده‌ای) معمولاً عامل موردبررسی، عاملی است که افراد با آن مواجهه دارند. درک اختلاف این دو طرح از نظر تفسیر نتایج اهمیت زیادی دارد. در مطالعات هم‌گروهی (مشاهده‌ای) برخلاف مشاهدات تجربی (کارآزمایی‌های بالینی تصادفی شده) مواجهه مورد مطالعه به‌صورت تصادفی و با اراده محققان اختصاص نیافته‌اند. به همین علت هنگامی که در یک مطالعه تصادفی نشده رابطه‌ای بین مواجهه و بیماری به دست می‌آید ممکن است این رابطه بر اثر انتخاب تصادفی نشده (جهت‌دار) افراد مطالعه به وجود آمده باشد.

تجربه‌ها یا مداخله‌ها، قطعی‌ترین مدارک را برای بررسی یک رابطه سببی فراهم می‌سازند. مطالعات پزشکی نیز از این امر مستثنا نیستند. مطالعات مداخله‌ای برحسب وجود یا عدم وجود گروه‌های مقایسه‌ای (موازی و یا هم‌زمان)، انجام مطالعه به روی بیماران یا افراد سالم و انتساب تصادفی آزمودنی‌ها با یکدیگر متفاوت هستند.

۳-۴-۱ مطالعات تجربی

در مطالعات تجربی، افراد تحت مطالعه تنها هنگامی مداخله موردنظر را دریافت می‌کنند و یا در معرض مواجهه معین و از قبل قرار می‌گیرند که محقق، بر اساس یک پروتکل دقیق و احتمالاً با استفاده از روش‌های انتساب تصادفی، این مداخلات را به هرکدام از اعضا اختصاص دهد. در این مطالعات، تجویز هرکدام از مداخلات برای افراد تحت مطالعه، بر اساس نیاز یا انتخاب افراد تحت مطالعه و یا شرایط محیطی ویژه و غیره صورت نمی‌گیرد بلکه بر اساس پروتکل تحقیق، محققان مداخلات را به افراد تحت مطالعه اختصاص می‌دهند. برای مثال هنگامی که پزشک برای درمان یک بیماری، نظیر سردرد برای بیمارانی که از لحاظ مالی توانمند هستند، برای کسب اطمینان بیشتر دارویی با مارک تجاری تجویز کند و به بیماران فقیر، به علت صرفه‌جویی در هزینه‌های درمان همان دارو را با نام ژنریک تجویز کند این مطالعه تجربی نیست. زیرا تعیین تجویز دارو به بیماران بر اساس سلیقه و قضاوت‌های شخصی صورت گرفته است. برای انجام یک مطالعه تجربی، لازم است پزشک بر

اساس یک پروتکل مدون، دارو را به بیماران تجویز کند تا هرگونه تغییرات بالقوه که ناشی از عوامل دیگر است، خنثی شود. درواقع اختصاص درمان‌ها در مطالعات تجربی به‌گونه‌ای است که از اثر تغییرات سایر عوامل بر مقایسه‌ها تا حد امکان جلوگیری شود. اگر در انتساب مداخله قضاوت‌های شخصی، شرایط بالینی بیماران و سلیقه مجریان مطالعه تأثیرگذار باشد، باز هم مقایسه روش‌ها با یکدیگر از طریق طراحی یک مطالعه نیمه تجربی یا غیرتجربی امکان‌پذیر است.

ملاحظات اخلاقی زیادی در طرح و اجرای مطالعات تجربی وجود دارد. علت این است که هدف مطالعه نحوه انتساب‌ها را مشخص می‌کند و نیازهای بالینی بیماران در نظر گرفته نمی‌شود. مطالعات تجربی از لحاظ اخلاقی تنها زمانی قابل اجرا هستند که انجام آن‌ها برای بیماران، مفید تشخیص داده شود. به‌طور خاص باید دلایل علمی قابل‌قبولی وجود داشته باشد که نشان دهد دو یا چند درمان موجود که در پروتکل مطالعه پیشنهاد شده‌اند بهترین درمان‌های موجود هستند. به علت این نکته، هرکدام از درمان‌ها یا مداخله‌های پیشنهاد شده باید دارای حداقل‌هایی از لحاظ قدرت درمان یا پیش‌گیری باشند. این محدودیت‌ها، برای مثال باعث می‌شوند اکثر مطالعات سم‌شناسی از نوع غیرتجربی طراحی شوند. با رعایت تمامی این مسائل رضایت کاملاً آگاهانه افراد تحت مطالعه، مجوز اصلی برای انجام مطالعات تجربی است.

باوجود این محدودیت‌ها، مطالعات زیادی از این نوع در زمینه‌های مختلف پزشکی طراحی و اجرا می‌شود. اکثر این مطالعات را تحت عنوان کارآزمایی‌های بالینی طبقه‌بندی می‌کنند. کارآزمایی‌های بالینی مطالعاتی هستند که برای مقایسه درمان‌های مختلف در بیماران طراحی می‌شوند. مطالعات تجربی که باهدف پیشگیری اولیه طراحی می‌شوند، کارآزمایی‌های میدانی یا مداخله‌های اجتماعی خوانده می‌شوند.

۳-۴-۲ کارآزمایی (مطالعه یا پژوهش) بالینی

کارآزمایی بالینی، مطالعه‌ای است که اعضای تحت بررسی را بیماران تشکیل می‌دهند. هدف اصلی هر کارآزمایی بالینی، ارزیابی یک درمان جدید برای بیماری و یا یافتن روش‌های مناسب برای پیشگیری از عوارض یک بیماری نظیر مرگ یا ناتوانی جسمی است. یک تعریف جامع از کارآزمایی‌های بالینی می‌تواند به این صورت زیر باشد.

”کارآزمایی بالینی عبارت است از: هرگونه پژوهش بر روی انسان‌ها که به‌منظور کشف، یا تأیید اثرات بالینی، دارویی و یا سایر اثرات فارماکودینامیکی یک فرآورده تحقیقاتی و یا به‌منظور شناسایی هرگونه واکنش نامطلوب یک فرآورده تحقیقاتی و یا مطالعه جذب، توزیع، سوخت‌وساز و دفع یک فرآورده تحقیقاتی باهدف اثبات بی‌خطری و یا تأثیر آن صورت می‌گیرد.“

هدف محققان در کارآزمایی‌های بالینی، آزمایش روش‌های درمان (تحت عنوان مداخله) بر روی افراد تحت مطالعه و مشاهده اثرات آن است. مهم‌ترین نقطه قوت کارآزمایی‌ها نسبت به مطالعات مشاهده‌ای، قابلیت انجام استنباط‌های سببی است. در عمل اختصاص تصادفی مداخله، می‌تواند باعث حذف اثر متغیرهای مخدوش‌کننده و کور کردن باعث حذف اثرات احتمالی ناشی از سوگرایی در قضاوت می‌شود.

با این حال در حالت کلی، کارآزمایی‌های بالینی گران، وقت‌گیر و گاهی اوقات می‌توانند برای افراد تحت

مطالعه مخاطره‌آمیز باشند و معمولاً مطالعات بالینی تنها به یک سؤال پژوهش بسیار اختصاصی پاسخ می‌دهند. به همین علت، پژوهش‌های بالینی تنها در مواردی که اهداف و سؤالات پژوهش به حد معینی از تکامل رسیده باشد، به کار گرفته می‌شود [۲۵].

اکنون مراحل مختلف برای طراحی یک مطالعه آزمایشی را مورد بررسی قرار می‌دهیم.

۳-۴-۲-۱ انتخاب بیماران

در هر کارآزمایی بالینی به تعریف دقیقی از بیماران واجد شرایط برای ورود به مطالعه نیاز داریم. هدف اصلی این است که مطمئن شویم بیمارانی که وارد کارآزمایی می‌شوند، معرف طبقه‌ای از بیماران هستند که در آینده یافته‌های کارآزمایی در مورد آنان به کار گرفته می‌شوند. به علاوه مطلوب این است که درمان با داروی جدید را در آن دسته از بیماران که احتمالاً بیشترین نفع را از دارو می‌برند متمرکز کنیم. در عین حال، ورود بیماران نباید به حدی محدود شده و کارآزمایی آن قدر کوچک بماند که یافته‌ها آن قابلیت تعمیم خود را از دست دهد. جنبه‌های اساسی که توجه به آن‌ها ضروری است عبارت‌اند از:

- ♦ منبع انتخاب بیمار
- ♦ وضعیت بیماری تحت مطالعه
- ♦ معیارهای خاص خروج بیماران

موضوع انتخاب منبع بیماران باید با دقت ویژه مورد توجه قرار گیرد. برای مثال در مطالعه بیماری افسردگی اگر به سراغ بیماران بستری شده در بیمارستان برویم به گروهی غیرمعمول دست می‌یابیم. این بیماران بیشتر موارد مزمن و شدید را شامل می‌شوند. در حالی که داروهای ضدافسردگی تحت مطالعه را معمولاً گروه بزرگ‌تری از بیماران افسرده تحت مراقبت پزشکان مصرف می‌کنند.

تعیین معیارهای ورود و خروج:

در تعیین معیارهای ورود و خروج هر کارآزمایی بالینی، باید میزان کلی پیامد مورد نظر، میزان تقریبی اثربخشی مداخله فعال، قابل تعمیم بودن یافته‌ها، دسترسی آسان به افراد مورد مطالعه، عوارض جانبی احتمالی و امکان پیگیری افراد نمونه مورد توجه قرار گیرند.

اگر پیامد مورد نظر نظیر سرطان پستان، نادر باشد، معمولاً لازم است افراد تحت مطالعه تنها از میان افراد پرمخاطره انتخاب شوند تا حجم نمونه و مدت پیگیری در حد قابل قبولی قرار گیرند؛ اما از طرف دیگر محدود کردن معیار ورود، به افراد پرمخاطره، باعث کاسته شدن قابلیت تعمیم نتایج و همچنین مشکل شدن نحوه دسترسی به اعضای نمونه می‌شود.

پژوهشگران باید برآوردهای قابل اعتمادی از میزان کلی پیامد مورد مطالعه در دست داشته باشند. این نتایج ممکن است از آمارهای جاری، مطالعات طولی و یا میزان پیامد در گروه بدون مداخله (گروه کنترل) فراهم شوند. محدود شدن به بررسی افراد پرمخاطره می‌تواند باعث کاهش اعضای نمونه شود. اگر عوامل خطر پیامد از قبل مشخص شده داشته باشند، معیارهای ورود به مطالعه را می‌توان بر اساس دارا بودن حداقل برآوردی از مواجهه با خطر پایه‌ریزی کرد. برای مثال، در مطالعه بالینی درباره اثر تاموکسیفن در پیشگیری از سرطان پستان، ترکیبی از عوامل خطر در ابتدا به این سرطان به عنوان یکی از معیارهای ورود به مطالعه به کار رفت. تنها زنانی

وارد این مطالعه می‌شدند که به مدت پنج سال، خطر ابتلا به سرطان بالاتر از ۱/۷٪ داشتند [۲۶]. راه دیگر در افزایش دسترسی به اعضای نمونه میزان پیامدها، محدود کردن مطالعه به بررسی افراد مبتلا است. به‌عنوان مثال، در مطالعه‌ای که به‌منظور بررسی افزایش میزان هورمون‌های استروژن و پروژسترون در کاهش میزان موارد ابتلا به بیماری عروق کرونر (CHD)، از ۲۷۶۳ زنی که قبلاً مبتلا به این عارضه بودند استفاده شد. این روش‌ها عملی‌تر و کم‌هزینه‌تر از مطالعه‌ای مشابه با زن‌هایی که دچار عارضه CHD نبودند (که حجم نمونه‌ای بالغ بر ۲۵،۰۰۰ نفر لازم دارد) است.

محدود کردن معیارهای ورود، به گروه‌های پرخطر دارای دو محدودیت عمده است. اول اینکه نتایج به‌دست‌آمده از مطالعه بر روی گروه‌های پرخطر، ممکن است قابل‌تعمیم به گروه‌های کم‌خطر نباشد. برای مثال، در مطالعه استروژن و پروژسترون که تنها زنان دارای عارضه CHD بررسی شدند، تضمینی برای کارایی این روش در زنان بدون عارضه CHD وجود ندارد. به‌علاوه جمع‌آوری مدارک و اطلاعات کافی در مورد گروه‌های پرخطر، نظیر سابقه وجود CHD، ممکن است نیاز به جمع‌آوری داده‌ها، اطلاعات و یا اندازه‌گیری‌های متعددی داشته باشد، که باعث پیچیده، وقت‌گیر و پرهزینه شدن مطالعه شود.

در تعیین معیارهای خروج یک مطالعه بالینی، باید به‌گونه‌ای تصمیم‌گیری شود که تا حد امکان از حذف غیرضروری بیماران از مطالعه جلوگیری شود. زیرا حذف غیرضروری بیماران از مطالعه باعث کاهش قابلیت‌تعمیم مطالعه و پیچیده و پرهزینه شدن جمع‌آوری نمونه‌ها می‌شود. دلایل عمده حذف بیماران از مطالعه در جدول ۳-۱ آورده شده است.

۳-۲-۴-۲ اندازه‌گیری متغیرهای پایه

به‌طور کلی در هر کارآزمایی بالینی، اطلاعات مربوط به چهار دسته از متغیرها باید جمع‌آوری شود: متغیرهای مربوط به شناسایی افراد مطالعه؛ نظیر نام و نام خانوادگی، شماره تلفن تماس، شماره شناسنامه، آدرس منزل و غیره، متغیرهای دموگرافیک نظیر سن، جنس و غیره، متغیرهای مخدوش‌کننده، متغیر پاسخ و سطوح مختلف مداخله.

اندازه‌گیری‌هایی که روی بیماران انجام می‌شود باید دقیق و تکرارپذیر بوده و به‌ویژه به مسئول اندازه‌گیری یا مشاهده‌گر بستگی نداشته باشد. متأسفانه بسیاری از ارزیابی‌ها کپی هستند و به وجود مشاهده‌گران کارآزموده نیاز است. مثلاً، بیماری روانی اساساً بر مبنای تشخیص بالینی ارزیابی می‌شود. مسئله اساسی این است که اطلاعات گردآوری‌شده ساختارمند باشند. برای نمونه، در مورد بیماری افسردگی دستیابی به یک عبارت کلی و پراکنده در مورد وضعیت بیمار مناسب نیست. در عوض، روش‌های خاص تشخیص افسردگی (مثلاً مقیاس اندازه‌گیری همیلتون) بسط یافتند که بکار بردن آن با انجام مصاحبه ساختاری منظم برای ثبت اطلاعات به وجود می‌آید. در موارد مشکل برای ارزیابی بالینی، تشکیل گروهی مرکب از دو یا سه نفر مشاهده‌گر که وظیفه اصلی آن‌ها توافق روی یک نظر مشترک است (پس از ارزیابی مستقل اولیه و سپس مشاوره با یکدیگر باشد) بسیار مفید است. این موضوع در مورد ارزیابی‌هایی که نیازی به حضور بیمار ندارد، مثلاً تفسیر CT اسکن بسیار مناسب و عملی است.

جدول ۳-۱: دلایل حذف افراد تحت مطالعه در کارآزمایی‌های بالینی

علت
درمان فعال انجام شده دارای اثر جانبی مخرب غیرمنتظره است. خطر غیرمنتظره به علت دریافت دارونما، افراد مطالعه را تهدید می‌کند.
درمان فعال مؤثر نیست زیرا فرد مورد مطالعه: - اساساً در معرض خطر ابتلا قرار ندارد. - به بیماری خاصی مبتلا است که اثربخشی درمان مورد مطالعه را از بین می‌برد. - داروهایی مصرف می‌کند که با درمان فعال، مداخله دارویی دارد. - به‌طور مستمر و دقیق از درمان فعال یا غیرفعال تبعیت نمی‌کند.
ناتمام ماندن پیگیری
مسائل و مشکلات اساسی بیمار برای شرکت یا ادامه همکاری در مطالعه

در بعضی از بیماری‌ها امکان ندارد اثرات درمان را ارزیابی کنیم مگر آنکه عقیده بیمار را جویا شویم. مثلاً تسکین درد در کارآزمایی داروهای ضد روماتیسم نمونه‌ای از این موضوع است.

ارزیابی پاسخ بیمار: پس از شروع کارآزمایی، پیشرفت درمانی هر بیمار به انجام ارزیابی عینی، دقیق و پایدار نیاز دارد تا بر اساس آن بتوان مزایای نسبی درمان‌ها را ارزیابی کرد. بنابراین، روش‌های ارزیابی و ثبت پیشرفت وضعیت بیماران به تعریف دقیقی در پروپوزال مطالعه نیاز دارد. در حقیقت برای ارزیابی بیمار در یک کارآزمایی بالینی، یادداشت‌های موردی معمولی کلاً نامناسب هستند. زیرا معمولاً بسیار مبهم، نامنظم، ناپایدار و ذهنی نگارش یافته‌اند.

بیماران حاضر در کارآزمایی به دلایل زیر ارزیابی می‌شوند:

- ۱- ارزیابی پایه پیش از آغاز درمان
- ۲- معیارهای اصلی واکنش بیمار
- ۳- معیارهای ثانویه مانند عوارض جانبی
- ۴- سایر جنبه‌های پایش بیمار

در هر کدام از این چهار مرحله برنامه‌ریزی دقیقی در مورد جمع‌آوری اطلاعات و ثبت داده‌های در فرم‌ها یا طرح‌ریزی فرم‌های جدید لازم است. درعین حال ابتدا باید تصمیم گرفت که کدام جنبه‌ها را اندازه‌گیری یا مشاهده کرد.

معیارهای اصلی پاسخ بیمار: کارآزمایی‌های بالینی ممکن است به مشاهدات گسترده‌ای برای هر بیمار نیاز داشته باشد، این موضوع ممکن است تفسیر داده‌ها را دشوار سازد. از این رو پیش از آن که کارآزمایی شروع شود باید در خصوص اهمیت نسبی اندازه‌گیرهای گوناگون راهکارهایی ارائه شود. در حقیقت اگر یک مقیاس مشخص برای بررسی پاسخ بیمار و مقایسه درمان‌ها به کار گرفته شود، کار مفیدی انجام گرفته است.

معیارهای فرعی و عوارض جانبی: پس از آنکه معیار اصلی ارزیابی بیمار به روشنی تعریف شد، جنبه‌های عمده دیگری نیز وجود دارد که محققان تمایل دارند آن‌ها را بررسی کنند. مثلاً در کارآزمایی دارو ارزیابی بی‌خطر

بودن و کارایی دارو از اهمیت برخوردار است. بنابراین نیاز به مقایسه درمان‌ها برای تعیین عوارض جانبی آن‌ها وجود دارد. اگر عوارض جانبی شناخته شده باشند، این موضوع واضح است (مثلاً کاهش ضربان قلب در اثر استعمال بتابلاکر، کاهش افتراقی گلبول سفید در اثر داروهای دارای اثر سمی روی سلول)؛ اما ثبت عوارض جانبی در مورد داروی نسبتاً جدید مشکل‌تر است. در چنین شرایطی باید تأکید زیادی به ارزیابی خود بیمار از عوارض جانبی داشته باشیم. یک روش قراردادی این است که یک فهرست بررسی از تمام علائم و بیماری‌ها را در فواصل معینی تهیه کرده و از بیمار سؤال کنیم. برای مثال، جدول ۳-۲ می‌تواند در یک کارآزمایی از داروهای ضدافسردگی مفید واقع شود:

جدول ۳-۲: نمونه‌ای از فهرست عوارض جانبی برای ارزیابی یک دارو جدید

سوزش دست‌ها	حالت تهوع	سردرد
رعشه	استفراغ	خستگی
تعرق زیاد	سوهاضمه	بی‌خوابی
تورم دست‌ها و پاها	خشکی دهان	گیجی
تشنج	خشکی پوست	اسهال
پرخاشگری	درد مفاصل	یبوست

توجه داشته باشید که این فهرست باید با کلماتی بیان شود که بیماران آن را درک کنند. روش دیگری هم این است که از بیمار سؤال می‌شود هر رویداد نامطلوبی را که تجربه کرده توضیح دهد. این برای هر بیمار رویدادها را روی یک فرم مخصوص درج کرده و برای رده‌بندی بعدی نگه می‌دارند [۲۵].

۳-۲-۴-۳ تصادفی کردن

مرحله بعدی در مطالعات بالینی، انتساب تصادفی گروه‌های مختلف درمانی به بیماران است. در ساده‌ترین طرح، یک گروه درمان فعال و گروه دیگر دارونما دریافت می‌کند. انتساب تصادفی بیماران به هر کدام از درمان‌ها، پایه و اساس انجام آزمون‌های آماری بر روی داده‌ها است.

۳-۲-۴-۳ روش‌های تصادفی کردن انتساب درمان‌ها

روش تصادفی ساده

در هر کارآزمایی تصادفی با دو درمان (A و B) ساده‌ترین روش استفاده از جدول اعداد تصادفی است. جدول ۳-۳ یک نمونه از این جدول‌ها را با ارقام تصادفی ۰ تا ۹ نشان می‌دهد. برای اختصاص تصادفی دو درمان به بیماران باید به این ترتیب عمل کرد.

نقطه‌ای اختیاری از جدول اعداد تصادفی انتخاب کرده و به تعداد موردنیاز از اعداد مندرج در جدول استخراج می‌نماییم. مثلاً امکان دارد دنباله زیر به دست آید.

۹ ۰ ۶ ۲ ۳ ۷ ۰ ۰ ۰ ۹ ۳ و غیره

ارقام ۰ تا ۴ را برای درمان A و ارقام ۵ تا ۹ را برای درمان B در نظر می‌گیریم. در این صورت فهرستی با ترتیب تصادفی زیر به وجود می‌آید:

A B A A A B A A B B A B و غیره

تعمیم این قاعده به تعداد درمان‌های بیشتر بسیار ساده است و مثلاً در مورد سه درمان A، B و C به این صورت عمل می‌کنیم:

ارقام ۱ تا ۳ برای درمان A

ارقام ۴ تا ۶ برای درمان B

ارقام ۷ تا ۹ برای درمان C

و از رقم ۰ صرف‌نظر می‌کنیم. در این صورت اگر توالی

۱ ۰ ۹ ۱ ۲ ۸ ۶ ۶ ۵ ۹ ۸ ۳ ۲ ۶ ۴ ۶ ۱ ۱ و غیره

از جدول اعداد تصادفی استخراج شود، انتساب‌های تصادفی سه درمان به‌صورت زیر می‌شود:

A A B B B A A C C B B B A C - C A و غیره.

حسن این روش ساده بودن انجام آن و غیرقابل‌پیش‌بینی بودن درمان بعدی است؛ اما یکی از مشکلات این روش نابرابر بودن تعداد افراد در هر گروه‌درمانی است. این موضوع هنگامی که از ۲۰ بیمار تحت بررسی ۸ بیمار درمان A و ۱۲ بیمار درمان B دریافت کنند، تفاوت آشکاری نیست؛ اما هنگامی که ۴ بیمار درمان A و ۱۶ بیمار درمان B دریافت کنند، عملاً تا حد زیادی از توان مطالعه کاسته می‌شود. برای اجتناب از این مسئله روش‌های کارآمدتری نظیر تصادفی‌سازی با جایگذاری، بلوک‌های جایگشتی تصادفی و تصادفی‌سازی طبقه‌بندی‌شده که شامل روش سکه اریب و تصادفی‌سازی طبقه‌بندی‌شده، ابداع شده است.

۴-۲-۴-۴ کور کردن

در صورت امکان‌پذیر بودن، کور کردن دوسویه آزمودنی‌ها راه بسیار مناسبی برای بالا بردن اعتبار مطالعه است. در کور کردن دوسویه، بیماران و گروه تحقیق از این موضوع که کدام‌یک از افراد مطالعه چه درمانی دریافت کرده بی‌اطلاع هستند، پس از اتمام مطالعه و جمع‌آوری نتایج، کدهای مربوط به هر بیمار در اختیار گروه تحقیق قرار می‌گیرد. معمولاً در مطالعاتی که یکی از گروه‌ها درمانی دریافت نمی‌کنند کور کردن با تجویز دارونماهایی مشابه داروی اصلی انجام می‌شود.

جدول ۳-۳: جدول اعداد تصادفی

جدول ۳-۳: جدول اعداد تصادفی

۰۳	۴۷	۲۳	۷۳	۸۶	۳۶	۹۶	۲۷	۳۶	۶۱	۴۶	۹۸	۶۳	۷۱	۶۲	۳۳	۲۶	۱۶	۸۰	۴۵	۶۰	۱۱	۱۴	۱۰	۹۵
۹۷	۷۴	۲۴	۶۷	۶۲	۴۲	۸۱	۱۴	۵۷	۲۰	۴۲	۵۳	۳۲	۳۷	۳۲	۲۷	۰۷	۳۶	۰۷	۵۱	۲۴	۵۱	۷۹	۸۹	۷۳
۱۶	۷۶	۶۲	۲۷	۶۶	۵۶	۵۰	۲۶	۷۱	۰۷	۳۲	۹۰	۷۹	۷۸	۵۳	۱۳	۵۵	۳۸	۵۸	۵۹	۸۸	۹۷	۵۴	۱۴	۱۰
۱۲	۵۶	۸۵	۹۹	۲۶	۹۶	۹۶	۶۸	۲۷	۳۱	۰۵	۰۳	۷۲	۹۳	۱۵	۵۷	۱۲	۱۰	۱۴	۲۱	۸۸	۲۶	۴۹	۸۱	۷۶
۵۵	۵۹	۵۶	۳۵	۶۴	۳۸	۵۴	۸۲	۴۶	۲۲	۳۱	۶۲	۴۳	۰۹	۹۰	۰۶	۱۸	۴۴	۳۲	۵۳	۲۳	۸۳	۰۱	۳۰	۳۰
۱۶	۲۲	۷۷	۹۴	۳۹	۴۹	۵۴	۴۳	۵۴	۸۲	۱۷	۳۷	۹۳	۲۳	۷۸	۸۷	۳۵	۲۰	۹۶	۴۳	۸۴	۲۶	۳۴	۹۱	۶۴
۸۴	۴۲	۱۷	۵۳	۳۱	۵۷	۲۴	۵۵	۰۶	۸۸	۷۷	۰۴	۷۴	۴۷	۶۷	۲۱	۷۶	۳۳	۵۰	۳۵	۸۳	۹۲	۱۲	۰۶	۷۶
۶۳	۰۱	۶۳	۷۸	۵۹	۱۶	۹۵	۵۵	۶۷	۱۹	۹۸	۱۰	۵۰	۷۱	۷۵	۱۲	۸۶	۷۳	۵۸	۰۷	۴۴	۳۹	۵۲	۳۸	۷۹
۳۳	۲۱	۱۲	۳۴	۲۹	۷۸	۶۴	۵۶	۰۷	۸۲	۵۲	۲۲	۰۷	۴۴	۳۸	۱۵	۵۱	۰۰	۱۳	۲۲	۹۹	۶۶	۰۲	۷۹	۵۴
۵۷	۶۰	۸۶	۳۲	۴۴	۰۹	۴۷	۲۷	۹۶	۵۴	۴۹	۱۷	۴۶	۰۹	۶۲	۹۰	۵۲	۸۴	۷۷	۲۷	۰۸	۰۲	۷۳	۴۳	۲۸
۱۸	۱۸	۰۷	۹۲	۴۶	۴۴	۱۷	۱۶	۵۸	۰۹	۷۹	۸۳	۸۶	۱۹	۶۲	۰۶	۷۶	۵۰	۰۳	۱۰	۵۵	۲۳	۶۴	۰۵	۰۵
۲۶	۶۲	۳۸	۹۷	۷۵	۸۴	۱۶	۰۷	۴۴	۹۹	۸۳	۱۱	۴۶	۳۲	۲۴	۲۰	۱۴	۸۵	۸۸	۴۵	۱۰	۹۳	۷۲	۸۸	۷۱
۲۳	۴۲	۴۰	۶۴	۷۴	۸۲	۹۷	۷۷	۷۷	۸۱	۰۷	۴۵	۳۲	۱۴	۰۸	۳۲	۹۸	۹۴	۰۷	۷۲	۹۳	۸۵	۷۹	۱۰	۷۵
۵۲	۳۶	۲۸	۱۹	۹۵	۵۰	۹۲	۲۶	۱۱	۹۷	۰۰	۵۶	۷۶	۳۱	۳۸	۸۰	۲۲	۰۲	۵۳	۵۳	۸۶	۶۰	۴۲	۰۴	۵۳
۳۷	۸۵	۹۴	۳۵	۱۲	۸۳	۳۹	۵۰	۰۸	۳۰	۴۲	۳۴	۰۷	۹۶	۸۸	۵۴	۴۲	۰۶	۸۷	۹۸	۳۵	۸۵	۲۹	۴۸	۳۹
۷۰	۲۹	۱۷	۱۲	۱۳	۴۰	۳۳	۲۰	۳۸	۲۶	۱۳	۸۹	۵۱	۰۳	۷۴	۱۷	۷۶	۳۷	۱۳	۰۴	۰۷	۷۴	۲۱	۱۹	۳۰
۵۶	۶۲	۱۸	۳۷	۳۵	۹۶	۸۳	۵۰	۸۷	۷۵	۹۷	۱۲	۲۵	۹۳	۴۷	۷۰	۳۳	۲۴	۰۳	۵۴	۹۷	۷۷	۴۶	۴۴	۸۰
۹۹	۴۹	۵۷	۲۲	۷۷	۸۸	۴۲	۹۵	۴۵	۷۲	۱۶	۶۴	۳۶	۱۶	۰۰	۰۴	۴۳	۱۸	۶۶	۷۹	۹۴	۷۷	۲۴	۲۱	۹۰
۱۶	۰۸	۱۵	۰۴	۷۲	۳۳	۲۷	۱۴	۳۴	۰۹	۴۵	۵۹	۳۴	۶۸	۴۹	۱۲	۷۲	۰۷	۳۴	۴۵	۹۹	۲۷	۷۲	۹۵	۱۴
۳۱	۱۶	۹۳	۳۲	۴۳	۵۰	۲۷	۸۹	۸۷	۱۹	۲۰	۱۵	۳۷	۰۰	۴۹	۵۲	۸۵	۶۶	۶۰	۴۴	۳۸	۶۸	۸۸	۱۱	۸۰
۶۸	۳۴	۳۰	۱۳	۷۰	۵۵	۷۴	۳۰	۷۷	۴۰	۴۴	۲۲	۷۸	۸۴	۲۶	۰۴	۳۳	۴۶	۰۹	۵۲	۶۸	۰۷	۹۷	۰۶	۵۷
۷۴	۵۷	۲۵	۶۵	۷۶	۵۹	۲۹	۹۷	۶۸	۶۰	۷۱	۹۱	۳۸	۶۷	۵۴	۱۳	۵۸	۱۸	۲۴	۷۶	۱۵	۵۴	۵۵	۹۵	۵۲
۲۷	۴۲	۳۷	۸۶	۵۳	۴۸	۵۵	۹۰	۶۵	۷۲	۹۶	۵۷	۶۹	۳۶	۱۰	۹۶	۴۶	۹۲	۴۲	۴۵	۹۷	۶۰	۴۹	۰۴	۹۱
۰۰	۳۹	۶۸	۲۹	۶۱	۶۶	۳۷	۳۲	۲۰	۳۰	۷۷	۸۴	۵۷	۰۳	۲۹	۱۰	۴۵	۶۵	۰۴	۲۶	۱۱	۰۴	۹۶	۶۷	۲۴
۲۹	۹۴	۹۸	۹۴	۲۴	۶۸	۴۹	۶۹	۱۰	۸۲	۵۳	۷۵	۹۱	۹۳	۳۰	۳۴	۲۵	۲۰	۵۷	۲۷	۴۰	۴۸	۷۳	۵۱	۹۲

۳-۴-۳ گزارش استاندارد کارآزمایی‌های بالینی

در مطالعات کارآزمایی بالینی علاوه بر طراحی علمی و اصولی نکات مهم دیگری حائز اهمیت هستند. این مراحل نحوه هدایت و گزارش نتایج مطالعات است. متأسفانه جزییات مراحل اجرایی، نحوه برخورد با خطاهای منظم پیش‌بینی‌نشده، نکاتی که برای بالا بردن کیفیت مطالعه و افزایش اعتبار و تکرارپذیری مطالعه لحاظ شده در مطالعه گزارش نمی‌شود و یا بسیار به‌اختصار بیان می‌شوند. در طول دهه اخیر تلاش‌های زیادی برای استاندارد شدن کیفیت گزارش‌ها در مطالعات کارآزمایی بالینی انجام شده که حاصل آن دستورالعمل CONSORT^A در سال ۱۹۹۶ میلادی و نسخه به‌هنگام شده آن در سال ۲۰۰۱ میلادی است [۲۷]-[۲۸]. اطلاعات بسیار جالب و مفیدی در این مورد در صفحه اینترنتی

<http://www.consort-statement.org>

وجود دارد (تاریخ دسترسی ۱۳۸۶/۷/۲۹).

نمودار ۱-۳ ساختار کلی موارد موردبررسی و جدول ۳-۴ چک‌لیستی از موارد مهم بررسی‌شده در دستورالعمل CONSORT را نشان می‌دهد.

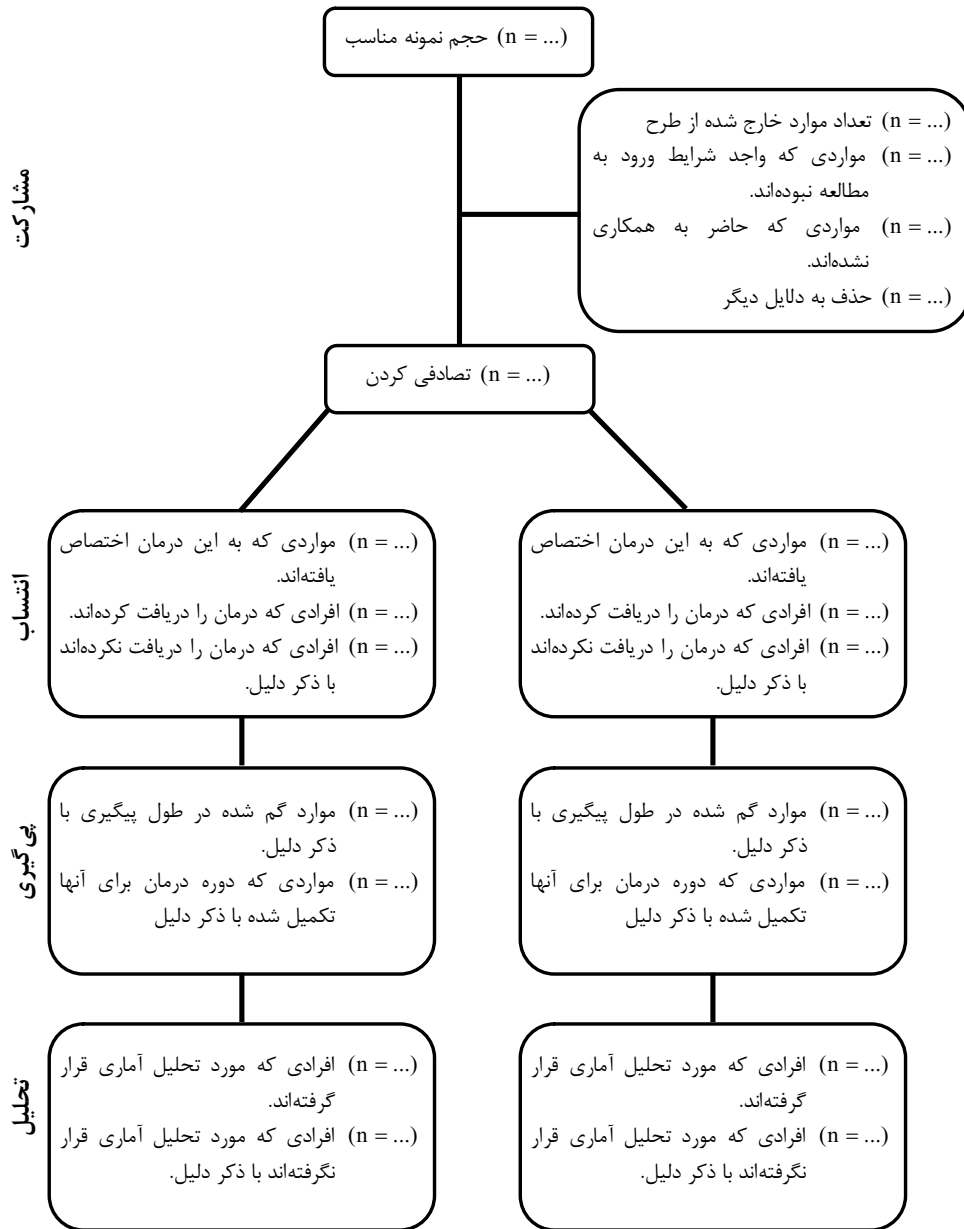
هرچند که به‌طورکلی دستورالعمل CONSORT برای تحلیل طرح‌هایی با دو گروه موازی تهیه‌شده اما به‌کارگیری آن در موارد پیچیده‌تر نظیر طرح‌های فاکتوریل، خوشه‌ای یا چندمرکزی و متقاطع نیز امکان‌پذیر است. نسخه‌های جدید دستورالعمل CONSORT برای گزارش این‌گونه مطالعات در حال تهیه و تنظیم است.

۱. **CON** solidate **S**tandards **O**f **R**eporting **T**rials. (یکپارچه‌سازی گزارش کارآزمایی‌ها)

جدول ۳-۴: چک‌لیست دستورالعمل CONSORT برای گزارش کارآزمایی‌های بالینی تصادفی شده

موردی باید کاملاً توضیح داده شود	فهرست مطالب
	عنوان و چکیده
۱- نحوه انتساب درمان‌ها به افراد مطالعه. (مثلاً انتساب تصادفی یا به روش بلوک‌های جایگشتی)	
	مقدمه
۲- مرور منابع، دلیل انجام مطالعه	کلیات
	روش‌ها
۳- معیارهای انتخاب بیماران، زمان و مکان جمع‌آوری اطلاعات	اعضای شرکت‌کننده در مطالعه
۴- جزئیات دقیق مداخله اعمال‌شده در هر گروه، چگونگی و زمان دقیق پذیرش بیماران	مداخله (درمان)
۵- اهداف اختصاصی و آزمون فرضیه‌های مربوط به مطالعه	اهداف
۶- تعریف دقیق پیامدهای اولیه و ثانویه، زمان و روش اندازه‌گیری این متغیرها. روش‌های به‌کاررفته برای اطمینان از کیفیت اندازه‌گیری‌ها.	پیامدهای موردنظر
۷- محاسبه حجم نمونه مطالعه و در صورت لزوم تعداد تحلیل‌های دوره‌ای هنگام انجام مطالعه و قواعدی برای اتمام زودتر از موعد مطالعه	حجم نمونه
	تصادفی کردن
۸- روش‌های به‌کاررفته برای تولید دنباله اعداد تصادفی و جزئیات مربوط به محدودیت‌های موجود (مثلاً بلوک‌بندی یا طبقه‌بندی داده‌ها)	تولید دنباله اعداد تصادفی
۹- مشخص کردن روش‌های به‌کاررفته برای مخفی نگه‌داشتن دنباله تصادفی انتساب‌ها پس از شروع مطالعه	روش‌های مخفی نگه‌داشتن دنباله تصادفی انتساب درمان‌ها
۱۰- ذکر افرادی که دنباله اعداد تصادفی را تولید کرده‌اند و افرادی که مسئول انتساب بیماران به هرکدام از درمان‌ها بوده‌اند.	اجرا
۱۱- اجرا یا عدم اجرای روش کورسازی، افرادی که نسبت به نوع درمان به‌کاررفته کور بوده‌اند و نحوه ارزیابی میزان موفقیت در اجرای کور کردن.	کور کردن
۱۲- روش‌هایی که برای مقایسه گروه‌ها از لحاظ پیامدهای موردنظر به کار گرفته می‌شوند. انجام تحلیل‌های جانبی نظیر مدل‌سازی، نحوه تعدیل مخدوش‌کننده‌ها و تحلیل در زیرگروه‌ها.	روش‌های آماری
	نتایج
۱۳- تعداد دقیق شرکت‌کنندگان در مطالعه در هر مرحله (ترجیحاً در یک دیاگرام نمایش داده شود). به‌طور خاص گزارش تعداد افراد موردنیاز مطالعه در هر گروه درمان، تعداد افرادی که به‌صورت تصادفی اختصاص داده شدند، تعداد افرادی که درمان	تعداد شرکت‌کنندگان در مطالعه

مواردی باید کاملاً توضیح داده شود	فهرست مطالب
موردنظر را دریافت کردند، تعداد افرادی که در پیگیری گم شدند و تعداد افرادی که در تحلیل‌ها موردبررسی قرار گرفتند باید ذکر شود. موارد انحراف از پروتکل مطالعه با ذکر دلیل باید گزارش شود.	
۱۴- زمان پذیرش هر بیمار در مطالعه و جزئیات مربوط به پیگیری‌ها	پذیرش
۱۵- مشخصات دموگرافیک و بالینی افراد تحت مطالعه	داده‌های زمینه‌ای
۱۶- تعداد افراد و شرکت‌کننده در هر گروه مطالعه که مورد تجزیه و تحلیل قرار گرفته‌اند به صورت اعداد دقیق (مثلاً ۱۰/۲۰ نه ۵۰٪) ذکر شود.	تعداد بیماران که مورد تحلیل قرار گرفته‌اند
۱۷- برای هر کدام از پیامدهای اولیه و ثانویه، خلاصه‌ای از نتایج در هر گروه و برآوردی از اندازه اثر درمان به کاررفته به همراه دقت آن (مثلاً فاصله اطمینان ۹۵٪) ذکر شود.	برآورد اثرها و پیامدها
۱۸- تحلیل‌های جانبی که می‌توانند شامل تحلیل چند متغیره تحلیل زیرگروه‌ها و یا تحلیل طبقه‌بندی شده باشد باید به همراه نتایج ذکر شود.	تحلیل‌های فرعی
۱۹- تمام پیامدهای نامطلوب مهم و یا اثرات جانبی مهم در هر گروه درمانی باید فهرست شود.	پیامدهای نامطلوب
	بحث
۲۰- تفسیر نتایج باید با توجه به آزمون فرضیه‌های مطالعه، منابع تورش بالقوه و خطرهای ناشی از سازگار نبودن نتایج تحلیل با یافته‌های قبلی انجام شود.	تفسیر نتایج
۲۱- بررسی میزان تعمیم (اعتبار خارجی) یافته‌ها	قابل تعمیم بودن نتایج
۲۲- تفسیر و استنباط نهایی از نتایج مطالعه با توجه به یافته‌های قبلی	استنباط کلی از مطالعه



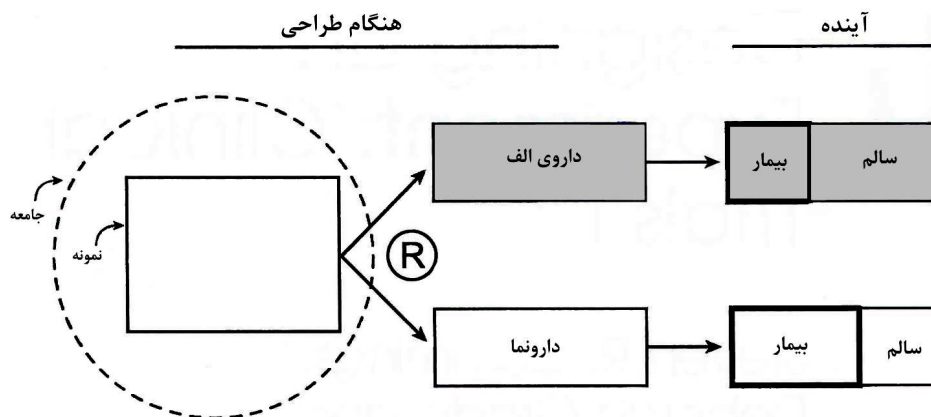
نمودار ۳-۱: تکمیل فرآیند پذیرش بیماران منطبق با دستورالعمل CONSORT برای گزارش کارآزمایی‌های بالینی تصادفی

۳-۴-۴ انواع مطالعات کارآزمایی بالینی

بررسی دقیق و جز به جز انواع مطالعات بالینی خارج از اهداف این کتاب است و در این قسمت تنها دسته‌بندی‌های کلی مطرح شده است خوانندگان محترم برای مطالعه بیشتر می‌توانند به منابع [۲۵] و [۲۹] تا [۴۰] مراجعه نمایند.

۳-۴-۴-۱ کارآزمایی با شاهد های هم‌زمان مستقل

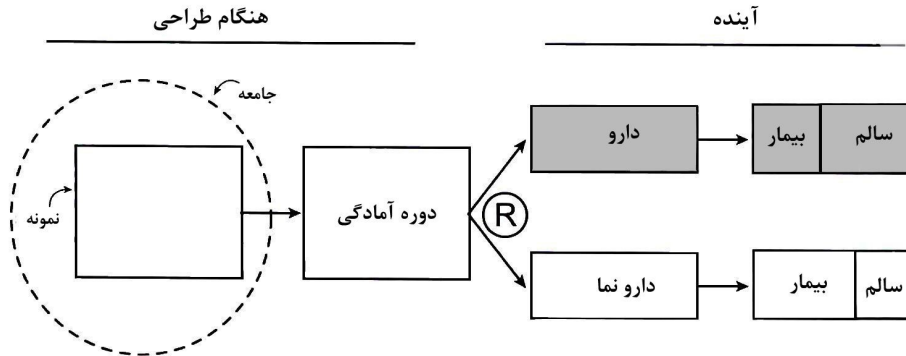
برای بررسی دقیق و نظارت مستقیم کارآزمایی، داشتن دو گروه الزامی است: گروه اول، درمان جدید را دریافت می‌کنند (گروه آزمایش) و به گروه دیگر، دارونما یا درمان استاندارد (گروه شاهد) تجویز می‌شود. گروه شاهد و گروه آزمایش باید از تمام جنبه‌ها، به جز روش درمانی موردنظر، یکسان باشند و تفاوت بین پاسخ‌ها در دو گروه فقط به واسطه روش درمانی (و نه عامل دیگری) اعمال شده، ایجاد شده باشد. بهترین راه، برای اطمینان از یکسان بودن درمان دو گروه، برنامه‌ریزی هم‌زمان و در یک دوره زمانی یکسان در هر دو گروه است. به این نوع طرح، مطالعه با شاهد های هم‌زمان موازی می‌گویند. شکل ۳-۷ شمایی از این مطالعه را نمایش می‌دهد



شکل ۳-۷: در یک کارآزمایی تصادفی شده محقق نمونه‌ای از جامعه جمع‌آوری می‌نماید و متغیرهای زمینه‌ای را اندازه‌گیری می‌کند. سپس افراد شرکت کننده در کارآزمایی به تصادف یکی از مداخله‌ها را دریافت می‌کنند. در طول پیگیری این افراد، متغیرهای مربوط به پیامدهای بیماری جمع‌آوری و تحلیل می‌شوند (R مخفف تصادفی سازی است).

۳-۴-۴-۲ طراحی موازی

اکثر کارآزمایی‌های بالینی تصادفی طراحی موازی دارند. در این مطالعات هر گروه از شرکت کنندگان در معرض یکی از مداخلات مطالعه قرار می‌گیرند. (شکل ۳-۸)



شکل ۳-۸: گاهی اوقات در انجام مطالعات بالینی تصادفی شده وجود یک دوره آمادگی برای بررسی عوارض احتمالی هر کدام از مداخله‌ها (و یا تجویز دارونما) تجویز می‌شود.

۳-۴-۳ مطالعه فاکتوریل

هنگامی که محققان مایل به بررسی ترکیبی یا مجزای دو یا چند مداخله با سطوح مختلف (بیش از ۲ سطح دارو) به همراه گروه شاهد هستند، یکی از گزینه‌های کارا و بسیار مناسب استفاده از طرح فاکتوریل است. آسان‌ترین روش یادگیری مطالعه فاکتوریل بررسی آن به وسیله ساده‌ترین حالت طرح فاکتوریل، یعنی مطالعه‌ای با دو مداخله (A و B) در دو سطح و چهار گروه درمانی است. در این مطالعه تعداد n بیمار در هر کدام از گروه‌های درمانی مطالعه می‌شوند. یکی از گروه‌ها هیچ‌کدام از درمان‌های A و B را دریافت نمی‌کنند، گروه بعدی هر دو درمان A و B را دریافت می‌کنند و دو گروه بعدی یکی تنها درمان A و دیگری تنها درمان B را دریافت می‌کنند. احتمال دارد گروه‌هایی که یکی از درمان‌ها یا هر دو درمان را دریافت نمی‌کنند، دارونما دریافت کنند. این مطالعه را طرح فاکتوریل ۲×۲ می‌نامند. به‌رغم سادگی، این طرح اکثر ویژگی‌های یک آزمایش فاکتوریل را دارا است. با استفاده از این مطالعه اطلاعات کافی برای بررسی اثرات درمان A و B به‌تنهایی و تأثیر هم‌زمان درمان‌های A و B فراهم می‌شود. نمایی ساده از این مطالعه در جدول ۳-۵ نمایش داده شده است.

جدول ۳-۵: چهار گروه‌درمانی با حجم نمونه برابر در هر گروه (طرح متعادل) در یک مطالعه فاکتوریل ۲×۲

مجموع	درمان B		درمان A
	دریافت کرده	دریافت نکرده	
۲n	n	n	دریافت نکرده
۲n	n	n	دریافت کرده
۴n	۲n	۲n	مجموع

تعمیم مطالعه فاکتوریل ۲×۲ به طرح‌های پیچیده‌تر به‌سادگی امکان‌پذیر است.

یکی از مشهورترین مطالعات فاکتوریل با مرتبه 2×2 مطالعه سلامت پزشکان^{۱۰} است. در این مطالعه تقریباً ۲۲,۰۰۰ پزشک از ایالات متحده برای بررسی اثرات ۱-آسپیرین در کاهش مرگ و میر ناشی از بیماری‌های قلبی و عروقی و ۲-بتا-کاروتن در کاهش بروز سرطان بررسی شد. هر پزشک در یکی از چهار گروه زیر قرار می‌گرفت:

۱- آسپیرین و بتا-کاروتن مصرف می‌کند.

۲- آسپیرین و بتا-کاروتن مصرف نمی‌کند.

۳- آسپیرین مصرف می‌کند ولی به جای بتا-کاروتن دارونما مصرف می‌کند.

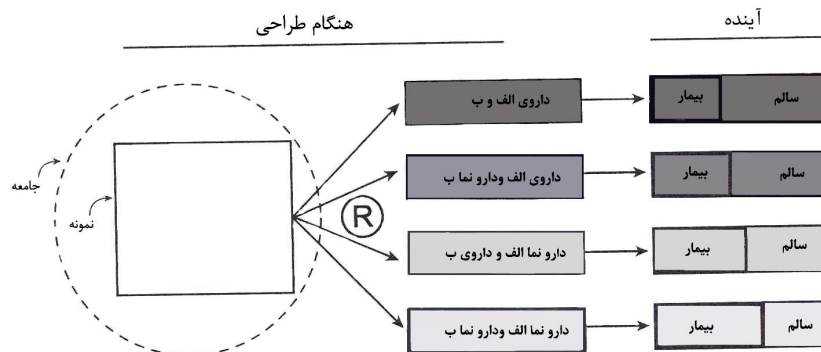
۴- بتا-کاروتن مصرف می‌کند ولی به جای آسپیرین دارونما مصرف می‌کند.

جدول ۳-۶ انتساب تصادفی پزشکان در این چهار گروه را نشان می‌دهد [۴۱] و [۴۲].

جدول ۳-۶: طرح فاکتوریل برای مطالعه تأثیر آسپیرین و بتا-کاروتن

مجموع	بتا-کاروتن		آسپیرین
	دریافت کرده	دریافت نکرده	
۱۱۰۳۷	۵۵۲۰	۵۵۱۷	دریافت نکرده
۱۱۰۳۴	۵۵۱۴	۵۵۲۰	دریافت کرده
۲۲۰۷۱	۱۱۰۳۴	۱۱۰۳۷	مجموع

مهم‌ترین محدودیت در انجام این مطالعات امکان وجود اثر متقابل بین درمان‌ها و پیامد موردنظر است. طراحی مطالعات فاکتوریل با وجود اثرات متقابل امکان‌پذیر است، اما این مطالعات پیچیده هستند و نیاز به حجم نمونه بسیار زیادی است، اجرای آنها مشکل است و تفسیر آنها به آسانی صورت نمی‌گیرد. بهترین حالت‌ها برای استفاده از طراحی فاکتوریل هنگامی است که مایل به بررسی دو سؤال پژوهش نسبتاً مجزا از یکدیگر هستیم. تصادفی کردن زوج‌های همسان شده یکی از راهبردهای بسیار مناسب برای متعادل کردن مخدوش‌کننده‌های پایه، نظیر سن، جنس است. در این نوع مطالعات به صورت تصادفی به هر کدام از زوج‌های همسان شده یکی از دو درمان موردنظر اختصاص داده می‌شود. یکی از کاربردهای جالب این مطالعه، استفاده از اعضای زوج بدن به عنوان زوج‌های همسان شده است. برای مثال در مطالعه‌ای بر روی بیماران رتینوپاتی دیابتی، در هر کدام از بیماران، به تصادف یک چشم لیزر درمانی و چشم دیگر بدون درمان (گروه کنترل) باقی می‌ماند. شکل ۳-۹ یک نمونه ساده از طراحی فاکتوریل 2×2 را نمایش می‌دهد.



شکل ۳-۹: در مطالعات فاکتوریل محققان نمونه‌ای از جامعه انتخاب می‌کنند و متغیرهای زمینه‌ای را اندازه‌گیری می‌نمایند. سپس به تصادف دو مداخله مؤثر و دارونماهای مشابه آن‌ها را به چهار گروه اختصاص می‌دهند. در طول پیگیری پس از تجویز داروها پیامدهای موردنظر اندازه‌گیری می‌شوند.

۳-۴-۴- طراحی متقاطع

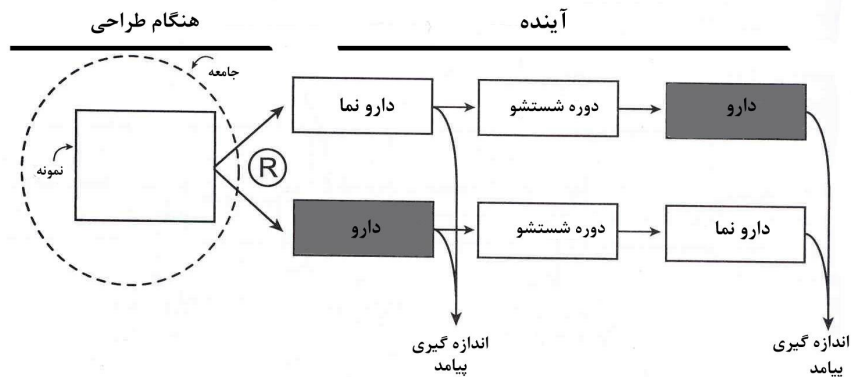
با انجام تغییرات اندکی در طرح مطالعاتی متوالی، می‌توان ترکیبی از شاهد‌های هم‌زمان و متوالی را به دست آورد. در این طرح از دو گروه بیمار استفاده می‌شود: یک گروه جهت دریافت درمان آزمایشی و گروه دوم جهت دریافت درمان داروی استاندارد یا دارونما.

در این مطالعات در ابتدای مطالعه، نیمی از شرکت‌کنندگان به تصادف، به‌عنوان گروه کنترل و بقیه جزء گروه درمان فعال قرار می‌گیرند. در مرحله دوم مطالعه، که معمولاً پس از یک دوره شستشو همراه است، جای دو گروه عوض می‌شود. این روش (که در هنگام وجود چند درمان تبدیل به طرح مربع لاتین می‌شود) عملاً امکان تجزیه و تحلیل بین گروه‌ها و درون گروه‌ها را فراهم می‌کند. شکل ۳-۱۰ نحوه انجام این طرح را نشان می‌دهد.

در این مطالعه چون هر عضو، به‌عنوان شاهد در مرحله بعدی مورد استفاده قرار می‌گیرد با حذف برخی متغیرهای مخدوش‌کننده بالقوه، توان آماری نسبت به طرح‌های موازی بیشتر خواهد بود. با این حال انجام طرح برای هر بیمار به حداقل دو مرحله نیاز است و پیچیده شدن تحلیل و تفسیر نتایج به علت امکان وجود اثر انتقالی از معایب این مطالعات هستند. اثر انتقالی به علت باقی ماندن اثرات درمان اولیه از دوره اول و تأثیر آن بر پیامد موردنظر بر ارزیابی‌های مربوط به مرحله دوم درمان، اتفاق می‌افتد. به‌عنوان مثال اثر درمان فشارخون با استفاده از درمان‌های دیوریتیک احتمالاً ماه‌ها باقی می‌ماند که این موضوع باعث می‌شود در ارزیابی درمان جدیدی که بلافاصله بعد از این درمان آغاز می‌شود سطح فشارخون مشاهده‌شده، هنوز متأثر از درمان‌های دیوریتیک باقی بماند.

به‌طور کلی در کارآزمایی‌های متقاطع، تنها پاسخ‌های کوتاه‌مدت در طول درمان و آخر دوره درمانی مدنظر است. هر اثر انتقالی، طولانی‌تر ناشی از درمان اول در دوره بعدی نامطلوب است و باعث مخدوش شدن نتایج می‌شود. اگر وجود اثر انتقالی محتمل باشد، باید پس از هر دوره درمان یک دوره شستشو با زمان مناسب وجود داشته باشد. در عین حال در بسیاری از موارد به علت جنبه‌های اخلاقی توقف دوره درمان امکان‌پذیر نیست. برای

- طراحی مطالعات متقاطع ویژگی‌های زیر در مورد مداخله و بیماری باید وجود داشته باشد:
- ♦ مداخلات معمولاً در بیماری‌های مزمن که دارای درمان سریع نیستند، استفاده می‌شود.
 - ♦ اثرات مداخله باید شروع سریع و دوره کوتاه داشته باشد.
 - ♦ وضعیت بیماری باید ثابت باشد.



شکل ۳-۱۰: در مطالعات متقاطع محققان نمونه‌ای از جامعه گرفته و متغیرهای زمینه‌ای را اندازه‌گیری می‌نمایند سپس به تصادف یکی از مداخله‌های موردنظر (دارو، دارونما) را به افراد مورد مطالعه اختصاص می‌دهند و پس از اعمال مداخله پیامدهای مربوط به آن را اندازه‌گیری می‌کنند. پس از یک دوره شستشو تخصیص مداخله‌ها برعکس شده و مراحل بالا تکرار می‌شود.

این مطالعه از جهاتی مزایای طرح‌های موازی و متوالی را در بردارد. گاهی اوقات طراحی متقاطع را می‌توان به بیش از دو درمان و در هر بیمار به دوره‌های متوالی بسط داد. مطالعاتی که برای توسعه داروها در مراحل اول آزمایش قرار دارند به علت کوتاه بودن فاصله زمانی بین درمان‌ها و اجرا به روی داوطلبان سالم اغلب از این دسته هستند. برای مثال می‌توان به مطالعه اثرات کوتاه‌مدت گازهای آلاینده بر کارکرد ریه داوطلبان اشاره کرد. در این مطالعه در مورد هر آلاینده (مثلاً دی‌اکسید سولفور) چهار سطح هیچ، کم، متوسط و زیاد وجود داشت که هر داوطلب به ترتیب تصادفی در چهار روز متوالی آن‌ها را دریافت می‌کرد [۲۳]. این مطالعه دارای طراحی مربع لاتین در بلوک‌های چهارتایی بود که برای توالی مقادیر تجویز از انتساب‌های تصادفی با اعداد ۱ تا ۴ استفاده شد. دوزهای مختلف آلاینده قبلاً به صورت تصادفی با حروف A، B، C یا D مشخص شده بودند. بدین ترتیب جدول زیر برای تکمیل مطالعه طراحی شد.

روزهای انجام مطالعه				توالی درمان‌ها
۴	۳	۲	۱	
D	C	B	A	۱
C	A	D	B	۲
B	D	A	C	۳
A	B	C	D	۴

۳-۴-۵ ارزیابی پیشرفت کار آزمایی

در بسیاری از کارآزمایی‌های بالینی بیماران تک‌تک وارد مطالعه می‌شوند و پاسخ‌های آنان به درمان به صورت تدریجی مشاهده می‌شود. به دلایل زیر پژوهشگران موظف هستند پیشرفت کارآزمایی را ارزیابی کنند:

- ♦ اطمینان از رعایت پروتکل مطالعه
- ♦ ارزیابی اثرات ناسازگار به‌ویژه عوارض جانبی و واکنش‌های سمی شدید درمان‌های جدید
- ♦ داده‌پردازی و توجه سریع به هرگونه خطا، ناسازگاری یا موارد گم‌شده در فرم‌های اطلاعاتی
- ♦ اطلاعات کلی در مورد چگونگی پیشرفت کارآزمایی
- ♦ مقایسه‌های میان‌دوره‌ای درحالی‌که کارآزمایی هنوز در جریان است به‌منظور این‌که از لحاظ اخلاقی از دریافت درمان‌های ناکارآمد به بیماران جلوگیری شود
- ♦ ارزیابی پاسخ‌های کوتاه‌مدت درمان‌ها

در مقایسه‌های میان‌دوره‌ای تنها باید به تعداد خاصی متغیر توجه داشت، زیرا در غیر این صورت تفسیر مقایسه‌های چندگانه مشکل می‌شود. توصیه می‌شود که تنها یک مقایسه درمانی بررسی شود و برای آن دستور توقف رسمی تعریف شود. مقایسه‌های دیگر را می‌توان به‌صورت غیررسمی برای هر مشخصه مهم به کار برد. تصمیم بر توقف یا تغییر یک کارآزمایی یک عمل صرفاً آماری نیست. در این مورد، مسائل دیگری نظیر اندازه اثر، میزان معنی‌داری تفاوت‌ها، جنبه‌های کاربردی درمان (نظیر سهولت، قابلیت پذیرش و هزینه) و ایده‌های جدید نیز باید در نظر گرفته شوند. در صورت وجود تحلیل‌های میان‌دوره‌ای در مطالعه تعریف دقیق قاعده توقف در کارآزمایی‌ها به روش‌های خاص آماری نیاز دارد. در این روش‌ها سطوح معنی‌داری آزمون‌ها بر اساس انجام مقایسه‌های چندگانه تعدیل شده است، نیاز است. این نوع تحلیل به تحلیل دنباله‌ای گروهی مشهور است.

۳-۴-۶ اندازه نمونه در کارآزمایی بالینی

پژوهشگران در طراحی کارآزمایی‌های بالینی همواره با این سؤال مواجه هستند که "چه تعداد بیمار برای انجام کارآزمایی نیاز است؟" به‌منظور دستیابی به اهداف اصلی و علمی هر کارآزمایی می‌توان روش‌های آماری برای تعیین تعداد کافی از بیماران به کار برد. **ضمیمه ج** به این امر اختصاص دارد.

درعین‌حال خاطر نشان ساختن این نکته که، این روش‌ها تنها به‌صورت رهنمودهایی به کاربرده می‌شود بسیار ضروری است. درواقع مسائل عملی مانند دستیابی به بیماران، منابع و ضرورت‌های اخلاقی پژوهش را نیز باید در نظر گرفت.

معمولاً وقتی روش‌های آماری برای تعیین حجم نمونه به کار گرفته می‌شود، پژوهشگران از تعداد زیاد و غیرقابل‌انتظار تعداد بیماران موردنیاز در مطالعه شگفت‌زده می‌شوند. در چنین مواقعی امکان دارد تصمیم گرفته شود اصول آماری فراموش شده و کارآزمایی به هر شکل ممکن انجام شود. متأسفانه این روش اغلب اوقات به کارآزمایی‌های کوچک که از امتیاز علمی کمی برخوردار هستند، منجر می‌شود. روش منطقی شروع کار با روش‌های آماری برای تعیین اندازه نمونه و سپس بررسی منابع مالی، تعداد بیماران در دسترس و سایر منابع است. گام بعدی ارزیابی میزان ورود بیماران شرکت‌کننده در کارآزمایی است. معمولاً این کار با برآورد تعداد

بیماران واجد شرایط در یک سال انجام می‌شود. پس از برآورد اندازه بیماران موردنیاز در کارآزمایی و میزان ورود بیماران، دوره زمانی لازم برای ورود بیماران قابل برآورد است. این نکته که ممکن است یک دوره اضافی برای انجام پیگیری‌ها لحاظ شود را نیز باید به خاطر سپرد.

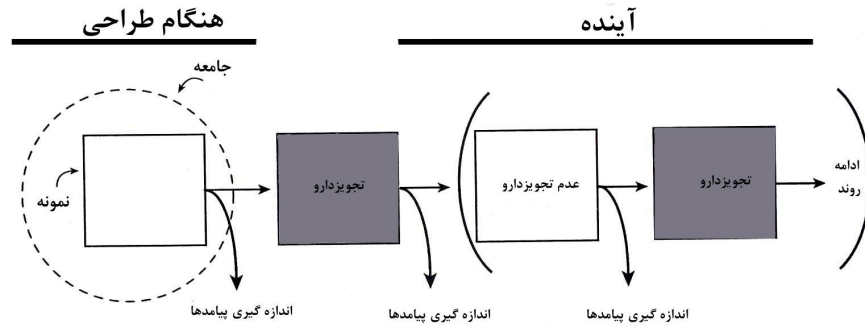
غالباً منابع مجزای بیماران اعم از بیمارستان، مطب‌های خصوصی یا مراکز تحقیقاتی به‌تنهایی برای اجرای کارآزمایی‌های بالینی با اندازه مناسب، ناکافی هستند بنابراین، در زمان طراحی کارآزمایی‌های بالینی، تشخیص این که آیا انجام مطالعه تنها در یک مرکز امکان‌پذیر است از اهمیت خاصی برخوردار است. اگر نیاز به اجرا کارآزمایی‌های چندمرکزی وجود داشته باشد باید دلایل، مزایا و معایب آن به‌دقت بررسی شود.

برخی مزایا و معایب مطالعات کارآزمایی‌های چندمرکزی عبارت‌اند از:

- ♦ حسن عمده کارآزمایی‌های چندمرکزی ورود سریع‌تر بیماران است. به همین دلیل حجم نمونه در کارآزمایی را می‌توان بزرگ‌تر در نظر گرفت و یا در زمان کوتاه‌تری به نتایج نهایی دست یافت.
- ♦ انجام کارآزمایی در چند مرکز باعث افزایش منابع بیماران و همچنین تنوع گروه درمانی می‌شود به همین علت تعمیم نتایج در این مطالعات با اطمینان بیشتری انجام می‌شود.
- ♦ برنامه‌ریزی و اداره کارآزمایی‌های چندمرکزی پیچیده‌تر از کارآزمایی‌های معمولی است.
- ♦ اجرای کارآزمایی‌های چندمرکزی از نظر نیروی انسانی و منابع بسیار گران است.
- ♦ اطمینان از این که تمام مراکز از پروتکل مطالعه پیروی خواهند کرد، بسیار حائز اهمیت است.
- ♦ کنترل کیفیت اندازه‌گیری‌ها، مشاهدات بالینی و ثبت داده‌ها در این مطالعات بسیار ضروری است به این منظور آموزش و ارائه توضیحات کافی به گروه تحقیق ضروری است.
- ♦ جمع‌آوری داده‌ها و داده‌پردازی در کارآزمایی‌های چندمرکزی نیاز به یک گروه مجرب و سازمان‌یافته دارد.
- ♦ به‌طور کلی هر چه تعداد مراکز در یک کارآزمایی بالینی بیشتر باشد، مشکلات و پیچیدگی کارآزمایی بیشتر می‌شود.
- ♦ تحلیل آماری داده‌های حاصل از کارآزمایی‌های چندمرکزی روش‌های مخصوصی نیاز دارد.

۳-۴-۷ طراحی با شاهد‌های متوالی (مطالعه نیمه تجربی)

یک مطالعه مداخله‌ای زمانی طراحی متوالی دارد، که هر یک از شرکت‌کنندگان تمام یا تعدادی از مداخلات مطالعه را در دوره‌های متوالی دریافت کنند. این که کدام شرکت‌کننده، کدام‌یک از مداخلات را دریافت کند به‌طور تصادفی معین می‌شود. در طراحی متوالی هر یک از شرکت‌کنندگان، شاهد خود هستند. در پزشکی زمانی استفاده از این طرح ضرورت می‌یابد که به علت عوارض جانبی خطرناک دارو یا احتمال تأثیر بسیار جدی درمان در بیماری‌های خطرناک، از لحاظ اخلاقی باید کمترین تعداد بیماران در پژوهش بررسی شوند. شکل ۳-۱۱ شمایی از نحوه انجام این مطالعه را نشان می‌دهد.



شکل ۳-۱۱: در مطالعات با شاهد‌های متوالی محققان نمونه از جامعه انتخاب کرده و متغیرهای زمینه‌ای را اندازه‌گیری می‌کنند سپس به همه اعضای گروه یک نوع مداخله تجویز می‌کنند و متغیرهای مربوط به بیماری یا پیامد موردنظر را در طول پیگیری اندازه‌گیری می‌کنند. گاهی اوقات در ادامه نوع مداخله تغییر یافته و مراحل بالا تکرار می‌شود.

۳-۴-۸ کارآزمایی با شاهد‌های خارجی

روش دیگر برای کنترل آزمایش‌ها، استفاده از شاهد‌هایی خارج از دایره مطالعه است. ممکن است در این روش، از نتیجه تحقیق‌های دیگر به‌عنوان مقایسه استفاده می‌شود. گاهی، افراد شاهد که شاهد تاریخی نامیده می‌شود، بیمارانی هستند که پژوهشگر قبلاً به طریق دیگری آن‌ها را درمان کرده است. از شاهد‌های تاریخی، اغلب در پژوهش‌های سرطان‌شناسی استفاده می‌شود. پژوهشگر در جریان مطالعات با شاهد تاریخی، باید به تغییرات سایر عوامل (زمان درمان افراد شاهد) نیز توجه نماید، زیرا ممکن است تفاوت حاصل، ناشی از تغییر در این عوامل باشد و نه خود در درمان.

۳-۴-۹ مطالعات بدون شاهد

در این دسته از مطالعات، مداخله وجود دارد ولی شاهد وجود ندارد. به این دسته از مطالعات در یک تعریف دقیق‌تر کارآزمایی یا تجربه نمی‌گوییم. به هنگام سنجش یک روش تشخیصی مطالعات بدون شاهد کاربرد فراوانی دارند.

ضعف عمده چنین مطالعاتی این است که متأسفانه پژوهشگران روش خود را بهترین می‌دانند. در تاریخ پزشکی مملو از نمونه‌هایی است که داروی به‌کاربرده شده، پس از انجام کارآزمایی بالینی شاهددار کنار گذاشته شده اس مهم‌ترین مسئله در مطالعات بدون شاهد، این است که متأسفانه روش‌های درمانی اثبات نشده، متداول گشته و این موضوع انجام مطالعات شاهددار را با مشکلات فراوانی مواجه کرده است.

۳-۴-۱۰ کارآزمایی تصادفی نشده

انتخاب بیماران برای روش‌های درمانی، همیشه به‌طور تصادفی نیست. مطالعاتی را که بیماران آن به‌طور تصادفی

انتخاب نشده‌اند، کارآزمایی تصادفی نشده می‌نامند. بسیاری از پژوهشگران در معرض انواع بسیار زیادی از تورش‌ها بوده، به‌نحوی که نتایج آن‌ها به میزان زیادی، مورد تردید است. مطالعات با افراد شاهد تصادفی نشده از نظر اعتبار بسیار ضعیف‌تر از مطالعات تجربی هستند. زیرا راهی برای جلوگیری از تورش‌های ناشی از نسبت دادن غیرتصادفی بیماران وجود ندارد.

۳-۴-۱۱ کارآزمایی‌های میدانی

کارآزمایی‌های میدانی از این لحاظ با کارآزمایی‌های بالینی متفاوت است که، در این مطالعات کارآزمایی بر روی افراد سالم متمرکز می‌شود. در کارآزمایی‌های بالینی احتمال بروز عوارض کوتاه‌مدت بیماری بسیار زیاد است. در صورتی که احتمال بروز عوارض کوتاه‌مدت در کارآزمایی‌های میدانی بسیار نادر است. به همین دلیل این مطالعات معمولاً بسیار گسترده بوده و به حجم نمونه زیاد نیاز دارند. از طرف دیگر، چون در این مطالعات افراد سالم مدنظر هستند، انتظار مراجعه این افراد به مراکز خاصی نظیر بیمارستان یا درمانگاه منطقی نیست و برای معاینه و دسترسی به افراد تحت مطالعه، غالباً لازم است به محل زندگی، کار یا تحصیل افراد مراجعه نموده و یا مراکز خاصی برای هدایت مطالعه و ترغیب افراد تحت مطالعه برای مراجعه به آن تدارک دیده شود. تمامی این مسائل باعث پرهزینه شدن این نوع مطالعات می‌شوند. به همین علت این مطالعات تنها در مورد پیشگیری از بیماری‌های بسیار شایع و یا فوق‌العاده خطرناک اجرا می‌شوند. برای مثال کارآزمایی‌های میدانی متعددی در مورد مصرف مقادیر زیاد ویتامین C و جلوگیری از سرماخوردگی طراحی و اجرا شده است [۴۳]-[۴۴].

بیماری فلج اطفال نیز یک بیماری نادر اما بسیار خطرناک است که به علت اهمیت عوارض آن، باعث اجرای بزرگ‌ترین مطالعه تجربی بر روی انسان‌ها گردید. در این کارآزمایی، واکسن تزریقی فلج اطفال یا دارونما به صدها هزار کودک مدرسه‌ای تزریق شد [۴۵].

در این مطالعات نیز نظیر کارآزمایی‌های بالینی، نحوه اختصاص روش‌های مورد مطالعه باید به نحوی باشد که گروه‌های تحت مطالعه قابل مقایسه باشند. انتساب تصادفی می‌تواند به‌عنوان یک گزینه ایده‌آل باشد؛ اما مشکلات عملی به کار بردن این رویه در مطالعاتی با حجم نمونه بسیار زیاد ممکن است موجب انصراف محققان از انجام این روش شود. به‌عنوان مثال تزریق محلول روغنی یددار به دانش‌آموزان مناطق کوهستانی و هیپراندمیک و مقایسه آن با گروه شاهد مناسب در کاهش درجه گواتر و افزایش غلظت هورمون‌های تیروئید و اثر آن بر پیشرفت بهره رشد و توانایی روانی - حرکتی نمونه مناسبی از این نوع کارآزمایی‌ها به شمار می‌رود [۴۶] و [۴۷].

۳-۴-۱۲ مداخله‌های اجتماعی و کارآزمایی‌های تصادفی خوشه‌ای

کارآزمایی‌های مینی بر مداخله‌های اجتماعی، حالت خاصی از مطالعات میدانی است که گستردگی آن تمام جامعه را در بر می‌گیرد. برای مثال افزودن فلوراید به آب آشامیدنی یک شهر برای بررسی امکان پیشگیری از پوسیدگی دندان در مقایسه با یک جامعه شهری دیگر که تنها از آب معمولی استفاده می‌کنند، یک کارآزمایی مبتنی بر مداخله اجتماعی است.

یک مثال بارز در این زمینه، برنامه پیشگیرانه کمیته کشوری مبارزه با کمبود ید در تولید، توزیع و ترویج مصرف نمک یددار در کل کشور است [۴۸].

گاهی، مداخله‌های تعریف شده برای مطالعه در گروه‌هایی کوچک‌تر از تمام جامعه به کار می‌روند. برای مثال مداخله‌هایی که مربوط به رژیم غذایی است، ممکن است در مورد اعضای یک خانواده اعمال شوند. برخی مداخله‌های محیطی ممکن است بر روی ساکنان یک واحد مسکونی و یا کارکنان یک اداره، آزمایش و بررسی شوند. یک وسیله جدید ورزشی ممکن است در یک تیم ورزشی مورد مطالعه قرار گیرند. گروه‌هایی که مداخله موردنظر بر روی آن‌ها آزمایش می‌شود، می‌توانند شامل واحدهای مختلف ارتش، مدارس، روستاهای یک شهرستان و یا هر گروه مشابه دیگری که دارای اعضای همانند هستند (از لحاظ مواجهه‌های مهم تأثیرگذار و مداخله مورد مطالعه)، باشند. اعمال مداخله در چنین گروه‌هایی به مراتب آسان‌تر و عملی‌تر از مطالعات میدانی وسیع و مداخلات اجتماعی است. چنین مطالعاتی که مداخله‌های مورد بررسی به تصادف به جای افراد به برخی گروه‌ها منتسب می‌شوند مطالعات تصادفی شده خوشه‌ای نامیده می‌شوند.

۳-۵ مطالعات بر مبنای داده‌های موجود

بسیاری از سؤالات پژوهش با استفاده از داده‌هایی که قبلاً جمع‌آوری شده‌اند با سرعت و کارایی مناسبی قابل بررسی هستند. به‌طور کلی سه روش اصلی برای استفاده از داده‌های موجود در پژوهش وجود دارد: تحلیل داده‌های ثانویه، مطالعات فرعی و بازیابی منظم.

سریع و به‌صرفه بودن این روش‌ها مهم‌ترین مزیت استفاده از آن‌ها در پژوهش است. گاهی اوقات، بررسی یک سؤال پژوهش به طریق متداول ممکن است به منابع مالی زیاد و زمان بسیار طولانی نیاز داشته باشد؛ اما با استفاده از داده‌های موجود امکان دستیابی سریع و ارزان به اهداف مطالعه فراهم باشد. به‌عنوان مثال یک مطالعه مداخله‌ای با چند عامل خطر و مداخله متفاوت با نام اختصاری MRFIT برای بررسی عوامل مؤثر در شیوع بیماری‌های قلبی و عروقی طراحی شد. در طول مطالعه اطلاعات مربوط به الگوی مصرف سیگار در افراد تحت مطالعه، بررسی و ثبت شد. پس از اتمام مطالعه، یکی از محققان متوجه شود که امکان بررسی اثرات مواجهه غیرارادی با دود سیگار^{۱۱} و شیوع بیماری‌های قلبی و عروقی فراهم شده است. نتایج این مطالعه نشان داد که شیوع بیماری‌های قلبی در مردان غیر سیگاری دارای همسر سیگاری دو برابر مردان متأهل با همسر غیر سیگاری است. نتیجه‌ای که بسیار جالب‌تر و جدیدتر از اهداف و یافته‌های اصلی مطالعه بود [۴۹].

با این حال مطالعات مبتنی بر داده‌های موجود دارای محدودیت‌هایی نیز هستند. انتخاب منابع مناسب برای جمع‌آوری داده‌ها، از قسمت‌های پیچیده این مطالعات است [۵۰]. گذشته از این نحوه اندازه‌گیری و ثبت اطلاعات در این داده‌ها، از اختیار محققان خارج است. مثلاً امکان دارد یک متغیر دوحالتی برای ابتلا یا عدم ابتلا به بیماری پرفشاری خون جمع‌آوری شده باشد. هم‌چنین این امکان وجود دارد که کیفیت داده‌ها مناسب نباشد و موارد گم‌شده یا اشتباه ثبت‌شده فراوان باشد [۵۱]. احتمال عدم ثبت و جمع‌آوری عوامل مداخله‌گر مهم، پیامدهای

۱. Passive smoking

اساسی و نتایج پایانی افراد تحت مطالعه نیز وجود دارد. تمام این محدودیت‌ها را می‌توان در این جمله خلاصه کرد که در تحقیقاتی مبتنی بر داده‌های موجود محققان هیچ نظارتی بر نحوه جمع‌آوری داده‌ها ندارند.

۳-۵-۱ تحلیل مجدد داده‌ها

در تحلیل مجدد داده‌ها، اطلاعات به دو صورت فردی و کلان در دسترس هستند. داده‌های فردی، داده‌هایی هستند که اطلاعات هر فرد را به‌طور جداگانه در بردارند. این داده‌ها را می‌توان از مطالعات قبلی، سوابق پزشکی، پرونده‌های بیمارستانی، گواهی‌های فوت و بسیاری از منابع دیگر به دست آورد. در چنین داده‌هایی، وجود ارتباط بین ویژگی‌های موردنظر و نمونه با استفاده از اطلاعات به‌دست‌آمده از پرونده‌ها و مدارک بررسی می‌شود. مراحل بعدی تحقیق، دقیقاً مشابه فرآیندی است که محققان، اندازه‌گیری‌ها و اطلاعات را مستقیماً از اعضا مطالعه جمع‌آوری می‌نمایند. یکی از منابع مهم جمع‌آوری اطلاعات اعضای تحت مطالعه، داده‌های جمع‌آوری شده از مطالعات قبلی است. در بسیاری از مطالعات، داده‌های بیشتری نسبت به آنچه محققان می‌توانند تحلیل کنند جمع‌آوری می‌شود. نوع دوم شامل مجموعه بزرگی از داده‌های ملی و منطقه‌ای است که عموماً در دسترس هستند و منحصر به گروه خاصی نیست. برای مثال مراجعه به مرکز آمار ایران امکان دسترسی به این داده‌ها را سریع‌تر می‌کند. هنگامی که داده‌های فردی در دسترس نباشند، داده‌های کلان گاهی مفید هستند. در داده‌های کلان فقط اطلاعات مربوط به گروه‌های بزرگ جامعه (یک شهر یا استان) در دسترس هستند (مثلاً میزان‌های مرگ ناشی از سرطان روده بزرگ در تمام استان‌های کشور). با چنین داده‌هایی ارتباطات بین گروه‌های جمعیتی به‌وسیله مقایسه اطلاعات یک عامل خطر (مانند مصرف سرانه گوشت) و اندازه وقوع پیشامد موردنظر (سرطان روده بزرگ) بررسی می‌شود.

تحلیل داده‌های موجود به دو روش ممکن است:

- ♦ محقق ممکن است علاقه‌مند به بررسی یک سؤال تحقیق خاص باشد و سعی در یافتن مجموعه داده‌های داشته باشد که بتواند به آن سؤال پاسخ دهد. این روش، روش معمول تحقیقات بالینی است.
 - ♦ محقق به بررسی و تحلیل داده‌های موجودی که در دسترس دارد می‌پردازد، سپس سؤالاتی که به‌وسیله این مجموعه داده‌ها قابل پاسخگویی است را موردبررسی قرار می‌دهد. مشکل این روش یافتن روابط معنی‌دار در میان انبوهی از اطلاعات است.
- مراحل مختلف پیدا کردن یک بانک اطلاعاتی مناسب با سؤال پژوهش به‌صورت زیر است:
- ♦ پیدا کردن یک سؤال پژوهش و بررسی متون علمی مرتبط
 - ♦ تهیه کردن فهرستی از متغیرهای پاسخ و مخدوش‌کننده
 - ♦ پیدا کردن بانک‌های اطلاعاتی مناسب، که اطلاعات موردنیاز را در برداشته باشد.
 - ♦ انتخاب بهترین مجموعه داده‌ها که با اهداف پژوهش سازگار است.
 - ♦ تنظیم فرضیه‌های تحقیق و مشخص کردن روش‌های آماری مناسب برای تحلیل نتایج
 - ♦ تحلیل داده‌ها
- مراحل مختلف طراحی سؤالات تحقیق، متناسب با داده‌های موجود به این صورت است:

- ♦ انتخاب بانک اطلاعاتی مناسب
- ♦ آشنایی با متغیرهای جمع‌آوری شده و نحوه اندازه‌گیری آن‌ها
- ♦ مشخص کردن متغیرها و یا زیرگروه‌هایی از اطلاعات که برای پاسخ به سؤالات تحقیق مناسب به نظر می‌رسد.
- ♦ مرور منابع و مشورت با سایر محققان درباره اهمیت و جنبه‌های جدید سؤال تحقیق
- ♦ تنظیم فرضیه‌های تحقیق و مشخص کردن روش‌های آماری تحلیل نتایج
- ♦ روش تحلیل داده‌ها

۳-۵-۲ مطالعات فرعی

طراحی مطالعه به روش تحلیل مجدد داده‌ها نیازمند پاسخگویی به سؤال تحقیقی است که اطلاعات مربوط به آن واقعاً در دسترس است. در مطالعه فرعی، برای محقق این امکان وجود دارد که تعداد محدودی متغیر جدید اضافه بر آنچه در تحقیق پیش‌بینی شده را اندازه‌گیری و ثبت نماید. مطالعات فرعی، بسیاری از فواید تحلیل داده‌های ثانویه را دارا هستند. در ضمن نسبت به این مطالعات، از محدودیت کمتری برخوردارند. مطالعات فرعی می‌توانند به هر نوع مطالعه (مثلاً مطالعات مقطعی و مطالعات مورد-شاهدی) اضافه شوند؛ اما مطالعات هم‌گروهی آینده‌نگر بزرگ و کارآزمایی‌های تصادفی به‌خوبی برای طراحی‌های چنین مطالعاتی سازگار هستند. اندازه‌گیری‌ها در مطالعه فرعی طراحی شده برای کارآزمایی‌های تصادفی معمولاً هنگامی که قبل از شروع کارآزمایی برنامه‌ریزی می‌شوند، مفید و اطلاع‌رسان هستند ولی جمع‌آوری این اطلاعات برای افرادی که از ابتدا عضو گروه تحقیق نبوده‌اند بسیار مشکل است. وقتی متغیرها در ابتدا اندازه‌گیری شده باشند، اندازه‌گیری آن‌ها در طول یا پایان کارآزمایی می‌تواند اطلاعات مفیدی را در اختیار محقق قرار دهد. در بیشتر مطالعات هم‌گروهی آینده‌نگر محقق به‌طور متناوب اندازه‌گیری‌های جدید را اضافه می‌کند. این موضوع یک موقعیت ایده‌آل برای طراحی مطالعات فرعی است

امکان طراحی مطالعات فرعی باید به‌طور مداوم پیگیری شوند. بخصوص به‌وسیله محققان تازه‌کار که دارای محدودیت‌های زمانی و منابع محدود هستند. مثلاً محقق که علاقه‌مند به بررسی اثرات کاهش وزن بر روی درد آرتروز است ممکن است کار را با تعریف مداخله‌هایی از قبیل رژیم، ورزش، تغییر عادت و استفاده از داروهای مؤثر بر کاهش وزن شروع کند. در این مطالعه ممکن است متغیرهای مداخله‌گر به‌وسیله مشاوره متخصصان کاهش وزن و یا مرور مقالات معین شوند. گام بعدی رده‌بندی درجات مختلف آرتروز در بیماران و هم‌چنین تعریف نحوه اندازه‌گیری متغیرهای پاسخ یعنی میزان درد و تغییر وزن (به‌عنوان اندازه فرعی) است. بعداً این مرحله محقق باید در جستجوی مطالعه‌ای باشد که فرصت خوبی را برای اندازه‌گیری‌های فرعی مهیا می‌کند. جلب نظر محققان اصلی تحقیق، مهم‌ترین مرحله این نوع تحقیقات است. بیشتر محققان تمایل دارند در مطالعه اصلی برخی معیارها و اندازه‌گیری‌های فرعی اضافه شوند. البته در صورتی که سؤال مهمی را بیان کنند و اساساً مانع اجرا و هدایت مطالعه اصلی نشوند. مثلاً اکثر محققان از افزودن یک پرسشنامه و یا یک سری اندازه‌گیری‌های جالب، کم‌هزینه و سریع استقبال می‌کنند درحالی که تمایل چندانی به گنجاندن اندازه‌گیری‌های

وقت گیر، پرهزینه، خطرناک و ناخوشایند یا غیراخلاقی نظیر آندوسکوپی ندارند. به طور کلی گنجاندن مطالعه فرعی به فرآیند تحقیق نیازمند اجازه از محقق اصلی یا کمیته مطالعه است.

معایب مطالعات فرعی کم است؛ اما در بعضی موارد مشکلات عملی برای کسب اجازه رسمی برای انجام مطالعه یا آموزش دادن به کسانی که اندازه‌گیری را انجام می‌دهند و یا گرفتن رضایت‌نامه از تک‌تک شرکت‌کنندگان وجود دارد. ممکن است مجری مطالعه فرعی از اهداف و نحوه اجرای مطالعه اصلی اطلاعی نداشته باشد و در نتیجه برای دستیابی به همه اطلاعات نیز مشکل داشته باشد.

۳-۵-۳ بازبینی منظم و متاآنالیز

در بازبینی منظم به بررسی، ارزیابی و نقد مطالعات تکمیل‌شده‌ای که درصدد پاسخ به یک سؤال تحقیق هستند، توجه می‌شود. هدف این مطالعه جمع‌بندی کلی و نهایی یک سؤال پژوهش خاص است. در مقایسه با روش‌های دیگر مرور منابع، بازبینی منظم، روش استاندارد و شناخته‌شده‌ای برای تعریف همه مطالعات مرتبط با سؤال تحقیق، بیان نتایج در مطالعات موجه و در صورت امکان محاسبه برآوردهای خلاصه به‌دست‌آمده از تمام نتایج است.

روش‌های آماری بازبینی منظم (محاسبه برآورد اثر و واریانس، آزمون‌های آماری برای بررسی همگنی و بررسی‌های آماری تورش نشر) متاآنالیز نامیده می‌شود. بازبینی منظم فرصت خوبی برای یک محقق تازه‌کار است. با وجود اینکه بازبینی منظم دقت و تلاش زیادی لازم دارد، اما بار مالی و یا منابع اضافی نیاز ندارد. برای انجام یک بازبینی منظم محقق باید با ادبیات موضوع تحقیق و سابقه سؤال تحقیق کاملاً آشنا باشد. بازبینی‌های منظم بر اساس جستجوی مقایسه‌ای و بدون تورش در مطالعات تکمیل‌شده طرح‌ریزی می‌شوند. پروپوزال بازبینی منظم باید دارای معیارهای ورود و خروج مناسب مطالعه باشند و این معیارها باید پیش از شروع به کار مشخص شوند. معیارهای ورود و خروج در مطالعات متاآنالیز به‌طور معمول مواردی نظیر محدوده زمانی انجام و انتشار مطالعه، جامعه تحت مطالعه، بیماری یا شرایط موردنظر، مداخله مطالعه، لزوم کور کردن مطالعه، تصادفی کردن آزمون‌ها گروه‌های کنترل موردقبول، تعداد داده‌های گم‌شده در پیگیری و کمترین مدت پیگیری موردقبول را مشخص می‌کنند. زمانی که این معیارها مشخص شدند مطالعه به‌وسیله دو یا چند محقق بازبینی می‌شود.

به‌طور کلی بازبینی‌های منظم شامل اطلاعات زیر است.

- ♦ خلاصه شاخص‌های مهم هر مطالعه، که اغلب شامل حجم نمونه مطالعه‌ها، درصد پاسخ‌گویی، مدت پیگیری، شاخص‌های جمعیتی مطالعه و روش‌های آماری استفاده شده است.
 - ♦ بازبینی نتایج اختصاصی هر مطالعه، که شامل برآوردهای خطر و فواصل اطمینان است.
- لطفاً برای کسب اطلاعات بیشتر در خصوص مطالعات فراتحلیل یا متاآنالیز به منابع شماره ۵۲ و ۵۳ مراجعه شود.

منابع

1. Kleinbaum DG, Kupper LL, Morgenstern H. Epidemiologic research: Principles and quantitative methods. New York, Van Nostrand Reinhold, 1982.
2. MC Mahan B, Pugh TF. Epidemiology: Principles and methods. Boston, Little Brown 1970:137-198, 175-184.
3. Morgenstern H, Thomas DC. Principles of study design in environmental epidemiology. Environ Health Perspect 1993, 101:23-38
4. Walter SD. The analysis of regional patterns in health data I, II. Am J Epidemiol 1992, 136:736-759.
5. Prentice RL, Sheppard L. Aggregate data studies of disease risk factors: Biometrika 1995, 85:113-125.
6. Rose G, Sick individuals and sick population. Int J Epidemiol 1985, 14:32-38.
7. Prentice RL, Kakar F, Hurstings et al. Aspects of the rationale for the women's health trial. J Natl Cancer Inst 1988, 80:802-14.
8. Clayton D, Kaldor J. Empirical Bayes estimates of age-standardized relative risks for use in disease mapping. Biometrics 1987, 43:671-81.
9. Greenland S, Robin J. Invited commentary: ecologic studies biases, misconceptions, and counter examples. Am J Epidemiol 1994, 139:747-60.
10. Koopman JS, Longin, IM Jr. The ecological effects of individual exposures and nonlinear disease dynamics in populations. Am J public Heath 1994, 84:836-42.
11. Wacholder S, Mc Laughlin JK, Silverman DT. Mandel JS. Selection of controls in case-control studies I: Principles Am J Epidemiol 1992, 135:1019-28.
12. Wacholder S, Mc Laughlin JK, Silverman DT. Mandel JS. II Type of controls. Am J Epidemiol 1992, 135:1029-41.
13. Soori H. Developmental risk factors for unintentional childhood poisoning in Iran. Saudi Medical Journal. 2001; Vol.22: 227-230.
۱۴. امینی نیک س، شجاعی تهرانی ح، کلیشادی ر و همکاران. اپیدمیولوژی در پزشکی تألیف چارلی اچ و بورینگ جی، مرکز تحقیقات قلب و عروق دانشگاه علوم پزشکی اصفهان، اصفهان، ۱۳۷۶
15. Kaplan N., Stamler J. Prevention of coronary heart disease. Philadelphia, Saunders, 1983.
16. Kannel WB. Coronary heart disease risk factors: A Framingham study update. Hos Pract, 1990, 25: 93-104.
17. Fron Dawber TR, Kannel WB, Lyell LP, An approach to longitudinal studies in a community: The Framingham study. Ann NY Acad Sci, 1993, 107: 539-561.

18. Doll R, Hill AB. Mortality in relation to smoking. The 25 year observation of British doctors. *Br Med J*, 1964, 1:1399.
 19. Gordon T, Moore FF, Shurtleff D, et al. Some epidemiologic problems in the long-term study of cardiovascular disease. Observations on the Framingham study. *J chronic Dis*, 1959, 10:186.
 20. Kitigawa EM, Hauser PM. Different mortality in the United States. Cambridge, MA: Harvard University Press, 1973.
 21. Lathrop GD, Wolfe WH, Albanese RA et al. Project Ranch Hand II. An epidemiologic investigation of health effects in Air Force personnel following exposure to Herbicides: Baseline morbidity study results. San Antonio, TX: US Air Force School of Aerospace Medicine, Aerospace Medical Division, Brooks Air Force Base, 1984.
 22. Matan Oski GM, Seltzer R, Sartweel PE, et al. The current mortality rates of radiologists and other physician specialists: specific causes of death. *Am J Epidemiol*, 1975, 101:199.
 23. Mc Michael AJ, Haynes SG, Tyroler HA. Observations on the evaluation of occupational mortality data, *J Occup Med*, 1975, 17:128.
 24. Rosenberg L, Hennekens CH, Rosner B et al. Oral Contraceptive use in relation to non-fatal myocardial infarction. *Am J Epidemiol*, 1980, 111:59.
 ۲۵. آیت‌اللهی م ت؛ کارآزمایی‌های بالینی تألیف استوارت پوکاک. شیراز، انتشارات دانشگاه علوم پزشکی شیراز ۱۳۸۰.
 26. Fische B, Costuntino JP, Wicherham DL et al. Tamoxifen for prevention of breast cancer: Report of the National Surgical Adjuvant Breast and Bowel Project P-1 Study. *J Natl Cancer Inst*, 1998, 90. 1371-88.
 27. Begg C, Cho M, Eastwoods Horton R, Moher D, Olkin I, et al. Improving the quality of reporting of randomized controlled trials. The CONSORT statement. *J A M A*, 1996, 276: 637-9.
 28. Moher D, Schulz KF, Altman D. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials. *JAMA*, 2001, 285:1987-91.
 29. Friedman LM, Furberg C, Demets DL. Fundamentals of clinical trials. Springer, New York 1998.
 30. Smith PG, Marrow R. Field trials of health interventions in developing countries. Mac Milan, London 1996.
 31. Donner A, Klar N. Design and analysis of cluster randomization trials in health research. Arnold, London 2000.
۳۲. سوری ح و همکاران. اپیدمیولوژی نوین بیماری‌های عفونی تألیف جیسک یوهان، تهران: اداره کل پیشگیری و

مبارزه با بیماری‌ها، ۱۳۸۰.

33. Breslow NE, Day NE. Statistical methods in cancer research-volume1- The analysis of case-control studies. International Agency for Research on cancer, Lyon 1987.
 34. Schlesselman JJ, Stolley PD. Case-control studies: Design, conduct, analysis. Oxford University Press, New York 1982.
 35. Chow SC, Liu JP. Design and analysis of clinical trials: Concepts and methodology (2nd ed.). Wiley, New Jersey, 2004.
 36. Redmond CK, Biostatistics in clinical trials. Wiley, Chichester 2001.
 37. Piantadosi S. Clinical trials: A methodologic Perspective. Wiley, New York 1997.
 38. Kirkwood BR, Jonathan ACS. Medical statistics (2nd ed.). Blackwell Science, Massachusetts 2003.
 39. Rothman KJ. Epidemiology: An introduction. Oxford University Press, New York 2002.
 40. Szklo M, Nieto FJ. Epidemiology: Beyond the basics. Aspen, Maryland, 2000.
 41. Heinekens CH, Burring JE, Monson JE, Lack of long term supplementation with beta carotene on the incidence of malignant neoplasm's and cardiovascular disease. N Eng J Med, 1996, 334: 1149.
 42. Data from steering committee of the physician's health study research group: Final report on the aspirin component of the ongoing physician's health study. N Engl J Med, 1989, 321: 129-135.
 43. Dykes MHM, Meier P. Ascorbic acid and the common cold: Evaluation of its efficacy and toxicity. JAMA, 1975, 231:1073-9.
 44. Karlowski TR, Chalmers TC, Frenkel LD et al. Ascorbic acid for common cold: A prophylactic and therapeutic trial. JAMA, 1975, 231:1038-42.
 45. Franci ST, Napier JA, Voigh BS, et al. Evaluation of the 1954 field trial of Poliomyelitis vaccine: final report. Ann Arbor, MI: Poliomyelites Vaccine Evaluation Center, University of Michigan, 1957.
۴۶. میر میران پ، کیمیاگر م نظرآبادی م ط و همکاران. اثرات تزریق محلول روغنی یددار در دانش‌آموزان مبتلا به اختلالات ناشی از کمبود ید، بررسی سه‌ساله. پژوهش در پزشکی، سال دو شماره فروردین و خرداد ۱۳۷۶، ۶۴-۵۴،
۴۷. عزیزی ف و شیخ‌الاسلام ر. برنامه کشوری مبارزه با کمبود ید. طب و تزکیه شماره‌های ۱۹ و ۲۰، تابستان ۱۳۷۴، ۱۸-۲۲.
48. Akbari ME, Naghavi M, Soori H. Epidemiology of deaths from injuries in the Islamic Republic of Iran. East Mediterr Health J. 2006 May-Jul;12(3-4):382-90.

49. Svendsen KM, Kuller LH, Martin MJ, et al. Effects of passive smoking in the multiple risk factor intervention trial (MRFIT). *Am J Epidemiol*, 1987, 126:783-95.
 50. Mohammadi MR, Ghanizadeh A, Rahgozart M, Noorbala AA, Malekafzali H, Davidian H, Naghavi H, Soori H, Yazdi SA. Suicidal attempt and psychiatric disorders in Iran. *Suicide Life Threat Behav*. 2005 Jun;35(3):309-16.
 51. Soori H & Bhopal R.S. Parental permission for children's independent outdoor activities: Implications for injury prevention. *European Journal of Public Health*. 2002; 12: 104-109.
۵۲. حقدوست ع ا. مرور ساختاریافته و متآنالیز، مفاهیم، کاربردها و محاسبات. تهران، انتشارات تیمورزاده، بهار ۱۳۸۶.
۵۳. سوری ح و همکاران. نرم‌افزارهای مورد استفاده در کارگاه‌های متآنالیز. دانشگاه علوم پزشکی شهید بهشتی. ۱۳۸۶.
- (برای دسترسی مراجعه شود به سایت مدیریت پژوهشی دانشگاه www.sbm.ac.ir)



دقت، اعتبار و تکرارپذیری اندازه‌گیری‌ها و آزمون‌های تشخیصی

۴-۱ انواع خطاها در اندازه‌گیری

منابع خطا در انجام فرایند اندازه‌گیری ممکن است به دلایل زیر باشند.

اشتباه: اشتباهات به علت بی‌دقتی و ممکن الخطا بودن انسان رخ می‌دهند. اشتباه می‌تواند ناشی از استفاده ناصحیح از ابزار، خطای ثبت و یا خطای محاسبه باشد. ماهیت اشتباه غیرقابل تشخیص است. در واقع هنگامی که اشتباه به حداقل می‌رسد اندازه‌های به‌دست‌آمده درست‌تر هستند.

تصادف: خطاهای تصادفی به علت طبیعت اندازه‌گیری، بی‌ثباتی شرایط محیط، محدودیت‌های ابزار اندازه‌گیری و نقصان در مفروضات اتفاق می‌افتند. در این‌گونه خطاها مقادیر مثبت و منفی در درازمدت یکدیگر را خنثی می‌کنند.

نمونه‌گیری: خطاهای نمونه‌گیری بر اثر برآوردهای به‌دست‌آمده از نمونه‌گیری به وقوع می‌پیوندد. اگر نمونه‌گیری با روش‌های مبتنی بر احتمال انجام گیرد، خطاهای مثبت و منفی گرایش به خنثی کردن یکدیگر خواهند داشت. اندازه خطای نمونه‌گیری به‌وسیله واریانس جامعه و حجم نمونه گرفته شده برآورد می‌شود.

تورش: هرگونه تحریف و خطای منظم در اندازه‌گیری، باعث ایجاد تورش در نتایج می‌شود. خطاهای ناشی از تورش در درازمدت یکدیگر را خنثی نمی‌کنند. خطاهای منظم از نقص در ابزار و تجهیزات، اشکالات فنی یا سوگیری محقق سرچشمه می‌گیرند.

وقتی مجموع خطاهای اتفاق افتاده کوچک باشد، اندازه‌گیری‌ها و یا برآوردها را صحیح در نظر می‌گیریم؛ اما به‌طور دقیق و مشخص هیچ تعریف معینی در مورد طبقه‌بندی خطاها به رده‌های مختلف از نظر اندازه آن‌ها وجود ندارد. در مجموع خطا دارای دو مؤلفه خطای منظم و خطای تصادفی است. این فصل به نحوه ارزیابی، برآورد و کنترل این خطاها در اندازه‌گیری‌ها و آزمون‌های تشخیصی در مطالعات پزشکی اختصاص دارد.

۴-۲ گزارش علمی اندازه‌ها

یکی از اشتباهات گمراه‌کننده در نوشتن متون علمی، گزارش اندازه‌گیری‌ها به‌صورت اعداد مطلق هستند. چون

هیچ اندازه‌گیری عاری از خطا نیست بنابراین گزارش هراندازه باید شامل محدوده‌ای که انتظار داریم مقدار واقعی اندازه در آن قرار گیرد باشد. برای مثال، هنگامی که با یک خط کش ۳۰ سانتی‌متری چوبی مشغول اندازه‌گیری هستیم، بر اثر تغییر درجه حرارت محیط ممکن است تغییر طولی به‌اندازه ۳ میلی‌متر در طول خط‌کش به وجود آید؛ بنابراین هنگامی که طول یک شیء را ۳۰۰ میلی‌متر مشاهده می‌کنیم اندازه واقعی مقداری بین ۲۹۷ تا ۳۰۳ میلی‌متر است. در این مثال اگر گزارش اندازه‌گیری طول یک ضلع با خط‌کش چوبی بین مقادیر ۳۰۳-۲۹۷ میلی‌متر و گزارش اندازه‌گیری طول دیگر به‌وسیله یک خط‌کش فلزی به‌صورت مقادیر بین ۳۱۵-۲۸۵ میلی‌متر گزارش شود، مساحت مربعی که یک‌طرف آن ۳۰۳-۲۹۷ میلی‌متر و طرف دیگر آن ۳۱۵-۲۸۵ است باید به‌صورت ۹۵۴۴۵-۸۴۶۴۵ میلی‌متر مربع گزارش شود.

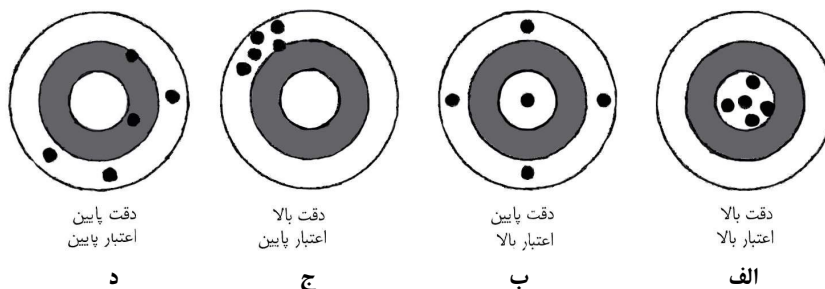
۴-۳ اعتبار و دقت در مطالعات بالینی

هر مطالعه کاربردی معمولاً شامل فرآیند اندازه‌گیری برخی پارامترها و جمع‌آوری اطلاعات مربوط به آنها است. خطای اندازه‌گیری یکی از عوامل مهمی است که محققان باید در مهار آن دقت و کوشش کافی به عمل آورند. متأسفانه در متون علمی واژه استاندارد برای بیان این مفاهیم وجود ندارد. اصطلاحات رایجی، اعتبار و صحت معادل هم و واژه‌های دقت، تکرارپذیری، سازگاری و پایایی معادل یکدیگرند و تنها بنا به سلیقه شخصی هرکدام از این واژه‌ها در متون علمی به‌کاربرده می‌شود.

یکی از بهترین روش‌ها برای توصیف این مفاهیم شکل (۴-۱) است. این شکل نتایج تیراندازی چهار فرد مختلف به چهار سیبل مشابه را نمایش می‌دهد شلیک‌های فرد "الف"، حول مرکز هدف و نزدیک هم جمع شده‌اند، با مشاهده این رخداد، می‌توان استدلال کرد که نشانه تفنگ به‌درستی تنظیم شده، بنابراین گلوله‌ها به‌جایی که نشانه‌گیری شده‌اند اصابت کرده‌اند و هم‌چنین تیرانداز "الف" فرد ماهری بوده که با دقت بالا نشانه‌گیری کرده است. این اطلاعات نشانگر دقت و اعتبار بالا در هدف‌گیری است. تیرانداز "ب"، شلیک‌های خود را در یک خوشه نزدیک به هم انجام نداده است. هرچند آشکارا به‌طور متوسط محل اصابت شلیک‌های او متمایل به مرکز هدف است. ممکن است چنین فرض شود که نشانه تفنگ به‌درستی تنظیم بوده ولی تیرانداز مهارت زیادی نداشته است که موجب پراکندگی شلیک‌ها شده است. این نتایج نشانگر دقت پایین و اعتبار بالا در هدف‌گیری است. از طرف دیگر فرد "ج" تیرانداز ماهری بوده زیرا شلیک‌های او نزدیک به هم است، ولی تمرکز دقیقی به هدف وجود ندارد، که احتمالاً به تنظیم نبودن دستگاه نشانه‌گیری تفنگ مربوط می‌شود. این حالت نشانگر بالا بودن دقت و پایین بودن اعتبار نشانه‌گیری‌ها است؛ اما فرد "د" نه تیرانداز ماهری بوده و نه دستگاه نشانه‌گیری تفنگ او به‌خوبی تنظیم شده است. زیرا شلیک‌های او غیر مجتمع و آشکارا دورتر از مرکز سیبل به تخته هدف برخورد کرده‌اند. این موضوع می‌تواند بیانگر دقت و اعتبار پایین در تیراندازی باشد، نشانه‌گیری‌های اشتباه می‌تواند با عواملی نظیر شدت و جهت وزش باد، عملکرد سیستم نشانه‌گیری اسلحه و یا شدت و صحت بینایی تیرانداز نیز مرتبط باشد. این عوامل می‌توانند باعث ایجاد خطای منظم در جهتی خاص شده و تیرها را از خال وسط سیبل دور گردانند.

تورش نتیجه یک خطای منظم است که در چنین حالتی ثبت دقیق اندازه‌ها به‌طور منظم و پیوسته متمایل

به مقادیر بالاتر (و یا پایین‌تر) از مقدار حقیقی دارد. واژه خطا در این زمینه فقط هنگامی استفاده می‌شود که مقدار حقیقی اندازه‌گیری معلوم باشد و عموماً از واژه تغییرات بجای آن استفاده می‌شود. دقت در اندازه‌گیری‌ها به میزان تغییرات تصادفی از یک نقطه ثابت (بدون توجه به اعتبار مقدار اندازه‌گیری) مرتبط است. یک اندازه‌گیری بی‌دقت گاهی اوقات بیشتر از یک اندازه ثابت است و گاهی اوقات کمتر از آن است. اگر خطا را به‌عنوان تفاضل مقدار مشاهده‌شده از مقدار حقیقی تعریف کنیم، آنگاه میانگین خطاهای تصادفی برابر صفر می‌شود. می‌توان چنین در نظر گرفت که خطاها دارای توزیع نرمال و مستقل از یکدیگر هستند.



شکل ۴-۱: نمایش اعتبار و دقت نشانه‌گیری به هدف

۴-۳-۱ دقت

اندازه‌گیری‌های دقیق تکرارپذیرند، در این صورت نتایج تقریباً یکسانی در هر آزمایش یا اندازه‌گیری ثبت می‌شود. استفاده از یک ترازوی دقیق در اندازه‌گیری وزن، باعث ثبت اندازه‌های دقیق می‌شود. در صورتی که استفاده از روش مصاحبه برای اندازه‌گیری کیفیت زندگی، فرایندی است که دستخوش تغییرات بسیار وسیع‌تر می‌گردد در نتیجه تکرارپذیری این روش کمتر است. دقت یکی از عوامل مهم در توان آماری هر مطالعه است. دقت اندازه‌گیری متأثر از خطاهای تصادفی (شانس) است. هر قدر خطاهای تصادفی بیشتر و مؤثرتر باشند، دقت نیز به همان نسبت کاهش پیدا می‌کند.

به‌طور کلی سه عامل ابزار اندازه‌گیری، مشاهده‌گر و تغییرات ذاتی اعضای مورد مطالعه، در ایجاد تغییرات تصادفی دخالت دارند.

تغییرات ذاتی اعضای مورد مطالعه: این تغییرات ناشی از تغییر لحظه‌ای بیولوژیکی در افراد است. خلق و خو و آمادگی جسمانی که مدام دستخوش دگرگونی و تغییرند، از این دست به شمار می‌روند.

تغییرات مشاهده‌گر به تغییراتی در اندازه‌گیری که ناشی از نوسانات مشاهده‌گر هستند نظیر واژه‌های به‌کاررفته در مصاحبه‌ها و یا میزان مهارت اپراتور در به‌کار بردن دستگاه‌ها، مربوط می‌شود.

تغییرات ابزار اندازه‌گیری که به تغییرات ناشی از وسایل اندازه‌گیری بر اثر عواملی مانند تغییر درجه حرارت، ارتفاع، استهلاک مرتبط می‌شود.

بنابراین فرایند اندازه‌گیری ابزارها غیرقابل اعتماد است اگر:

محقق ویژگی یک فرد را در شرایط گوناگون اندازه‌گیری کند و نتایج متفاوتی به دست آورد. (تغییرات مشاهده‌گر) محققان مختلف با ابزار اندازه‌گیری یکسان نتایج متفاوتی راجع به هر عضو مطالعه به دست آورند. (تغییرات ابزار اندازه‌گیری)

فرد مورد مطالعه (پاسخگو یا کمیت مورد اندازه‌گیری) در شرایط مختلف به یک آزمون، پرسشنامه یا ابزار اندازه‌گیری، پاسخ‌های متفاوتی ارائه دهد. (تغییرات ذاتی افراد تحت مطالعه).

۴-۳-۱- ارزیابی دقت نتایج

ارزیابی دقت تنها در اثر تکرار آزمایش‌ها امکان‌پذیر است. دقت به‌وسیله سنجش سازگاری و همخوانی نتایج به‌دست‌آمده از تکرار اندازه‌گیری‌ها قابل ارزیابی است. برای دستیابی به این هدف ممکن است یک مطالعه تجربی با ثابت نگه‌داشتن همه شرایط و تغییر یکی از سه منبع تغییر، ذکر شده در بالا طراحی شود. در این صورت چهار حالت زیر قابل بررسی خواهد بود.

- ارزیابی میزان تکرارپذیری نتایج هر مشاهده‌گر

هر مشاهده‌گر (کاربر) یک پارامتر مشخص را در دفعات معینی بر روی یک مجموعه نمونه اندازه‌گیری می‌کند. (Within-Observer reproducibility)

- ارزیابی میزان تکرارپذیری نتایج چند مشاهده‌گر

یک گروه از مشاهده‌گرها (اپراتورها) یک پارامتر مشخص را دفعات معینی بر روی یک مجموعه نمونه اندازه‌گیری می‌کنند. (Between-Observer reproducibility)

- ارزیابی میزان تکرارپذیری ابزار اندازه‌گیری

در این حالت مشاهده‌گر به‌وسیله یک ابزار اندازه‌گیری، پارامتر موردنظر را به‌دفعات معین در یکسری نمونه اندازه‌گیری می‌کند. (Within-Instrument reproducibility)

- ارزیابی میزان تکرارپذیری ابزارهای مختلف اندازه‌گیری

فرایند اندازه‌گیری یک پارامتر مشخص در یک مجموعه نمونه به‌دفعات معین با ابزارهای مختلف اندازه‌گیری تکرار می‌شود. (Between-Instrument reproducibility)

۴-۳-۲- راهبردهای مختلف برای بالا بردن دقت

پنج روش برای کاهش خطاهای تصادفی و بالا بردن دقت اندازه‌گیری وجود دارند که عبارت‌اند از:

۱- استاندارد کردن روش‌های اندازه‌گیری: پروتکل هر مطالعه باید شامل تعاریف علمی و دقیق در مورد

پارامترهای مطالعه باشد. برای مثال باید دستورالعمل‌های خاص برای انجام اندازه‌گیری‌ها وجود داشته باشد. این تعاریف باید شامل راهنمای مکتوب برای نحوه آماده کردن شرایط محیط، اعضای مورد مطالعه، چگونگی ثبت و نگهداری نتایج اندازه‌گیری‌ها و مصاحبه‌ها، چگونگی تنظیم دستگاه‌ها و غیره باشد. چنین دستورالعمل‌هایی برای انجام مطالعات گسترده و پیچیده الزامی است و در مطالعاتی با ابعاد کوچک‌تر توصیه

به فراهم آوردن این دستورالعمل‌ها وجود دارد. حتی در مواردی که تنها یک پارامتر برای اندازه‌گیری وجود دارد، فراهم کردن یک دستورالعمل مختصر برای هر دستگاه اندازه‌گیری کمک بسیار مؤثری در یکنواخت شدن نتایج می‌کند.

۲- آموزش و تأیید مهارت‌های مشاهده‌گر: آموزش باعث بهبود و یکنواختی شیوه‌های اندازه‌گیری می‌شود. اهمیت این موضوع، هنگامی که نیاز به استفاده از چند مشاهده‌گر وجود دارد، بسیار زیادتر است.

۳- بهبود ابزارهای اندازه‌گیری: استفاده از وسایل مکانیکی یا الکترونیکی دقیق معمولاً باعث کاهش تغییرات می‌شود. هم‌چنین در هنگام به‌کارگیری مصاحبه و پرسشنامه، استفاده از فرم‌های مکتوب و روش‌های استاندارد باعث کاهش نکات نامعلوم و گیج‌کننده و افزایش شفافیت اطلاعات به‌دست‌آمده می‌شود.

۴- استفاده از وسایل خودکار: آن دسته تغییراتی که ناشی از خطای مشاهده‌گر در اندازه‌گیری است با استفاده از وسایل خودکار و یا پرسشنامه‌هایی که به‌وسیله افراد تحت مطالعه تکمیل می‌شود، قابل جلوگیری است. به‌طورقطع استفاده از این روش تنها هنگامی که ابزارهای خودکار، دارای دقت بالاتری از اندازه‌گیری توسط انسان باشد، باعث کاهش تغییرات خواهد شد.

۵- تکرار: اثر خطاهای تصادفی بدون توجه به منبع آن با تکرار اندازه‌گیری و استفاده از میانگین مشاهدات بجای اندازه‌های تکی قابل کاهش است. اساساً استفاده از این روش باعث افزایش دقت می‌شود. تنها محدودیت‌های این روش هزینه‌های احتمالی و مشکلات عملی در تکرار اندازه‌گیری‌ها است.

در فرایند اندازه‌گیری، محققان باید در نحوه استفاده از هرکدام از این پنج روش برای تضمین دقت اندازه‌گیری‌ها به توافق برسند. تصمیم‌گیری در این زمینه به اهمیت متغیر موردنظر، عوامل بالقوه مؤثر در بروز خطاهای تصادفی در مطالعه، هزینه و انجام‌پذیر بودن هر راهبرد بستگی دارد.

به‌طورکلی روش‌های استاندارد کردن و آموزش همیشه به‌کاربرده می‌شود و راهبرد تکرار گزینه‌ای است که در صورت امکان‌پذیر بودن، دقت بالا را تضمین می‌کند.

در جدول ۱-۴ راهبردهای مختلف برای کاهش خطای تصادفی و افزایش دقت در فرآیند اندازه‌گیری فشارخون که در یک مطالعه به‌منظور درمان پرفشاری خون امکان دارد استفاده شود، فهرست شده است.

جدول ۴-۱: راهبردهای مختلف برای بالا بردن دقت

مثال‌هایی از نحوه اجتناب از خطا	مثال‌هایی از خطای تصادفی	منبع خطای تصادفی	راهبرد برای کاهش خطای تصادفی
زمان فروکش کردن کاف دقیقاً در نقطه ۲ mmHg/sec تعیین شود.	تغییراتی در اندازه فشارخون که ناشی از ثبت مقدار فشارخون در زمان‌های مختلف پس از فروکش کردن کاف است (گاهی اوقات سریع‌تر و گاهی اوقات کندتر از ۲ mmHg/sec).	مشاهده‌گر	۱- استاندارد کردن روش‌های اندازه‌گیری در روند تحقیق
فراهم کردن شرایطی که هر فرد قبل از اندازه‌گیری فشارخون بتواند ۵ دقیقه در یک اتاق خلوت استراحت کند.	تغییر در اندازه فشارخون که از تغییر مدت‌زمان استراحت هر فرد قبل از اندازه‌گیری به وجود می‌آید.	افراد تحت مطالعه	۲- آموزش و تأیید مشاهده‌گر
به تمام کاربرها تکنیک یکسانی آموزش داده شود.	تغییر در اندازه فشارخون که از شیوه‌های متفاوت هر مشاهده‌گر نشئت می‌گیرد.	مشاهده‌گر	۳- بهبود وسایل اندازه‌گیری
استفاده از جدول اعداد تصادفی برای گرد کردن رقم آخر مشاهده‌شده.	تغییر در اندازه فشارخون که از ترجیح به گرد کردن به ارقام خاص حاصل می‌شود (گرد کردن به ۰ یا ۵ در هنگام ثبت).	وسيله اندازه‌گیری یا مشاهده‌گر	۴- اتوماتیک کردن وسایل اندازه‌گیری
استفاده از فشارسنج‌های دیجیتالی برای ثبت نتایج	تغییر در اندازه فشارخون به علت خطای فردی در به‌کارگیری خواندن یا ثبت اندازه‌ها	مشاهده‌گر	۵- تکرار اندازه‌گیری
استفاده از میانگین فشارخون به‌دست‌آمده از دو یا چند اندازه‌گیری.	تمام منابع خطای تصادفی	مشاهده‌گر، افراد تحت مطالعه، ابزار اندازه‌گیری	

۴-۳-۲ اعتبار

هر پژوهش به داده‌های دقیق، تکرارپذیر و قابل‌اعتماد نیاز دارد؛ اما این ویژگی‌ها هیچ تضمینی برای باکیفیت بودن یا اساساً مرتبط بودن داده‌ها با اهداف تحقیق ندارند. برای مثال کودکانی که دارای اختلالات بینایی یا شنوایی هستند ممکن است نمره‌های پایینی در یک آزمون استاندارد سنجش هوش کسب کنند. به‌طور یقین مواد و روش‌های نامناسب نمی‌تواند منجر به نتایج منصفانه‌ای شود. برای مثال اگر بیماری گواتر با لمس گلوی افراد تحت مطالعه به‌وسیله دست اندازه‌گیری شود نتایج اندازه‌گیری معتبر نخواهد بود.

اعتبار تابعی از خطاهای منظم است. هر قدر خطای منظم بیشتر باشد اعتبار نتایج نیز، ضعیف‌تر می‌شود. سه عامل عمده خطای منظم عبارت‌اند از:

- ♦ **تورش مشاهده‌گر (Observer bias):** هرگونه تحریف آگاهانه یا ناآگاهانه در فراهم‌سازی یا گزارش مشاهدات به‌وسیله محقق مانند استفاده از سوالات جهت‌دار، داوری سوگرایانه و غیره.

- ♦ **تورش اعضای نمونه (Subject bias):** تحریف اندازه‌گیری‌ها که به‌وسیله افراد تحت مطالعه صورت می‌گیرد. برای مثال خانم‌هایی که مبتلا به سرطان پستان شده‌اند و از این موضوع که مصرف بیش‌ازاندازه چربی یکی از عوامل خطر مرتبط با این سرطان است اطلاع دارند، ممکن است در هنگام مصاحبه به‌طور اغراق‌آمیزی میزان مصرف چربی در الگوی غذایی دوران جوانی خود را زیاد گزارش کنند.
 - ♦ **تورش ابزار اندازه‌گیری (Instrument bias):** این خطا می‌تواند از علل ساختاری ابزار مورداندازه‌گیری ناشی شود. مثلاً یک ترازو که به‌تازگی کنترل و تنظیم‌نشده، ممکن است به‌طور ثابت وزن افراد تحت مطالعه را کمتر یا بیشتر از حد واقعی نمایش دهد.
- جدول ۴-۲ تفاوت‌های موجود بین اعتبار و دقت را نمایش می‌دهد. بار دیگر تأکید می‌کنیم که اعتبار و دقت لزوماً با یکدیگر وابسته نیستند.

جدول ۴-۲: اعتبار و دقت در اندازه‌گیری یک ویژگی

اعتبار (Accuracy)	دقت (Precision)	تعریف
معیاری که معین می‌کند اندازه‌گیری‌های به‌عمل‌آمده از یک پارامتر تا چه اندازه صحت دارد. مقایسه با یک روش استاندارد	معیاری که معین می‌کند یک صفت در تکرار دفعات اندازه‌گیری تا چه اندازه تغییرپذیری دارد. مقایسه اندازه‌های تکراری	مهم‌ترین روش ارزیابی
بالا بردن صحت نتایج خطای منظم ناشی از: مشاهده‌گر افراد تحت مطالعه وسایل اندازه‌گیری	بالا بردن توان آماری در تشخیص اثرها خطای تصادفی ناشی از: مشاهده‌گر افراد تحت مطالعه وسایل اندازه‌گیری	ارزش بررسی عوامل تهدیدکننده

۴-۳-۱ ارزیابی اعتبار نتایج

مهم‌ترین روش برای ارزیابی اعتبار اندازه‌گیری‌ها، مقایسه نتایج حاصل با اندازه‌های به‌دست‌آمده از یک روش استاندارد مرجع است، که از صحت نتایج آن مطمئن هستیم. در اندازه‌هایی که دارای مقیاس پیوسته هستند، میانگین اختلاف بین روش موردبررسی و روش استاندارد مرجع قابل‌محاسبه است. در اندازه‌هایی که دسته‌بندی شده هستند، مقایسه با روش استاندارد مرجع از طریق محاسبه حساسیت و ویژگی امکان‌پذیر است. هنگامی که روش استاندارد وجود ندارد، معمولاً برای ارزیابی صحت تکنیک جدید محققان فرض را بر صحیح بودن شیوه‌های متداول قرار می‌دهند.

۴-۳-۲ راهبردهای مختلف برای بالا بردن اعتبار نتایج

روش‌های اصلی برای افزایش اعتبار نتایج شامل چهار راهبرد اول فهرست شده در افزایش دقت به همراه سه روش دیگر هستند که در اینجا شرح داده می‌شوند:

۱- استاندارد کردن روش‌های اندازه‌گیری

۲- آموزش و تأیید مهارت‌های مشاهده‌گر

۳- بهبود ابزارهای اندازه‌گیری

۴- استفاده از وسایل خودکار

۵- **اندازه‌گیری‌های نامحسوس:** گاهی اوقات اندازه‌گیری برخی ویژگی‌ها بدون آگاهی اعضای مورد مطالعه امکان‌پذیر است. این موقعیت می‌تواند باعث حذف برخی از منابع بالقوه تورش شود. برای مثال هنگام بررسی الگوی تغذیه کودکان دبستانی ممکن است به‌جای پرسش مستقیم از کودکان، تعداد کاغذهای شکلات موجود در سطل آشغال‌ها در پایان هرروز شمارش شود.

۶- **کور کردن:** این راهبرد صحت کلی اندازه‌گیری‌ها را افزایش نمی‌دهد بلکه می‌تواند باعث حذف تورش ترجیحاً در نحوه ارزیابی افراد تحت مطالعه شود. در یک مطالعه دوسویه کور فرد تحت مطالعه و تیم معالج از محتوی داروهای مصرف‌شده (یا حتی دارونما) اطلاع ندارند. در مطالعات مشاهده‌ای نیز امکان استفاده از این رویه برای افرادی که مسئول اندازه‌گیری پیامد موردنظر هستند وجود دارد.

۷- **تنظیم دستگاه‌های اندازه‌گیری:** اعتبار نتایج بسیاری از ابزارهای اندازه‌گیری به‌خصوص وسایل مکانیکی و الکتریکی در صورت تنظیم دوره‌ای، همواره در حد مطلوب باقی می‌ماند.

رعایت دو راهبرد استاندارد کردن و آموزش اجتناب‌ناپذیر هستند. هر زمان که امکان انجام رویه کورسازی وجود دارد، اجرای آن الزامی است و تنظیم دوره‌ای دستگاه‌هایی که به‌طور بالقوه باگذشت زمان حساسیت خود را از دست می‌دهند نیز از موارد اساسی است.

در جدول ۳-۴ راهبردهای مختلف برای کاهش خطاهای منظم و افزایش صحت در فرایند اندازه‌گیری فشارخون که در یک مطالعه به‌منظور درمان پرفشاری خون قابل‌استفاده است فهرست شده‌اند.

جدول ۴-۳: راهبردهای مختلف در کاهش خطای منظم به منظور افزایش اعتبار

راهبرد برای کاهش خطای منظم	منبع خطای منظم	مثال‌هایی از خطای منظم	مثال‌هایی از راهبردهای اتخاذ شده برای اجتناب از خطا
۱- استاندارد کردن روش‌های اندازه‌گیری در فرایند تحقیق	مشاهده‌گر	یکی از کاربرها به‌طور منظم هنگام اندازه‌گیری فشارخون دیاستولیک، درجه را هنگام قطع شدن صدا ثبت کند.	تعریف عملیاتی فشارخون دیاستولیک به‌طوری که درجه فشارسنج در لحظه شنیدن صدای رگ ثبت شود.
۲- آموزش و تأیید مهارت مشاهده‌گر	مشاهده‌گر	ثبت بیش‌ازاندازه فشارخون به علت تبعیت نکردن از دستورالعمل‌های اندازه‌گیری در طرح تحقیق	آموزش کاربر و تأیید صحت کار به‌وسیله یک فشارسنج دو عقربه‌ای
۳- بهبود دقت اندازه‌گیری	افراد تحت مطالعه	ثبت بیش‌ازاندازه فشارخون دیاستولیک، به علت اینکه اعضای شرکت‌کننده در طرح از پله‌های بیمارستان بالا می‌روند.	در نظر گرفتن یک اتاق خلوت برای افراد شرکت‌کننده در طرح به‌طوری که بتوانند ۵ دقیقه در آن استراحت کنند.
	ابزار اندازه‌گیری	ثبت بیش‌ازاندازه فشارخون به علت استفاده کاف استاندارد در افرادی که بازوی بسیار بزرگ دارند.	استفاده از کاف بزرگ‌تر در افراد چاق
۴- اتوماتیک کردن ابزار اندازه‌گیری	مشاهده‌گر	تمایل آگاهانه یا ناآگاهانه مشاهده‌گر برای ثبت فشارخون پایین‌تر از مقدار حقیقی در گروه درمان جدید	استفاده از فشارسنج دیجیتالی که نتایج را به‌طور اتوماتیک ثبت می‌کند.
	افراد تحت مطالعه	افزایش فشارخون به علت اضطراب پاسخگو در برخورد با کادر پزشکی	استفاده از فشارسنج دیجیتالی تا افراد تحت مطالعه به‌تنهایی فشارخون را اندازه‌گیری کنند.
۵- اندازه‌گیری نامحسوس	افراد تحت مطالعه	تمایل به اغراق در مصرف داروی فشارخون	اندازه‌گیری میزان داروی مصرفی در ادرار
	مشاهده‌گر	تمایل آگاهانه یا ناآگاهانه در ثبت کمتر از واقعی فشارخون در گروه درمان جدید	استفاده از یک رویه دوسوکور که در آن کاربرها و افراد تحت مطالعه از نوع داروی استفاده شده بی‌اطلاع باشند.
۶- کور کردن	افراد تحت مطالعه	تمایل افراد تحت مطالعه به اغراق درباره اثرات جانبی داروی جدید	استفاده از یک رویه دوسوکور که در آن کاربرها و افراد تحت مطالعه از نوع داروی استفاده شده بی‌اطلاع باشند.
۷- تنظیم ابزار اندازه‌گیری	ابزار اندازه‌گیری	اندازه‌گیری بیش از مقدار واقعی فشارخون به علت عدم تنظیم دستگاه فشارخون	کنترل و تنظیم مستمر دستگاه اندازه‌گیری فشارخون

۴-۳-۳ اعتبار داخلی و اعتبار خارجی

مناسب بودن روش‌ها و میزان اعتماد به هر پژوهش به وسیله ارزیابی اعتبار داخلی و خارجی آن مشخص می‌شود. اعتبار داخلی میزان صحت و درستی استنباط‌هایی است که می‌توان در خصوص رابطه سببی در مطالعه انجام داد. به عبارت دیگر اعتبار داخلی یک پژوهش زمانی بالا است که بتوانیم مطمئن شویم اثر مشاهده‌شده به‌طور کامل مربوط به مداخله اعمال شده است نه متغیرهای مزاحم (مخدوش‌کننده) دیگر.

اعتبار خارجی مربوط به قابلیت تعمیم‌یافته‌های تحقیق به جامعه هدف، در شرایط معمولی است. بالا بودن اعتبار داخلی، مستلزم ایجاد کردن شرایطی مصنوعی و غیرطبیعی (آزمایشگاهی) و یا محدود شدن به موارد خاص در پژوهش است این موضوع ممکن است باعث ناتوانی محقق در تعمیم‌یافته‌های تحقیق در شرایط عادی شود. هنگام ارزیابی اعتبار یک ابزار سنجش توجه به نکات زیر می‌تواند مهم باشد.

- ♦ **اعتبار ظاهری:** اندازه‌گیری‌ها باید مرتبط و شبیه به همان چیزی باشند که ادعای اندازه‌گیری آن شده است. البته این موضوع هیچ تضمینی مبنی بر علمی بودن آن روش نیست.
- ♦ **اعتبار محتوا:** ویژگی‌های اندازه‌گیری شده باید نماینده مناسبی از جامعه هدف باشند. به‌عنوان مثال هنگامی که قصد داریم میزان فعالیت بدنی روزانه در بیماران را بررسی کنیم، اعضای نمونه نمی‌توانند تنها بیمارانی که قادر به راه رفتن هستند، باشند. یک مثال دیگر، سنجش کیفیت زندگی است. اندازه‌گیری چنین کیفیتی باید شامل پرسش‌هایی راجع به عملکردهای اجتماعی، بهداشتی، عاطفی و ذهنی باشد.
- ♦ **اعتبار ساختاری:** وسیله اندازه‌گیری باید با مفروضات نظری و تئوری‌های علمی موجود مطابقت داشته باشد. برای مثال اگر اعتقاد داریم داروی جدید به علت ایجاد تغییرات در یکسری از ویژگی‌ها مؤثرتر از داروی استاندارد است، ابزار اندازه‌گیری به کار برده شده باید توانایی اندازه‌گیری این اختلاف را دارا باشد.
- ♦ **اعتبار تجربی:** نتایج به‌دست‌آمده از فرآیند اندازه‌گیری باید با نتایج به‌دست‌آمده به‌وسیله روش‌های اندازه‌گیری مشابه همخوانی داشته باشد.

۴-۳-۴ انتخاب روش‌های مناسب برای اندازه‌گیری متغیرهای تحقیق

آزمایش‌های پزشکی نظیر آن‌هایی که برای غربالگری عوامل خطر، تشخیص یک بیماری و یا پیش‌بینی روند بیماری استفاده می‌شوند از موضوعات بسیار مهم تحقیقات پزشکی هستند. این آزمایش‌ها که معمولاً پرهزینه هستند نقش اساسی در درمان و پیش‌گیری ایفا می‌کنند. به همین علت، به معیارهایی جهت انتخاب مناسب هر کدام از این روش‌ها نیازمندیم. جدول (۴-۴) سؤال‌های مهمی که میزان مفید بودن هر آزمایش پزشکی را مشخص می‌کند را به همراه طرح‌های مطالعاتی مناسب برای پاسخ به این سؤالات و شاخص‌های آماری مرتبط با هر کدام نشان می‌دهد.

جدول ۴-۴: نحوه ارزیابی آزمایش‌های پزشکی

سؤال	طرح مطالعه احتمالی	شاخص‌های آماری برای بیان نتایج
نتایج به‌دست‌آمده تا چه اندازه تکرارپذیر هستند؟	مطالعاتی که تغییرات درونی یا بین مشاهده‌گرها و یا تغییرات درونی یا بین دستگاه‌ها را ارزیابی می‌کند.	نسبت توافق، ضریب کاپا، ضریب تغییرات، میانگین و انحراف معیار اختلاف‌ها
آزمایش‌ها تا چه اندازه معتبر هستند؟	مطالعات مقطعی، مورد-شاهدی و یا هم‌گروهی که در آن‌ها نتایج آزمایش با روش استاندارد مرجع مقایسه می‌شوند.	حساسیت، ویژگی، ارزش اخباری مثبت و منفی، منحنی راک و نسبت درستی‌نمایی
آزمایش‌های انجام گرفته تا چه اندازه تصمیم‌های بالینی را تحت تأثیر قرار می‌دهند؟	ارزیابی صحت تشخیص‌های افتراقی به کمک آزمایش‌ها	نسبت آنومالی‌ها، نسبت نتایج ناهم‌هنگ، نسبت نتایج کلینیکی تغییر یافته به کمک آزمایش، هزینه هر مورد آنومالی و یا تغییر در هر تصمیم کلینیکی
آیا انجام آزمایش باعث بهبود پیامدهای بالینی می‌شود یا تنها دارای عوارض جانبی است؟	کارآزمایی‌های تصادفی شده، مطالعات هم‌گروهی یا مورد-شاهدی که متغیر مداخله‌گر، انجام یا عدم انجام آزمایش موردنظر و متغیر پاسخ شامل مرگ‌ومیر، بیماری‌زایی و هزینه مربوط به بیماری یا آزمایش جدید می‌شود.	نسبت خطر، نسبت شانس، نسبت مخاطره، تعداد افرادی که نیاز به درمان دارند، نسبت و میزان پیامدهای جانبی

نکات مهم برای طراحی مطالعاتی که هدف آن‌ها مقایسه آزمون‌های پزشکی است در این قسمت بررسی می‌شوند. اولین نکته در طرح این مطالعات بررسی نتایج آزمون برای طیف وسیعی از افراد که شامل افراد سالم تا افرادی با بیماری شدید می‌شود. از آنجاکه اکثر این مطالعات توصیفی هستند (در مراحل اولیه)، نحوه انتخاب اعضای نمونه تأثیر بسیار زیادی در نتایج دارد. به همین علت اگر افراد نمونه (سالم یا بیمار) در ویژگی‌های مهمی با جامعه هدف متفاوت باشند، نتایج حاصل دارای تورش می‌شود. به‌عنوان مثال امکان دارد اعضای بیمار در نمونه، از میان بیمارانی با وضعیت حاد که آسان‌تر قابل تشخیص هستند انتخاب شوند و افراد گروه شاهد (بدون بیماری) در نمونه از میان افراد سالم‌تر جامعه انتخاب شوند. در این صورت احتمالاً آزمون‌های تشخیصی زیادی دارای نتایج ایده‌آل هستند. نتایج گمراه‌کننده هنگامی بروز پیدا می‌کنند که مسئله تشخیص بین دو بیماری با علائم و نشانه‌ها مشابه مطرح باشد و یا تشخیص میان یک فرد سالم و یک فرد که دارای علائم خفیف اولیه هنگام ظهور بیماری است، مدنظر باشد. به همین علت افراد تحت مطالعه در چنین طرح‌هایی باید از تمام طیف بیماری انتخاب شوند تا قابلیت‌های آزمون از لحاظ بالینی مشخص شود.

نکته بعدی توجه به منابع تغییرات، قابلیت تعمیم نتایج و طرح نمونه‌گیری در مطالعه است. در برخی از سؤال‌های پژوهش منبع اصلی تغییرات، افراد تحت مطالعه (بیماران) هستند. به‌عنوان مثال گلوبول‌های سفید خون در نوزادانی که در خون آن‌ها عفونت باکتریایی وجود دارد افزایش پیدا می‌کند. تشخیص این مطلب به احتمال بسیار زیاد وابستگی چندانی به شخصی که از خون نمونه‌برداری می‌کند و یا آزمایشگاهی که تعداد گلوبول‌ها را اندازه‌گیری می‌کند، ندارد. در عوض در بسیاری از آزمون‌ها، نتایج به میزان بسیار زیاد به شخصی که آزمایش‌ها را

انجام می‌دهد و یا افرادی که نتایج را تفسیر می‌کنند، بستگی دارد. به‌عنوان مثال حساسیت، ویژگی و اعتبار تفسیر نتایج حاصل از ماموگرافی، شدیداً به کیفیت تجهیزات مورد استفاده و مهارت پاتولوژیست وابسته است. تکرار آزمایش به‌وسیله چند متخصص یا تفسیر نتایج آزمایشگاهی به‌وسیله چند متخصص، باعث افزایش تعمیم‌پذیری نتایج مطالعاتی که نیاز به مهارت و تجربه دارند می‌شود. هنگامی که اعتبار یا هزینه از مؤسسه‌ای به مؤسسه دیگر متفاوت است برای دستیابی به نتایج قابل تعمیم، لازم است نمونه‌ای از این مؤسسات در تحقیق مورد بررسی قرار گیرد.

مسئله مهم دیگر در طراحی این مطالعات، توجه به کور بودن آزمودنی‌ها است. محققان باید تمام تلاش خود را در مخفی نگه داشتن سایر نتایج کلینیکی و پاراکلینیکی مربوط به تشخیص بیماری در هنگام انجام آزمایش و تفسیر نتایج آن انجام دهند. مثلاً هنگامی که هدف مطالعه ارزیابی نتایج سونوگرافی در تشخیص آپاندیسیت است، افرادی که تفسیر نتایج را بر عهده دارند، نباید از سوابق قبلی بیمار و نتایج معاینات بالینی مطلع باشند. به همین ترتیب اشخاصی که وظیفه تصمیم‌گیری نهایی در مورد وجود یا عدم وجود آپاندیسیت در بیماران را بر عهده دارند (تشخیص نهایی کاملاً معتبری که نتایج سونوگرافی با آن مقایسه می‌شود)، نباید از نتایج مربوط به آزمایش سونوگرافی مطلع باشند. کور کردن، تورش‌ها، زمینه‌های ذهنی و اطلاعات اضافی را که می‌تواند در تصمیم‌گیری راجع به نتایج آزمون تأثیرگذار باشد را حذف می‌کند.

۴-۴ طراحی مطالعاتی که تکرارپذیری آزمون‌ها را بررسی می‌کنند

گاهی اوقات نتایج آزمون بر اساس اینکه کجا یا در چه زمانی آزمایش‌ها انجام شده‌اند و یا چه کسی آن‌ها را انجام داده است متغیر هستند. تغییر در نتایج، میزان تکرارپذیری آزمون‌ها را هنگامی که نتایج یک نمونه به‌وسیله یک کاربر یا آزمایشگاه در دفعات مختلف تکرار می‌شود، مشخص می‌کند. برای مثال اگر یک رادیولوژیست عکس‌های رادیوگرافی قفسه سینه را دو بار در زمان‌های مختلف تفسیر کند چه نسبتی از عکس‌ها دارای تفسیرهای مشابه در دومرتبه هستند؟ و یا اگر عکس‌های یکسانی به دو رادیولوژیست جهت تفسیر ارائه شود چه نسبتی از تفسیرها مشابه هستند؟

هنگامی که تکرارپذیری بین مشاهده‌گرها و یا یک مشاهده‌گر به‌تنهایی، ضعیف باشد آزمون‌های تشخیصی بر این مبنا ناکارآمد هستند. آزمون‌های که برای ارزیابی تکرارپذیری طراحی می‌شوند، نیاز به یک آزمون مرجع استاندارد برای تشخیص صحت نتایج ندارند. بدین معنی که دو مشاهده‌گر (یا تمام مشاهده‌گرها) می‌توانند به روی یک تفسیر اشتباه اتفاق نظر داشته باشند در این صورت هم نتایج تکرارپذیر است. به همین علت تکرارپذیر بودن نتایج، دقت را تضمین می‌کند ولی ارتباطی با اعتبار نتایج ندارد. مطالعات مقطعی برای مقایسه نتایج آزمون‌ها بین دو یا چند مشاهده‌گر و یا مقایسه نتایج یک مشاهده‌گر در موقعیت‌های متفاوت به کار می‌رود.

۴-۴-۱ تحلیل

در بخش قبل راجع به نحوه طراحی مطالعاتی که تکرارپذیری آزمون‌ها را بررسی می‌کنند توضیح داده شد. در این

قسمت برخی شاخص‌های آماری برای بیان میزان تفاوت بین نتایج حاصل از این مطالعات بررسی می‌شوند.

۴-۱-۱-۴ متغیرهای اسمی

ساده‌ترین معیار برای بیان توافق میان مشاهده‌گرها، نسبت نتایج موافق به کل مقایسه‌ها است. این نسبت را میزان توافق می‌نامند. با این حال هنگامی که بیش از دو رده موجود باشد تفسیر میزان توافق پیچیده می‌شود. معیار مناسب‌تر برای بیان میزان توافق ضریب کاپا است. این معیار میزان توافق خالص که برای مشاهده توافقی‌های ناشی از تصادف تعدیل شده است را بیان می‌کند. ضریب کاپا از ۱- به عنوان عدم توافق کامل تا ۱+ به عنوان توافق کامل تغییر می‌کند. کاپای برابر صفر نشان‌دهنده وجود توافقی‌های کاملاً تصادفی است [۱]. این معیار را با حروف یونانی κ نمایش می‌دهند. اگر A_{obs} نشان‌دهنده نسبت توافق بین دو مشاهده‌گر و A_{exp} ، نسبت توافق‌های حاصل از شانس باشد، ضریب κ به صورت زیر محاسبه می‌شود:

$$\kappa = \frac{A_{obs} - A_{exp}}{1 - A_{exp}}$$

ضریب کاپا بیشتر از ۰/۷۵، بیانگر توافق بسیار بالا است. ضریب کاپا بین ۰/۷۵ تا ۰/۴۵ نشان‌دهنده توافق متوسط و ضریب کاپا کمتر از ۰/۴۵ بیانگر توافق ضعیف است. جدول ۴-۵ خلاصه‌ای از نتایج عکس‌های ماموگرافی ۱۰۰ نفر را (زلحاظ طبیعی یا غیرطبیعی بودن) که توسط پاتولوژیست مختلف تفسیر شده، نمایش می‌دهد.

جدول ۴-۵: تفسیر نتایج عکس‌های پاتولوژی توسط دو پاتولوژیست

پزشک شماره ۲		پزشک شماره ۱	
مجموع	طبیعی	غیرطبیعی	مجموع
۳۵	۱۵	۲۰	غیرطبیعی
۶۵	۵۵	۱۰	طبیعی
۱۰۰	۷۰	۳۰	مجموع

ضریب کاپا برای تلخیص نتایج مربوط به میزان توافق بین دو پاتولوژیست به صورت زیر محاسبه می‌شود:

$$A_{exp} = \left[\frac{\frac{30 \times 35}{100} + \frac{70 \times 65}{100}}{100} \right] = 0/56$$

$$A_{obs} = \frac{20 + 55}{100} = 0/75$$

$$\kappa = \frac{0/75 - 0/56}{1 - 0/56} = \frac{0/19}{0/44} = 0/43$$

۴-۱-۲ متغیرهای پیوسته

معیارهای بیان‌کننده تغییرات بین مشاهده‌گرها در متغیرهای پیوسته تا حدود زیادی به طراحی مطالعه بستگی دارد. برخی از مطالعات تنها برای ارزیابی میزان توافق بین دو دستگاه یا دو روش (برای مثال درجه حرارت‌های به‌دست‌آمده از دو دماسنج مختلف) طراحی می‌شوند. بهترین روش برای توصیف این نوع داده‌ها گزارش میانگین اختلاف اندازه‌های هر نمونه است. گزارش دفعاتی که این اختلاف از مقدار حساس بالینی بزرگ‌تر شده است نیز می‌تواند مفید باشد. برای مثال هنگامی که اختلاف $0/5^{\circ}$ سانتی‌گراد در دمای اندازه‌گیری شده به‌وسیله دو دماسنج از لحاظ بالینی اهمیت دارد، گزارش میانگین اختلاف‌های مشاهده‌شده و دفعاتی که این اختلاف از $0/5^{\circ}$ بیشتر شده است مهم است. مطالعات دیگر بر این مبنای طراحی می‌شوند که تغییرات بین گروه‌های زیادی از مشاهده‌گرها، روش‌ها، آزمایشگاه‌ها یا دستگاه‌ها را ارزیابی کنند. در این موارد نتایج عموماً به‌وسیله محاسبه و اعلام ضریب تغییرات گزارش می‌شوند. (ضریب تغییرات از تقسیم انحراف معیار نمونه به میانگین محاسبه می‌شود). اگر نتایج به‌صورت نرمال توزیع شده باشند، 95% مشاهدات در فاصله‌ای بین دو انحراف معیار از میانگین $(\text{mean} \pm 2\text{SD}_{\text{mean}})$ قرار می‌گیرند.

معیار استاندارد دیگر محاسبه و گزارش ضریب همبستگی داخل رده‌ای (ICC) است. این ضریب به‌صورت زیر محاسبه می‌شود:

$$ICC = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_e^2}$$

که در این رابطه σ_u^2 واریانس اندازه‌گیری‌های واقعی داخل هر گروه و σ_e^2 واریانس خطاهای اندازه‌گیری است. هر قدر که خطای اندازه‌گیری کوچک‌تر باشد افزایش تغییرات خطای اندازه‌گیری کمتر شده و مقدار ICC به ۱ نزدیک‌تر می‌شود. هنگامی که تمام اندازه‌گیری‌ها صحیح انجام گرفته شده باشد $\sigma_e^2 = 0$ و در نتیجه $ICC = 1$ می‌شود. در این حالت تمام تغییرات مشاهده‌شده ناشی از واریانس بین اندازه‌گیری‌ها است. دامنه این معیار بین صفر تا یک است که $ICC = 1$ نشان‌دهنده توافق و تکرارپذیری کامل است.

برای مثال در قسمتی از یک مطالعه مورد-شاهدی برای بررسی وجود ارتباط میان آسم و مصرف آنتی‌اکسیدان در رژیم غذایی (که اطلاعات به‌وسیله تکمیل پرسشنامه توسط بیماران جمع‌آوری شد) دو اندازه‌گیری برای تعیین مقدار مصرف سلنیوم به فاصله سه ماه از یکدیگر انجام گرفت. نمودار ۴-۱ نتایج را برای ۹۴ بیمار ۱۵ تا ۵۰ ساله نمایش می‌دهد. همان‌طور که مشاهده می‌شود با اینکه همبستگی واضحی میان اندازه‌گیری اول و دوم وجود دارد، بین جواب‌ها در پرسشنامه اول و دوم اختلاف اساسی وجود دارد. میانگین و انحراف معیار لگاریتم مصرف سلنیوم در اندازه‌گیری اول برابر $3/826 \pm 0/401$ و در اندازه‌گیری دوم $3/772 \pm 0/768$ مشاهده شد. به نظر می‌رسد کاهش معنی‌داری در میانگین مشاهده‌شده برای اندازه‌گیری دوم نسبت به اندازه‌گیری اول وجود داشته باشد. (میانگین اختلاف $0/058$ با مقدار معنی‌داری $P = 0/083$) [۲].

برآوردهای مؤلفه‌های واریانس جهت محاسبه ICC برابر:

$$\sigma_e^2 = 0/0535, \text{ واریانس خطاهای اندازه‌گیری}$$

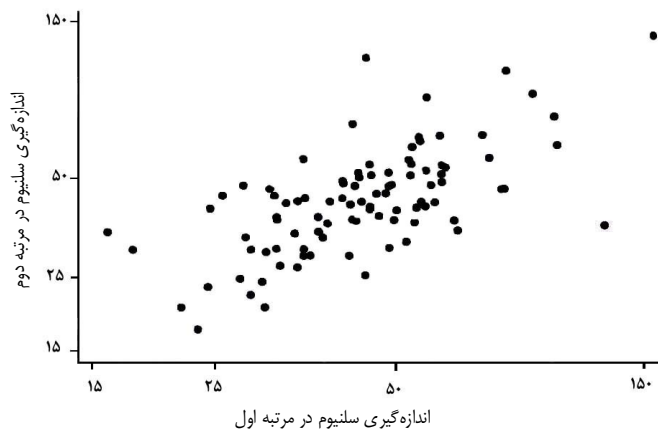
$$\sigma_u^2 = 0.0955, \text{ واریانس داخل هر گروه}$$

$$\sigma_u^2 + \sigma_e^2 = 0.1491, \text{ واریانس کل}$$

می‌شود و در نتیجه:

$$ICC = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_e^2} = \frac{0.0955}{0.1491} = 0.6410$$

بنابراین در این مثال ۶۴/۱ درصد از واریانس کل مربوط به تغییرات داخل گروه‌ها است و بنابراین تنها یک‌بار استفاده از پرسشنامه سنجش رژیم غذایی، مناسب به نظر می‌رسد



نمودار ۴-۱: نمودار پراکنش مقدار مصرف سلیوم ($\mu\text{g} / \text{week}$) در ۹۴ بیمار

۴-۵ مطالعاتی که صحت آزمون‌ها را بررسی می‌کنند

مطالعاتی که در این قسمت بررسی می‌شوند به این موضوع که «تا چه اندازه آزمون پیشنهادی منجر به نتایج درست می‌شود؟» می‌پردازد. برای دستیابی به جواب این سؤال، نیاز است روشی که به‌وسیله آن نتیجه صحیح آزمون‌ها مشخص می‌شود، وجود داشته باشد. برای برخی از بیماری‌ها چنین روش استاندارد وجود دارد. مثلاً نتیجه بیوپسی از بافت معمولاً وجود یا عدم وجود بیماری را تعیین می‌کند. (هرچند آزمون‌های هیستولوژیک هم ایده‌آل نیستند زیرا خطاهای نمونه‌گیری، تهیه نامناسب بافت و یا تشخیص اشتباه ممکن است نتایج این آزمون را مخدوش کند). برای برخی دیگر از بیماری‌ها نیز روش استاندارد قراردادی وجود دارد. مثلاً انسداد پنجاه درصدی حداقل یکی از عروق اصلی کرونر که به‌وسیله آنژیوگرافی مشخص شده باشد، به‌عنوان معیاری برای تشخیص بیماری انسداد عروق کرونر در نظر گرفته می‌شود؛ اما در بسیاری از بیماری‌ها نظیر امراض رماتیسمی، برای تشخیص بیماری لازم است تا فرد معاینه شونده دارای حداقل‌هایی از نشانه‌های بالینی، سندرم‌ها و یا موارد غیرعادی پاراکلینیکی باشد تا به‌عنوان بیمار تلقی شود. در این شرایط طبیعی است که صحت نشانه‌ها، سندرم‌ها یا آزمایش‌های تشخیصی، که خود جزئی از معیارهای تشخیص بیماری هستند، قابل ارزیابی نیست.

۴-۵-۱ طراحی

مطالعاتی که صحت آزمون‌های تشخیصی را بررسی می‌کنند، می‌توانند از نوع مقطعی یا مورد-شاهدی باشند. درحالی‌که مطالعاتی که صحت آزمون‌های پیش‌بینی‌کننده وضعیت بیماری را ارزیابی می‌کنند، معمولاً از نوع هم‌گروهی هستند. در مطالعات مورد-شاهدی از جامعه بیماران و افراد سالم به‌طور مجزا نمونه‌گیری شده و صحت نتایج آزمون در دو گروه مقایسه می‌شوند. متأسفانه در نظر گرفتن طیف وسیعی از افراد تحت مطالعه در نمونه (که از لحاظ بالینی افراد سالم تا بیماران را شامل شود) بسیار دشوار است. در نمونه‌گیری برای چنین مطالعاتی بیماران و افراد سالم نباید از قسمت‌های انتهایی طیف یعنی افراد کاملاً سالم و بیماران با نشانه‌های واضح بیماری که به‌آسانی قابل تشخیص هستند انتخاب شوند. افراد سالم هم باید شامل گروهی باشند که دارای علائم و سندرم‌های مشابه بیماران باشند. به همین علت مطالعات مورد-شاهدی در ارزیابی صحت آزمون‌های تشخیص باید محدود به بیماری‌های نادری شود که هیچ طرح نمونه‌گیری دیگری برای انجام آن‌ها وجود ندارد و یا انجام آن‌ها با دشواری‌هایی همراه است.

طرح مطالعات مقطعی در ارزیابی صحت، به‌طورمعمول نتایج معتبرتری به دست می‌دهند که تفسیر آسان‌تری دارند.

ارزیابی صحت آزمون‌هایی که روند بیماری را پیش‌بینی می‌کنند، نیاز به طراحی مطالعات هم‌گروهی (آینده‌نگر یا گذشته‌نگر) دارند. در مطالعات هم‌گروهی آینده‌نگر آزمون موردنظر، در زمان ورود اعضا به مطالعه انجام می‌شود. این اعضا پیگیری می‌شوند تا سیر تکامل بیماری در آینده مشاهده شود. در شرایط خاص استفاده از مطالعات هم‌گروهی گذشته‌نگر نیز امکان‌پذیر است. برای مثال صحت آزمون «Viral load» در بیماران مبتلا به ایدز را به‌وسیله یک هم‌گروه از این بیماران که قبلاً پیگیری شده‌اند و نمونه خون آن‌ها موجود است، قابل ارزیابی است. هنگامی که بیماری موردنظر نادر است و یا آزمون بسیار پرهزینه است، طراحی مطالعه مورد-شاهدی آشیانه‌ای نسبتاً کارا است. معمولاً انتظار داریم نتایج آزمون‌هایی با ماهیت اسمی، به‌صورت پاسخ‌های مثبت یا منفی باشد؛ اما پاسخ‌های رتبه‌ای یا پیوسته نیز محتمل است. هرگاه که امکان‌پذیر باشد پاسخ‌های رتبه‌ای یا پیوسته به علت اینکه اطلاعات بیشتری در بردارند ارجحیت دارند. در حالتی که متغیر پاسخ دوحالتی (مثبت/منفی) باشد ارزیابی آن به‌وسیله آزمون استاندارد به‌راحتی امکان‌پذیر است.

۴-۵-۲ تحلیل

در تعیین صحت آزمایش یا روش تشخیصی، دو نکته را باید در نظر گرفت:

اول آن که میزان کارایی روش در تشخیص وضعیت موردنظر چقدر است؟ (چند درصد از بیمارانی که دارای نتیجه آزمایش مثبت هستند، واقعاً مبتلا به بیماری هستند؟). این نکته را حساسیت آزمایش می‌گویند. اگر میزان حساسیت آزمایش بالا باشد، میزان منفی کاذب آن پایین است. به‌عبارت‌دیگر نتیجه آزمایش در تعداد زیادی از بیماران مبتلا به این بیماری به‌طور کاذب منفی نمی‌شود. حساسیت را به شکل‌های مختلف می‌توان تعریف کرد: احتمال مثبت بودن نتیجه آزمایش بیمارانی که، مبتلا به بیماری موردنظر هستند؛ نسبتی از بیماران مبتلا به بیماری با نتیجه آزمایش مثبت؛ یا میزان مثبت واقعی.

نکته بعدی این است که صحت یک آزمایش، میزان شناسایی صحیح آن دسته از افرادی است که به بیماری موردنظر مبتلا نیستند. این قسمت از صحت آزمایش را ویژگی (اختصاصی بودن) آزمایش می‌نامند. اگر ویژگی یک آزمایش بالا باشد، میزان مثبت کاذب آزمایش پایین است. ویژگی یک آزمایش را به روش‌های زیر می‌توان تعریف کرد. احتمال نتیجه منفی آزمایش در افرادی که به بیماری موردنظر مبتلا نیستند و پاسخ آزمایش در آن‌ها منفی است؛ یا میزان منفی واقعی.

برای تعیین حساسیت و ویژگی به ترتیب زیر عمل می‌شود:

یک گروه از بیماران که مطمئناً به بیماری (یا وضعیت) موردنظر مبتلا هستند، انتخاب می‌شوند. گروه دیگر که مطمئن هستیم به بیماری (یا وضعیت) موردنظر مبتلا نیستند نیز انتخاب می‌شوند. آزمایش موردنظر را در دو گروه انجام می‌دهیم. سپس نسبت (یا درصد) بیماران مبتلا به بیماری موردنظر را که دارای نتیجه آزمایش مثبت هستند به دست می‌آوریم. علاوه بر حساسیت و ویژگی، دانستن احتمال این موضوع که اعضای موردبررسی واقعاً بیمار باشند، در صورتی که نتیجه آزمون مثبت مشاهده شود و یا واقعاً سالم باشند، در صورتی که نتیجه آزمون منفی مشاهده شود بسیار مهم است. این کمیت‌ها را می‌توان به وسیله شاخص‌های ارزش اخباری آزمون مثبت و ارزش اخباری آزمون منفی نمایش داد. روش تکمیل جدول توافقی برای این آزمون‌ها در جدول (۴-۶) نمایش داده شده است.

جدول ۴-۶: نحوه تنظیم جدول ۲×۲ در آزمون‌های تشخیصی

بیماری		نتیجه آزمایش
وجود ندارد D ⁻	وجود دارد D ⁺	
مثبت کاذب: FP	مثبت واقعی: TP	مثبت: T ⁺
منفی واقعی: TN	منفی کاذب: FN	منفی: T ⁻

۴-۵-۳ ارزیابی آزمون تشخیصی

برای ارزیابی آزمون تشخیصی شاخص‌های حساسیت، ویژگی، ارزش اخباری مثبت و منفی آزمون و درصد مثبت کاذب و منفی کاذب حائز اهمیت هستند. جدول ۴-۷ نحوه محاسبه این شاخص‌ها را نشان می‌دهد.

جدول ۴-۷: شاخص‌های ارزیابی آزمون‌های تشخیصی

شاخص	نحوه محاسبه
- حساسیت	$\frac{TP}{TP + FP} \times 100$
- ویژگی	$\frac{TN}{TN + FN} \times 100$
- ارزش اخباری مثبت آزمون	$\frac{TP}{TP + FN} \times 100$
- ارزش اخباری منفی آزمون	$\frac{TN}{FP + TN} \times 100$
- درصد مثبت کاذب	$\frac{FP}{TP + FP} \times 100$
- درصد منفی کاذب	$\frac{FN}{FN + TN} \times 100$

جدول ۴-۸ نحوه محاسبه این شاخص‌ها را در یک مطالعه درباره توانایی والدین در به خاطر آوردن تزریق واکسن. ب ت ژ و مقایسه با پرونده‌های بهداشتی نشان داده است [۳].

جدول ۴-۸: مقایسه حافظه والدین با پرونده‌های بهداشتی در تزریق واکسن ب ت ژ

اطلاعات تزریق واکسن ب ت ژ به دست آمده از پرونده بهداشتی (مرجع استاندارد آزمون)			
مجموع	خیر	بله	مجموع
بله	۵	۵۵	۶۰
خیر	۲۵	۱۵	۴۰
مجموع	۳۰	۷۰	۱۰۰

حساسیت
 $\frac{55}{70} \times 100 = 78.6\%$

ویژگی
 $\frac{25}{30} \times 100 = 83.3\%$

ارزش اخباری مثبت
 $\frac{55}{60} \times 100 = 91.7\%$

ارزش اخباری منفی آزمون
 $\frac{25}{40} \times 100 = 62.5\%$

حساسیت و ویژگی یک آزمون نشان‌دهنده درستی آن آزمون هستند اما در مورد تصمیم‌گیری بر اساس نتیجه مثبت و منفی یک آزمون، حساسیت و ویژگی چندان مفید نیستند. ارزش اخباری مثبت احتمال وجود بیماری به شرط مثبت بودن نتیجه آزمون و ارزش اخباری منفی احتمال عدم وجود بیماری به شرط منفی بودن نتیجه آزمون را نشان می‌دهند. این دو احتمال اعتبار و سودمندی هر آزمون تشخیصی را در عمل می‌سنجند. نکته دیگر در مورد این شاخص‌ها توجه به شیوع بیماری موردنظر در جامعه و رابطه آن با این شاخص‌ها است. حساسیت و ویژگی شاخص‌هایی هستند که شیوع بیماری تأثیری بر مقدار آن‌ها ندارد؛ اما ارزش اخباری آزمون‌ها به شدت شیوع بیماری در جامعه موردبررسی وابسته است. برای درک این روابط به مثال فرضی زیر توجه نمایید. در این مثال جمعیت فرضی ۱۰۰۰۰۰ نفر در آزمونی با حساسیت ۹۰ درصد و ویژگی ۸۰ درصد با شیوع‌های مختلف بیماری تحت بررسی، در نظر گرفته شده است.

الف: شیوع برابر ۰/۰۰۱

مجموع	واقعیت		نتیجه آزمون
	سالم	بیمار	
۲۰،۰۷۰	۱۹،۹۸۰	۹۰	مثبت
۷۹،۹۳۰	۷۹،۹۲۰	۱۰	منفی
۱۰۰،۰۰۰	۹۹،۹۰۰	۱۰۰	مجموع

$$\text{ارزش اخباری مثبت} = \frac{90}{90 + 19980} = 0/0045$$

$$\text{ارزش اخباری منفی} = \frac{79920}{79920 + 10} = 0/9998$$

ب: شیوع برابر ۰/۰۱

مجموع	واقعیت		نتیجه آزمون
	سالم	بیمار	
۲۰،۷۰۰	۱۹،۸۰۰	۹۰۰	مثبت
۷۹،۳۰۰	۷۹،۲۰۰	۱۰۰	منفی
۱۰۰،۰۰۰	۹۹،۰۰۰	۱،۰۰۰	مجموع

$$\text{ارزش اخباری مثبت} = \frac{900}{900 + 19800} = 0/0434$$

$$\text{ارزش اخباری منفی} = \frac{79200}{79200 + 100} = 0/9987$$

ج: شیوع برابر ۰/۱

مجموع	واقعیت		نتیجه آزمون
	مثبت	منفی	
۲۷۰,۰۰۰	۱۸۰,۰۰۰	۹۰,۰۰۰	مثبت
۷۲۰,۰۰۰	۷۲۰,۰۰۰	۱۰,۰۰۰	منفی
۱,۰۰۰,۰۰۰	۹۰۰,۰۰۰	۱۰۰,۰۰۰	مجموع

$$\text{ارزش اخباری مثبت} = \frac{9000}{9000 + 18000} = 0/3333$$

$$\text{ارزش اخباری منفی} = \frac{72000}{72000 + 1000} = 0/9863$$

د: شیوع برابر ۰/۵

مجموع	واقعیت		نتیجه آزمون
	مثبت	منفی	
۵۵۰,۰۰۰	۱۰۰,۰۰۰	۴۵۰,۰۰۰	مثبت
۴۵۰,۰۰۰	۴۰۰,۰۰۰	۵۰,۰۰۰	منفی
۱,۰۰۰,۰۰۰	۵۰۰,۰۰۰	۵۰۰,۰۰۰	مجموع

$$\text{ارزش اخباری مثبت} = \frac{45000}{45000 + 10000} = 0/8181$$

$$\text{ارزش اخباری منفی} = \frac{40000}{40000 + 5000} = 0/8888$$

ه: شیوع برابر ۰/۹

مجموع	واقعیت		نتیجه آزمون
	سالم	بیمار	
۸۳۰,۰۰۰	۲۰,۰۰۰	۸۱۰,۰۰۰	مثبت
۱۷۰,۰۰۰	۸۰,۰۰۰	۹۰,۰۰۰	منفی
۱,۰۰۰,۰۰۰	۱۰۰,۰۰۰	۹۰۰,۰۰۰	مجموع

$$\text{ارزش اخباری مثبت} = \frac{81000}{81000 + 2000} = 0/9759$$

$$\text{ارزش اخباری منفی} = \frac{8000}{9000 + 8000} = 0/4705$$

و: شیوع برابر ۰/۹۹

مجموع	واقعیت		نتیجه آزمون
۸۹،۳۰۰	۲۰۰	۸۹،۱۰۰	مثبت
۱۰،۷۰۰	۸۰۰	۹،۹۰۰	منفی
۱۰۰،۰۰۰	۱،۰۰۰	۹۹،۰۰۰	مجموع

$$\text{ارزش اخباری مثبت} = \frac{۸۹۱۰۰}{۸۹۱۰۰ + ۲۰۰} = ۰/۹۹۷۸$$

$$\text{ارزش اخباری منفی} = \frac{۸۰۰}{۹۹۰۰ + ۸۰۰} = ۰/۰۷۴۸$$

ز: شیوع برابر ۰/۹۹۹

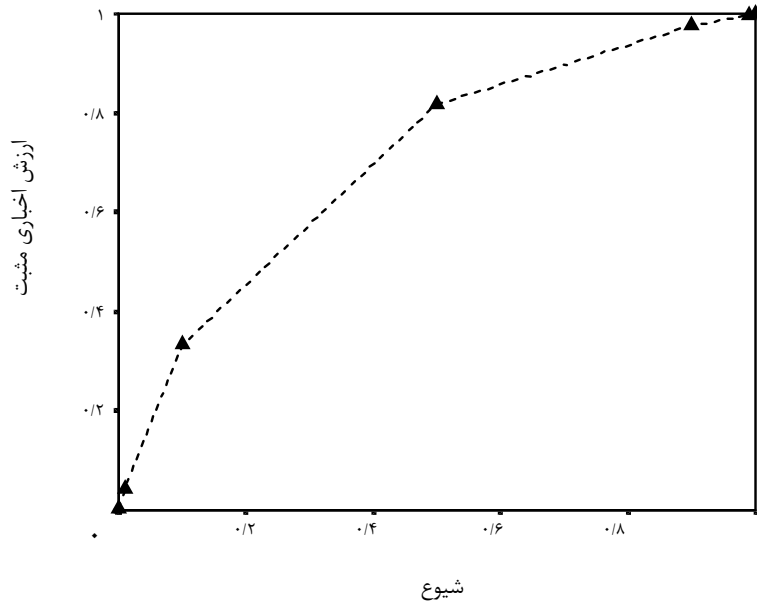
مجموع	واقعیت		نتیجه آزمون
۸۹،۹۳۰	۲۰	۸۹،۹۱۰	مثبت
۱۰،۰۷۰	۸۰	۹،۹۹۰	منفی
۱۰۰،۰۰۰	۱۰۰	۹۹،۹۰۰	مجموع

$$\text{ارزش اخباری مثبت} = \frac{۸۹۹۱۰}{۸۹۹۱۰ + ۲۰} = ۰/۹۹۹۸$$

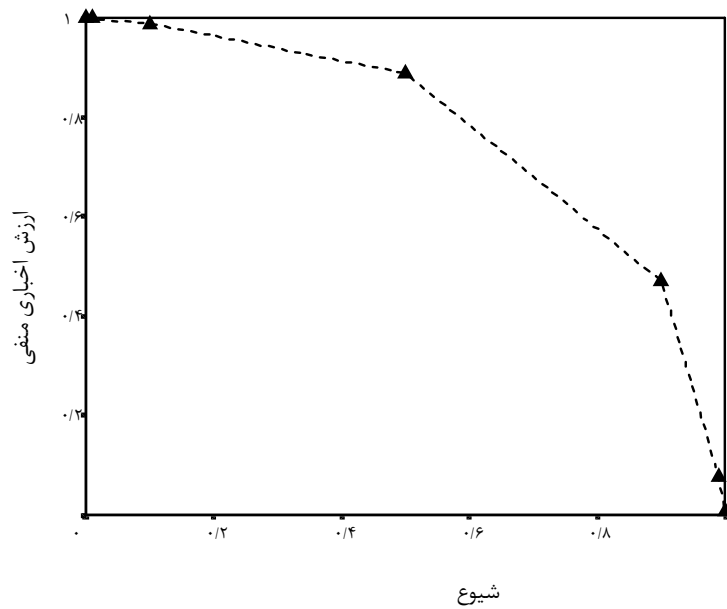
$$\text{ارزش اخباری منفی} = \frac{۸۰}{۹۹۹۰ + ۸۰} = ۰/۰۰۷۹$$

نمودارهای ۴-۲-الف و ب رابطه ارزش اخباری مثبت و منفی را با شیوع بیماری در این مثال فرضی نشان

می‌دهد.



نمودار ۴-۲-الف: رابطه شیوع و ارزش اخباری مثبت برای حساسیت ۹۰٪ و ویژگی ۸۰٪ در یک جامعه فرضی ۱۰۰،۰۰۰ نفری



نمودار ۴-۲-ب: رابطه شیوع و ارزش اخباری منفی برای حساسیت ۹۰٪ و ویژگی ۸۰٪ در یک جامعه فرضی ۱۰۰،۰۰۰ نفری

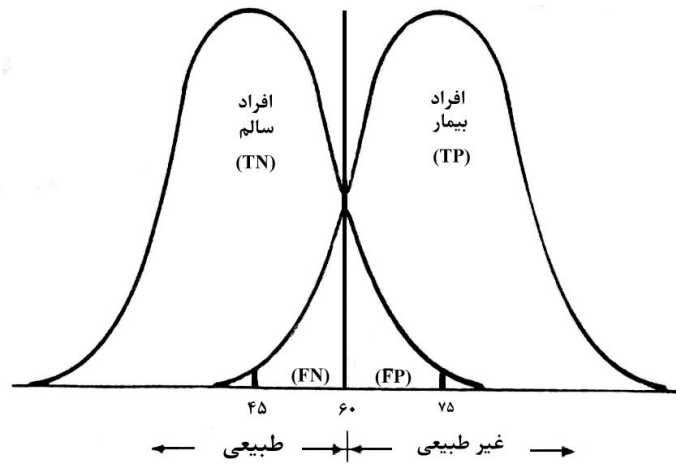
۴-۵-۴ منحنی‌های راک

روش‌های قبلی برای ارزیابی آزمون‌های تشخیصی هنگامی کاربرد دارند که نتیجه روش تشخیصی مثبت یا منفی باشد. با این وجود، در بسیاری از آزمایش‌ها کمیت‌هایی با مقیاس پیوسته اندازه‌گیری می‌شوند. در مواقعی که مقادیر نتیجه آزمایش به‌طور پیوسته تعیین می‌شود، میزان‌های حساسیت و ویژگی به محلی که نقطه تمایز بین مثبت و منفی شدن تشخیص تنظیم شده است بستگی دارد. با استفاده از دو منحنی توزیع نرمال مربوط به مقادیر نتیجه آزمایش (یک منحنی توزیع برای افراد مبتلا به بیماری و منحنی توزیع دیگر برای افراد غیر مبتلا به بیماری موردنظر) می‌توان این موقعیت را نمایش داد. در نمودارهای ۴-۳-الف و ب دو منحنی توزیع فرضی در حالتی که مقدار میانگین آزمودنی (مثلاً میزان غلظت کلسترول در خون) برای افراد مبتلا به بیماری ۷۵ و برای افراد بدون ابتلا به بیماری ۴۵ هستند، نشان داده شده است.

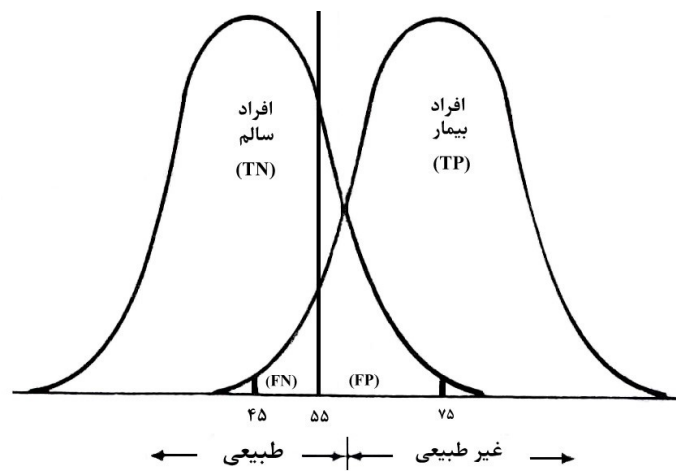
اگر نقطه تمایز تصمیم‌گیری را ۶۰ در نظر بگیریم، همان‌طور که در نمودار ۴-۳-الف ملاحظه می‌شود، حدود ۱۰ درصد افراد بدون ابتلا به بیماری، از نظر آزمایش «غیرطبیعی» قلمداد می‌شوند. زیرا مقدار نتیجه آزمایش آن‌ها از ۶۰ بالاتر است. به همین ترتیب، حدود ۱۰ درصد از افراد مبتلا به بیماری از نظر آزمایش «طبیعی» قلمداد می‌شوند. زیرا مقدار نتیجه آزمایش آن‌ها کمتر از ۶۰ است. به عبارت دیگر در این آزمایش، حساسیت ۹۰٪ و ویژگی ۹۰٪ است.

فرض کنیم پزشکی مایل به استفاده از آزمایشی با حساسیت بالاتر است. به عبارت دیگر، پزشک داشتن تعداد بیشتر مثبت کاذب را به افرادی که واقعاً به بیماری مبتلا هستند، ترجیح می‌دهد. در نمودار ۴-۳-ب این حالت، یعنی افزایش حساسیت، با کاهش نقطه تمایز تصمیم‌گیری بین مقدار طبیعی و غیرطبیعی از مقدار ۶۰ به مقدار ۵۵، نمایش داده شده است. میزان حساسیت آزمایش افزایش می‌یابد ولی درازای آن ویژگی آزمایش کاهش یافته است.

برای نمایش ارتباط بین ویژگی و حساسیت آزمایش‌هایی که نتیجه آن‌ها از نوع پیوسته هستند، روش مؤثرتری وجود دارد. در عرصه ارتباطات برای نمایش نسبت‌های سیگنال به سروصدای مزاحم، منحنی‌های راک ایجاد شده‌اند. برای درک کاربرد این منحنی، در یک روش تشخیصی، نتیجه مثبت واقعی را دریافت «سیگنال صحیح» و نتیجه مثبت کاذب را دریافت «پارازیت» در نظر می‌گیریم. منحنی راک نموداری است که از نسبت حساسیت (میزان مثبت واقعی) به میزان مثبت کاذب به دست می‌آید. خط نقطه‌چین در وسط نمودار ۴-۴، مربوط به آزمایشی است که به‌طور شانسی مثبت یا منفی است. هر چه منحنی راک به گوشه چپ و بالای نمودار نزدیک‌تر باشد، صحت آن بیشتر است. زیرا در آنجا میزان مثبت واقعی یک و میزان مثبت کاذب صفر می‌شود. ملاک‌های آزمایش مثبت هر چه دقیق‌تر تعریف شوند، نقطه مربوط به حساسیت و ویژگی (نقطه الف) در روی منحنی به طرف پایین و چپ (یعنی حساسیت کمتر و ویژگی بیشتر) جابه‌جا می‌شوند [۳].

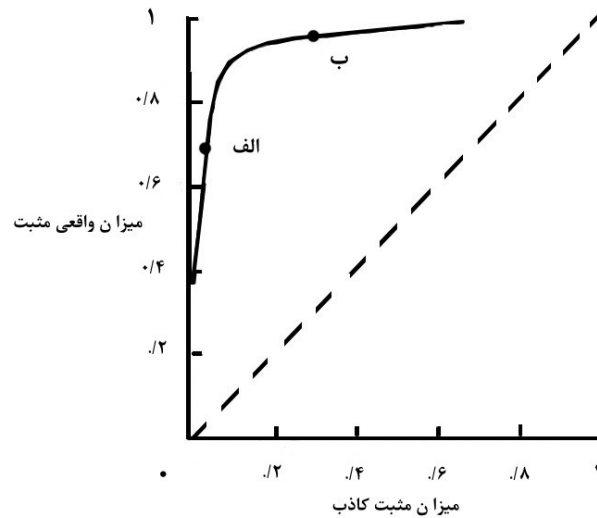


نمودار ۴-۳-الف: نمودار دو توزیع فرضی به ازای خط تصمیم‌گیری برابر با نقطه تمایز ۶۰



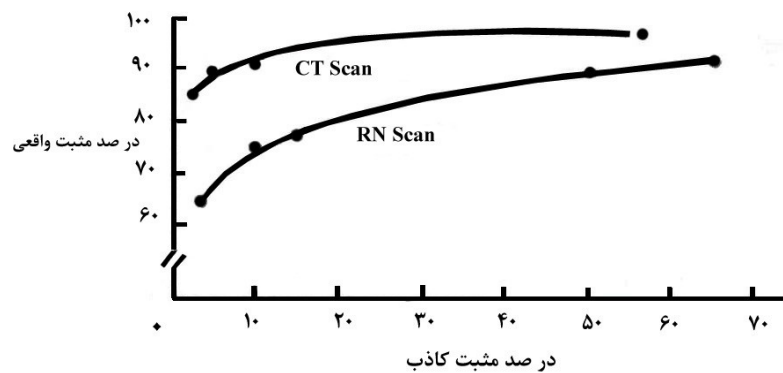
نمودار ۴-۳-ب: نمودار دو توزیع فرضی به ازای خط تصمیم‌گیری برابر با نقطه تمایز ۵۵

اگر ملاک‌های کمتری برای مثبت بودن نتیجه آزمایش موردنیاز باشد، نقطه مربوط به حساسیت و ویژگی (نقطه ب) در روی منحنی به طرف بالا و راست (یعنی حساسیت بیشتر و ویژگی کمتر) جابه‌جا می‌شود.



نمودار ۴-۴: منحنی راک

منحنی‌های راک روش‌های نموداری مفیدی برای مقایسه دو روش تشخیصی هستند. به‌عنوان مثال، از هر دو روش اسکن رادیونوکلئید (RN) و توموگرافی کامپیوتری (CT) برای کشف تومورهای مغزی استفاده می‌شود. منحنی‌های راک برای مقایسه این دو روش به‌کاربرده شده‌اند. نمودار ۴-۵ منحنی‌های راک این دو روش را نمایش می‌دهد [۴].



نمودار ۴-۵: نمایش دو منحنی راک برای ارزیابی دو روش تشخیصی

با انجام یک آزمون آماری می‌توان تعیین نمود که آیا دو منحنی راک با هم تفاوت قابل توجهی دارد یا خیر. در این روش، مساحت زیر هر منحنی راک تعیین شده و با استفاده از روش اصلاح شده حاصل جمع رتبه‌ای ویلکاکسون دو روش مقایسه می‌شود.

۴-۵-۵ نسبت‌های درستنمایی

هرچند که اطلاعات یک آزمون تشخیصی به وسیله حساسیت و ویژگی در متغیرهای دوحالتی و به وسیله منحنی‌های راک در متغیرهای پیوسته قابل بیان هستند اما روش بهتری نیز وجود دارد. نسبت درستنمایی برای آزمون مثبت و منفی به محقق اجازه می‌دهد تمام اطلاعات موجود در یک آزمون را به صورت خلاصه بیان کند. برای هر یک از آزمون‌های مثبت یا منفی نسبت درستنمایی به صورت زیر بیان می‌شود.

$$\text{نسبت درستنمایی} = \frac{\text{بیماری | نتیجه آزمون (احتمال)}}{\text{سالم بودن | نتیجه آزمون (احتمال)}}$$

در مورد متغیرهای اسمی نسبت درستنمایی آزمون مثبت برابر

$$\text{نسبت درستنمایی مثبت آزمون} = \frac{\text{حساسیت}}{\text{(ویژگی - ۱)}}$$

و نسبت درستنمایی آزمون منفی برابر

$$\text{نسبت درستنمایی منفی آزمون} = \frac{\text{(حساسیت - ۱)}}{\text{ویژگی}}$$

است.

علت استفاده از نسبت درستنمایی این است که وقتی آزمونی کاملاً معتبر نباشد مطلع شدن از نتیجه این آزمون مشخص نخواهد کرد که فرد بیمار است یا نه. با استفاده از نسبت‌های درستنمایی احتمال این وقایع قابل پیش‌بینی هستند. هر قدر میزان نسبت درستنمایی بزرگ‌تر باشد توانایی آزمون در تشخیص صحیح بیماری بیشتر می‌شود. نسبت‌های درستنمایی بالاتر از ۱۰۰ ایده‌آل هستند. هر قدر میزان نسبت درستنمایی کوچک‌تر باشد (به صفر نزدیک‌تر باشد) توانایی آزمون در تشخیص صحیح افراد سالم بیشتر می‌شود. نسبت درستنمایی وقتی برابر یک باشد نشان‌دهنده این است که هیچ اطلاع جدیدی از آزمون حاصل نمی‌شود.

برای مثال جدول ۴-۹ نتایج مطالعه را برای بررسی امکان پیش‌بینی مننژیت باکتریایی^{۱۲} و یا پیدایش باکتری در خون^{۱۳} به وسیله تعداد گلبول‌های سفید خون در نوزادان مبتلا به تب نمایش می‌دهد [۵].

۱. Bacterial Meningitis

۲. Bacteremia

جدول ۴-۹: رابطه بین مننژیت باکتری یا وجود باکتری در خون و تعداد گلبول‌های سفید نوزادان.

نسبت درستمایی	وجود مننژیت یا باکتری در خون		تعداد گلبول‌های سفید
	ندارد	دارد	
۲/۰	۹۶ (%۴)	۵ (%۸)	۵۰۰۰ >
۰/۷	۸۵۶ (%۳۹)	۱۸ (%۲۹)	۵۰۰-۹۹۹۹
۰/۳	۷۹۰ (%۳۶)	۸ (%۱۲)	۱۰۰۰-۱۴۹۹۹
۲/۱	۲۸۶ (%۱۳)	۱۷ (%۲۷)	۱۵۰۰-۱۹۹۹۹
۳/۴	۱۵۱ (%۷)	۱۵ (%۲۴)	۲۰۰۰۰ ≤
	۲۱۷۷ (%۱۰۰)	۶۳ (%۱۰۰)	مجموع (درصد)

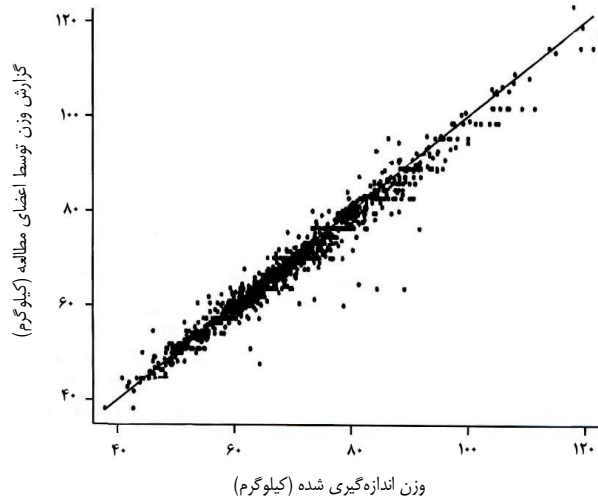
همان‌طور که ملاحظه می‌شود تعداد گلبول‌های سفید کمتر از ۵,۰۰۰ یا بیشتر از ۲۰,۰۰۰ در بین نوزادان مبتلا به مننژیت یا دارای باکتری در خون شایع‌تر است. نسبت درستمایی این موضوع را بسیار ساده نمایش می‌دهد. برای مثال، ۸٪ از نوزادانی که تعداد گلبول‌های سفید در خون آن‌ها کمتر از ۵,۰۰۰ است مبتلا به بیماری هستند در صورتی که تنها ۴٪ به بیماری مبتلا نیستند. بنابراین نسبت درستمایی برابر $\frac{۸\%}{۴\%} = ۲$ است.

۴-۵-۶ ارزیابی روش‌های تشخیصی در داده‌های پیوسته

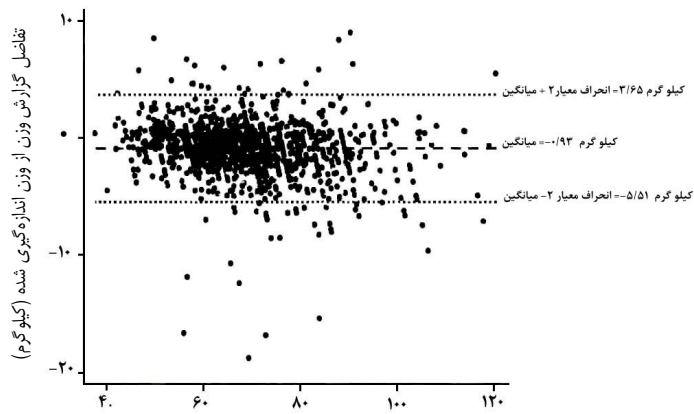
در این بخش، روش تحلیل مناسب برای مطالعاتی که روش‌های مختلف تشخیصی را برای اندازه‌گیری یک پارامتر یکسان مقایسه می‌کنند، بررسی می‌شوند. مشهورترین روش برای تحلیل این مطالعات استفاده از روش بلاند و آلمن است [۶] و [۷].

داده‌های نمایش داده شده در نمودار ۴-۶ به مطالعه بررسی سلامت قلب زنان بریتانیا اختصاص دارد. از ۱۲,۲۳۶ زن شرکت‌کننده در این مطالعه درخواست شد وزن خود را گزارش کنند و سپس وزن آن‌ها در مطالعه اندازه‌گیری شد. این نمودار پراکنش این دو اندازه را توصیف می‌کند. این دو اندازه همبستگی بالایی دارند. برای مثال ضریب همبستگی پیرسون بین این دو اندازه برابر ۰/۹۸۲ است. باین‌حال بالا بودن همبستگی نشان‌دهنده وجود توافق مناسب بین دو کمیت نیست. زیرا برای مثال اگر یک متغیر همیشه دو برابر متغیر دیگر باشد ضریب همبستگی آن‌ها دقیقاً برابر یک است. به‌جز این موضوع، ضریب همبستگی، کمیتی وابسته به محدوده داده‌ها است. هر قدر محدوده داده‌ها وسیع‌تر باشد همبستگی بالاتر و هر قدر داده‌ها نزدیک‌تر به هم باشند همبستگی

پایین تر می شود.



نمودار ۴-۶: نمودار پراکنش گزارش وزن توسط افراد تحت مطالعه در برابر اندازه گیری های انجام شده در ۱۲۳۶ زن شرکت کننده در مطالعه قلب زنان در بریتانیا



نمودار ۴-۷: نمودار بلاند و آلتمن از تفاضل گزارش از وزن اندازه گیری شده (محور عمودی) و میانگین وزن گزارش شده و وزن اندازه گیری شده (محور افقی) در ۱۲۳۶ زن شرکت کننده در مطالعه قلب زنان در بریتانیا

خط نظری در نمودار ۴-۶ ناحیه‌ای که دو اندازه دارای مقدارهای کاملاً برابر هستند را مشخص می‌کند. اگر تمام نقاط روی این خط قرار بگیرند، دو اندازه‌گیری دارای توافق کامل هستند. همان‌طور که از شکل مشخص است، درصد عمده‌ای از مشاهدات زیر این خط قرار گرفته‌اند. این موضوع نشان می‌دهد که معمولاً گزارش وزن توسط افراد مطالعه کمتر از مقدار واقعی آن است [۸].

بلاند و آلتمن [۹] پیشنهاد کردند که میزان توافق، به‌وسیله رسم نمودار تفاضل هر زوج از اندازه‌ها (در محور عمودی) در برابر میانگین هر زوج (در محور افقی) بررسی شود. این نمودار که معمولاً تحت عنوان نمودار بلاند و آلتمن شناخته می‌شود در نمودار ۴-۷ نمایش داده شده است. اگر یکی از اندازه‌گیری‌ها صحیح باشد، میانگین تفاضل مشاهده‌شده وجود تورش در اندازه‌گیری جدید را مشخص می‌کند. در این مثال، میانگین وزن گزارش شده ۶۸/۸۸ کیلوگرم و میانگین وزن اندازه‌گیری شده ۶۹/۸۵ کیلوگرم است. میانگین اختلاف بین این دو اندازه‌گیری برابر ۰/۹۳- کیلوگرم (بافاصله اطمینان ۰/۸۰- تا ۱/۰۷- کیلوگرم) است. به‌این ترتیب، گرایش بارزی در گزارش کمتر وزن وجود دارد که میانگین این کاهش برابر ۰/۹۳ کیلوگرم است. نمودار ۴-۷ میانگین به همراه فاصله اطمینان ۹۵٪ توافق بین دو اندازه را نمایش می‌دهد.

۴-۶ تأثیر خطای اندازه‌گیری در تحلیل نتایج

مسائل و مشکلاتی که از وجود خطا در اندازه‌گیری پیامد مورد مطالعه با متغیرهای مربوط به مواجهه رخ می‌دهد، در جدول ۴-۱۰ نشان داده شده است.

جدول ۴-۱۰: مشکلات ناشی از خطای اندازه‌گیری یا طبقه‌بندی

نوع خطا		نوع متغیر
خطای اندازه‌گیری (متغیرهای پیوسته)	سو طبقه‌بندی (متغیرهای اسمی)	
رگرسیون حول میانگین	تضعیف اثرها در مدل رگرسیون	پیامد
مهار مخدوش‌کننده‌ها	تضعیف اثرها در مدل رگرسیون مشکلات بالقوه برای مهار مخدوش‌کننده‌ها	مواجهه

۴-۶-۱ تضعیف اثرها در مدل رگرسیون

بر اثر وجود خطا در طبقه‌بندی پیامدها، اثرهای برآورد شده برای اندازه‌گیری اثر مواجهه، به سمت فرضیه صفر (بی‌اثر بودن مواجهه) گرایش (تورش) پیدا می‌کند. به‌این ترتیب اندازه همبستگی بین مواجهه و پیامد کمتر از اندازه واقعی برآورد می‌شود.

در مواجهه‌هایی که از نوع پیوسته هستند این تورش با ضریب همبستگی داخل رده‌ای (ICC) مرتبط است. در مدل رگرسیون خطی این رابطه به‌صورت زیر است.

$$ICC \times \text{اثر حقیقی} = \text{اثر برآورد شده}$$

این رابطه به‌طور تقریبی هنگامی که اندازه اثر حقیقی و واریانس خطا خیلی بزرگ نباشد در مدل‌های

رگرسیون کاکس و لجستیک نیز برقرار است.

در مواجهه رسته‌ای، اندازه اثر برآورد شده به‌وسیله انجام اندازه‌های مکرر در تمام یا برخی از افراد، قابل تصحیح است. با این حال روش‌های محاسبه اندازه اثر واقعی بسیار پیچیده‌تر از حالت قبلی هستند. علت اصلی همبستگی خطاها با مقادیر حقیقی اندازه‌گیری‌ها است [۱۰] و [۱۱].

۴-۶-۲ رگرسیون حول میانگین

پدیده رگرسیون حول میانگین برای اولین بار توسط گالتون گزارش شد. او هنگام آمیزش گرده‌های لوبیا متوجه شد محصول به‌عمل آمده از دو لوبیای بلند از هر دو لوبیا کوتاه‌تر است و برعکس.

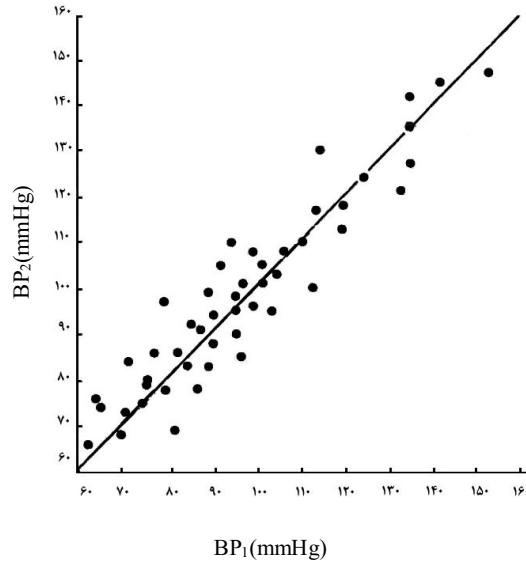
نظیر همین پدیده در اندازه‌گیری مکرر یک کمیت با وجود خطای اندازه‌گیری رخ می‌دهد. مقادیر بزرگ در اندازه‌گیری اول، در اندازه‌گیری‌های بعدی کوچک‌تر می‌شود و مقادیر کوچک در اندازه‌گیری اول در اندازه‌گیری‌های بعدی بزرگ‌تر می‌شود. در واقع بین اندازه اول و سایر اندازه‌ها همبستگی منفی وجود دارد.

نمودار ۴-۸ رابطه بین دو فشارخون دیاستولیک در یک فاصله شش‌ماهه را در ۵۰ داوطلب نمایش می‌دهد [۸] و نمودار ۴-۹ پراکنش تفاضل بین دو مشاهده (محور عمودی) در برابر اندازه اولیه مشاهده‌شده (محور افقی) را نشان می‌دهد. همان‌طور که از شکل‌ها مشخص است اندازه اولیه فشارخون بالا تمایل به کاهش در اندازه‌گیری دوم (شش ماه بعد) دارد و برعکس. اگر بین کاهش واقعی دو اندازه و اندازه حقیقی مرتبه اول رابطه‌ای وجود نداشته باشد، ضریب رگرسیون مربوط به همبستگی مشاهده‌شده (β_{obs}) بین تفاضل دو اندازه و اندازه مرتبه اول به‌صورت زیر محاسبه می‌شود.

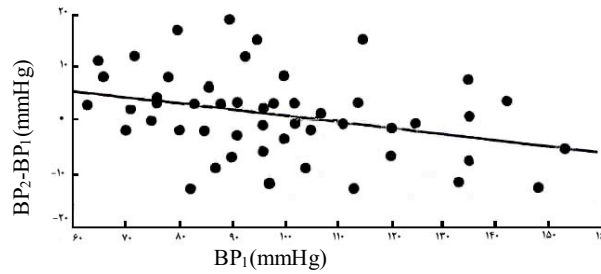
$$\beta_{obs} = ICC - 1$$

هرقدر واریانس خطای اندازه‌گیری بزرگ‌تر باشد، ICC کوچک‌تر می‌شود و ضریب رگرسیون مشاهده‌شده بزرگ‌تر می‌شود.

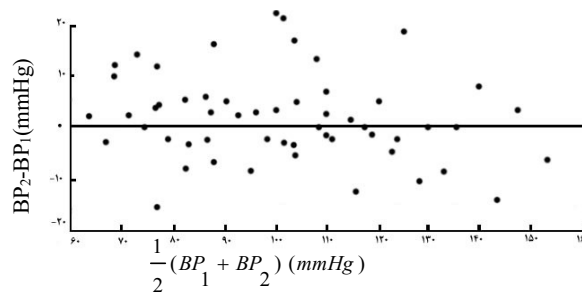
یکی از روش‌های متداول برای کاهش اثر رگرسیون حول میانگین استفاده از میانگین اندازه‌گیری‌ها است. نمودار ۴-۱۰ نمودار پراکنش تفاضل دو اندازه‌گیری فشارخون را در برابر میانگین دو اندازه نمایش می‌دهد. همان‌طور که ملاحظه می‌شود اثر رگرسیون حول میانگین تا حد زیادی تعدیل شده است [۱۲] و [۱۴].



نمودار ۸-۴: رابطه بین دو اندازه فشارخون دیاستولیک که در یک دوره شش‌ماهه اندازه‌گیری شده است.



نمودار ۹-۴: تغییرات فشارخون دیاستولیک در برابر مقادیر اولیه. ضریب همبستگی $r = -0.35$ به همراه خط رگرسیون مستقیم برازش شده به داده‌ها.



نمودار ۱۰-۴: تغییرات فشارخون دیاستولیک در برابر میانگین مقادیر اولیه و ثانویه. ضریب همبستگی $r = -0.19$ به همراه خط رگرسیون مستقیم برازش شده به داده‌ها.

۴-۷ مطالعاتی که اثر یک آزمون را در تشخیص بیماری بررسی می کنند

یک آزمون ممکن است معتبر باشد، اما اگر بیماری نادر باشد، آزمون تنها در موارد بسیار اندکی مثبت خواهد بود. این موضوع معادل این است که انجام این آزمون در موارد بسیار زیادی لازم به تجویز نیست. آزمون‌های دیگری نیز ممکن است اغلب اوقات دارای نتیجه مثبت باشند، اما هیچ تأثیری بر تصمیمات بالینی نداشته باشند. زیرا هیچ اطلاع‌گر جدیدی نسبت به آنچه از سابقه بیمار، معاینات بالینی و یا آزمون‌های دیگر به دست می‌آید به معلومات معالجه اضافه نمی‌کنند [۱۸-۱۴].

در این قسمت، طراحی مطالعاتی مورد بحث قرار می‌گیرد که میزان فایده آزمون‌ها در تشخیص بیماری و اثرات آنها بر تصمیمات بالینی را ارزیابی می‌کنند.

۴-۷-۱ طراحی

مطالعاتی که فایده آزمون مورد نظر را در تشخیص بیماری بررسی می‌کنند به سؤالاتی از این قبیل پاسخ می‌دهند.

- ♦ هنگامی که این آزمون برای یک موقعیت خاص تجویز می‌شود، چند درصد از موارد آنومالی مشخص می‌شود؟
 - ♦ آیا نتایج آزمون با استفاده از سایر اطلاعات موجود در مورد بیماری، قابل پیش‌بینی هستند؟
 - ♦ چه اتفاقی برای بیمارانی که دارای آنومالی هستند می‌افتد؟ آیا آنها بهره‌ای از نتایج آزمون می‌برند؟
- این مطالعات نسبت نتایج مثبت را در میان بیمارانی که دارای علائم خاص هستند بررسی می‌کنند. معمولاً منطقی است که فرض کنیم:
- ♦ احتمال مشاهده نتایج مثبت (وجود عارضه یا پیامد)، برای بیمارانی که این آزمون برای آنها تجویز شده است بیشتر از سایر افراد است.
 - ♦ افرادی که نتیجه آزمون آنها منفی است، فایده‌ای از آزمون نمی‌برند (زیرا فرایند درمان در آنها تغییر نمی‌کند).

با این مفروضات، اگر میزان نتایج مثبت یک آزمون کم باشد، می‌توان استدلال کرد که کاربرد این آزمون در این مورد خاص مفید نیست. برای ارزیابی تأثیر مستقیم نتایج آزمون بر تصمیمات بالینی، مطالعاتی از نوع مقایسه قبل و بعد طراحی می‌شود. در این مطالعات معمولاً نوع درمان‌ها و عملیاتی که تیم پزشکی قبل و بعد از مشخص شدن نتایج آزمون تجویز می‌کنند، مورد مقایسه قرار می‌گیرند [۱۵].

برای مثال در مطالعه‌ای که برای ارزیابی اثر سونوگرافی در تغییر تصمیمات بالینی، در دردهای حاد ناحیه شکم در کودکان طراحی انجام شده بود، تیم تحقیق، تشخیص بالینی و درمان‌های تجویز شده برای ۹۴ کودک دارای درد حاد ناحیه شکم را، قبل از آزمون سونوگرافی و پس از مشخص شدن نتایج سونوگرافی، مقایسه کردند. در این مطالعه، در ۴۶٪ از موارد اطلاع اضافی به دست آمده از سونوگرافی، تشخیص و نحوه درمان بالینی را تحت تأثیر قرار داد [۱۶].

واضح است که تغییر تصمیم‌های بالینی، به تنهایی سودمند بودن آزمون برای بیماران را تضمین نمی‌کند. این موضوع تنها هنگامی تأیید می‌شود که برای تشخیص‌های حاصل از آزمون، درمان‌های مفید و کارا وجود داشته

باشد. در مثال بالا احتمال بسیار زیادی وجود دارد که بیماران از آزمون جدید بهره‌مند گردند. زیرا برای مثال تغییر تصمیمات بالینی از «مرخص شدن از بیمارستان» به «عمل جراحی» در کودکانی که دچار آپاندیس شده‌اند، یا تغییر تصمیم از «عمل جراحی» به «تحت بررسی» بودن در بیمارستان در کودکانی که دچار دردهای مبهم در شکم هستند، نتایجی هستند که با استفاده از سونوگرافی قابل‌دستیابی هستند.

۴-۷-۲ تحلیل

تحلیل آماری و گزارش نتایج این مطالعات معمولاً ساده است. غالباً نسبت آزمون‌های مثبت و آزمون‌هایی که باعث تغییر در مدیریت و یا (به‌طور بالقوه) بهبود پیامدها می‌شوند، محاسبه و به همراه فاصله اطمینان‌های مربوطه گزارش می‌شوند. مطالعاتی که قابلیت اجرا، هزینه‌ها و خطرهای آزمون را بررسی می‌کنند، یکی از بخش‌های مهم در مطالعات بالینی است. این مطالعات قابل‌اجرا و عملی بودن آزمون‌های تشخیصی را بررسی می‌کنند و معمولاً توصیفی هستند. طرح نمونه‌گیری در این مطالعات بسیار مهم است [۱۹]، زیرا نتایج غالباً از مؤسسه‌ای به مؤسسه دیگر و همچنین در بین بیماران متفاوت است.

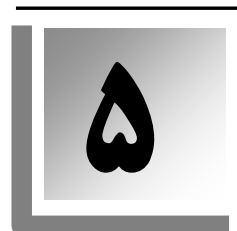
منابع

1. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*, 1986, 1:307-10.
2. Margetts BM, Shaheen SO, Sterne JAC, Thompson RL, Songhurst CE, Burney PGJ. Association between asthma and intake of dietary antioxidants. *Am J Resp Clin Care Med*, 2001, 164:1823-28.
۳. سرافراز ع، عطارزادگان ک. آمار پزشکی. تألیف داوسون - ساندرز ب و تراپ جی آر، مشهد: انتشارات دانشگاه علوم پزشکی مشهد، ۱۳۷۶، صفحات ۶۰۸-۵۶۹.
4. Griner PF, et al. Selection and interpretation of diagnostic tests and procedures. *Annals of Internal Medicine*, 1981, 94:553-600.
5. Puntell RH, Newman TB, Takayam JI, et al. The PROS febrile infant study. In preparation, 1999.
6. Bland JM, Altman DG. Measurement error and correlation coefficient. *BMJ*, 1996, 313:211-2.
7. Bland JM, Altman DG. Measurement error proportional to the mean. *BMJ*, 1990 313:106.
8. Kirkwood BR, Jonathan ACS. *Medical statistics* (2nd ed.). Blackwell Science, Massachusetts 2003.
9. Bland JM, Altman DG. Comparing methods of measurement: Why plotting difference against standard method is misleading? *Lancet*, 1995, 346:1085-87.
10. Frost C, Thompson SG. Correcting for regression dilution bias: Comparison of methods for a single predictor variable. *Journal of the Royal Statistical Society A*, 2000, 163:173-189.
11. White I, Frost C, Tokanug AS. Correcting for measurement error in binary and continuous variables using replicates. *Statistics in Medicine*, 2001 20: 3441-57.
12. Kazemnejad A, Sanagoo M, Mohebbi M. Estimating treatment effects in the presence of regression to the mean. 53rd Annual meeting of ISI, 2002.
13. Kazemnejad A, Mohebbi M, Sanagoo M, Estimating treatment effects in selected paired data. INTERSTAT 2003.
14. Weisberg S. *Applied linear regression*. New York Wiley, 1985.
15. Friedland DJ. *Evidence-based medicine: a framework for clinical Practice*. Stamford, CT. Appleton & Lange, 1998.

۱۶. سوری ح، جوادی م ع، رفعتی ن، محبی م ر. طراحی و اجرای مطالعه تعیین شیوع و علل نابینایی و کم بینایی در استان تهران در سال ۸۴-۱۳۸۳. مجله چشم‌پزشکی بینا، ۱۳۸۴؛ دوره ۱۱، پاییز، شماره ۱: صفحات ۲۲-۳.

17. Carrico CW, Fenton LZ, Taylor GA, et al. Impact of sonography on the diagnosis and treatment of acute lower abdominal pain in children and young adults. *Am J Roentgenol*, 1999, 772:513-6.
18. Soori H. Measuring health-related quality of life among schoolchildren in Ahwaz-Iran. *Primary Prevention*, September 2004, 25 (1): 125-132.
19. Mohammadi MR, Soori H, et al. Suicidal attempt and psychiatric disorders in Iran. *Suicide Life Threat Behav*. 2005 Jun;35(3):309-16.

مشکلات سبب‌شناسی در مطالعات پزشکی



انجام استنباط سببی از روابط مشاهده‌شده در مطالعات، از مسائل بسیار مهم در تحقیقات بالینی است. این فصل به بررسی مشکلات سبب‌شناسی در مطالعات پزشکی اختصاص دارد. ابتدا روش‌های اجتناب از روابط تصنعی را معرفی می‌کنیم و سپس به بررسی روابط حقیقی غیر سببی می‌پردازیم. در ادامه نقش شانس و مخدوش شدن در ایجاد روابط غیر سببی بررسی می‌شود.

۱-۵ روابط تصنعی

فرض کنید می‌خواهیم ارتباط بین مصرف قهوه و سکته قلبی را بررسی کنیم. یکی از حالت‌های ممکن این است که نوشیدن قهوه یکی از عوامل وقوع سکته قلبی است؛ اما قبل از استنباط چنین موضوعی باید چهار گزینه دیگر که باعث ایجاد ارتباط بین دو پدیده می‌گردد را نیز مدنظر داشت. جدول ۱-۵ استنباط‌های ممکن را پس از مشاهده یک رابطه بررسی می‌کند.

جدول ۱-۵: پنج استنباط ممکن که پس از مشاهده رابطه بین نوشیدن قهوه و سکته قلبی در نمونه مورد مطالعه امکان وقوع دارند.

الگوی سببی	آنچه واقعاً در جامعه برقرار است	نوع رابطه	گزینه‌های ممکن
-	نوشیدن قهوه و وقوع سکته قلبی با هم ارتباط ندارند.	تصنعی	۱. شانس (خطای تصادفی)
-	نوشیدن قهوه و وقوع سکته قلبی با هم ارتباط ندارند.	تصنعی	۲. تورش (خطای منظم)
سکته قلبی ← نوشیدن قهوه	سکته قلبی یکی از علت‌های نوشیدن قهوه است	حقیقی	۳. سببی (معلول - علت)
عامل خارجی سکته قلبی ← نوشیدن قهوه	نوشیدن قهوه با یک عامل خارجی که علت سکته قلبی است مرتبط است.	حقیقی	۴. مخدوش شدن
سکته قلبی → نوشیدن قهوه	نوشیدن قهوه یکی از علت‌های سکته قلبی است	حقیقی	۵. سببی (علت - معلول)

دو عامل اول این جدول یعنی شانس و تورش باعث ایجاد روابط تصنعی می‌شوند. یعنی رابطه مشاهده شده فقط در نمونه مورد مطالعه برقرار است، ولی در جامعه وجود ندارد؛ اما هنگامی که یک رابطه حقیقی (به دور از شانس و تورش) مشاهده می‌شود، به جز رابطه سببی مورد نظر، دو گزینه دیگر نیز احتمال وقوع دارند. در مثال مورد بحث، سیگار کشیدن یکی از عوامل شناخته شده در بروز سکته قلبی است که با نوشیدن قهوه نیز مرتبط است. به همین علت سیگار کشیدن یکی از عوامل مخدوش کننده در بررسی رابطه سببی نوشیدن قهوه و سکته قلبی است.

۲-۵ مهار اثر شانس در روابط

فرض کنید هیچ ارتباطی بین نوشیدن قهوه و سکته قلبی وجود نداشته باشد و ۶۰٪ افراد جامعه قهوه می‌نوشند. اگر نمونه‌ای ۲۰ نفره از بیماران دچار عارضه سکته قلبی انتخاب شود، انتظار داریم ۱۲ نفر از آنها قهوه مصرف کنند. تنها بر اثر شانس امکان دارد در نمونه ۲۰ نفری ۱۹ نفر مصرف کننده قهوه وجود داشته باشد. در این صورت اگر باز هم بر اثر شانس، تعداد افراد مصرف کننده قهوه در گروه شاهد ۱۹ یا بیشتر باشد یک رابطه تصنعی بین مصرف قهوه و وقوع سکته قلبی مشاهده می‌شود. این روابط که تنها بر اثر خطاهای تصادفی (شانس) رخ می‌دهد به خطای نوع اول مشهور است.

روش‌های کاهش نقش خطاهای تصادفی در استنباطها هنگام طراحی و تجزیه و تحلیل نتایج امکان پذیر است. (جدول ۲-۵) از جمله این راهکارها، افزایش دقت اندازه گیری و افزایش حجم نمونه است، که در فصل‌های ۱ و ۴ بررسی شدند. محاسبه مقدار احتمال و ساختن فاصله اطمینان برای برآوردها، روش‌های آماری ارزیابی نقش شانس در روابط مشاهده شده هستند، که در بخش تجزیه و تحلیل اطلاعات کاربرد فراوان دارند. برای مثال مشاهده مقدار احتمال برابر ۰/۱ نشان دهنده این موضوع است که تفاوتی به بزرگی آنچه در مطالعه بین گروه‌های آزمودنی مشاهده شده در ۱۰٪ موارد، تنها بر اساس تصادف ایجاد می‌شود.

۳-۵ مهار اثر تورش در روابط

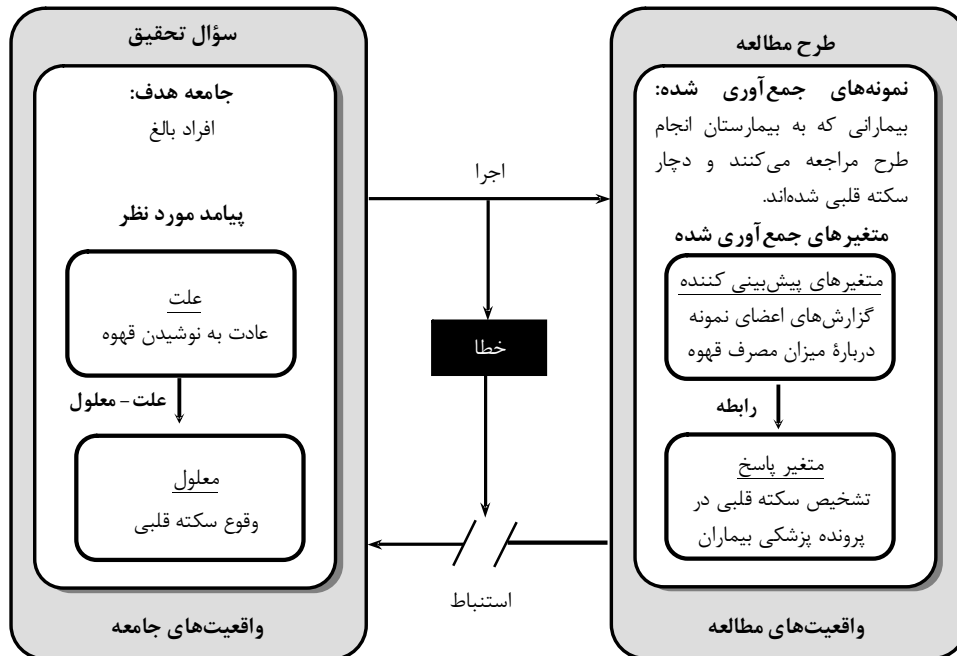
بروز رابطه تصنعی بر اثر تورش بسیار پیچیده تر و گمراه کننده تر است. تفاوت منظم بین سؤال تحقیق و آنچه محققان به بررسی و یافتن پاسخ برای آن می‌پردازند را تورش می‌نامند. راهکارهای مختلفی در مرحله طراحی و اجرای مطالعه وجود دارند که به کمک آنها می‌توان تورش را به حداقل رساند (جدول ۲-۵).

جدول ۲-۵: روش‌های کاهش مشاهده روابط تصنعی در مطالعات تحلیلی

منبع ایجاد رابطه تصنعی	روش مهار در مرحله طراحی	روش ارزیابی در مرحله تحلیل نتایج
تصادف	افزایش حجم نمونه به همراه راهکارهای ذکر شده در فصل‌های اول و چهارم	محاسبه و تعیین مقدار احتمال به عنوان یکی از مستندات مطالعه
تورش	توجه عمیق به منابع بالقوه وجود تفاوت بین سؤال پژوهش و طرح مطالعه که شامل افراد تحت مطالعه، متغیرهای مستقل و پیامد تحت بررسی است.	جمع‌آوری داده‌های تکمیلی جهت بررسی نقش تورش‌های بالقوه بر روابط مقایسه همخوانی نتایج به دست آمده با نتایج اطلاعات مشابه

دسته‌های مختلفی از تورش شناخته شده‌اند که نحوه تشخیص و کنترل آن‌ها در فصل‌های ۸ و ۹ بررسی می‌شوند. یک ایده جالب، نوشتن سؤال تحقیق و طرح مطالعه در کنار هم نظیر شکل ۵-۱ است. سپس سه مسئله زیر باید مورد توجه واقع شود.

- ۱- آیا نمونه‌های بررسی شده در مطالعه (مثلاً موردها و شاهد‌ها) نماینده مناسبی برای جامعه هدف هستند؟
- ۲- آیا اندازه‌گیری‌ها و ثبت متغیرهای زمینه‌ای و مستقل به اندازه کافی معتبر هستند؟
- ۳- آیا اندازه‌گیری‌ها و ثبت متغیر پاسخ به اندازه کافی معتبر است؟



شکل ۵-۱: به حداقل رساندن تورش بر اثر مقایسه سؤال تحقیق و طرح مطالعه

گام بعد بررسی راهکارهای مختلف در جلوگیری از وقوع هر کدام از تورش‌های بالقوه است. اگر تورش‌ها به آسانی قابل پیش‌گیری باشند، می‌توان طرح مطالعه را بهبود بخشید و سه سؤال بالا را مجدداً بررسی کرد و اگر به آسانی قابل پیش‌گیری نباشند باید ارزشمند بودن انجام مطالعه (مقایسه‌ها و استنباط‌های صورت گرفته) با وجود تورش‌های شناخته شده بالقوه مورد ارزیابی قرار گیرد.

معمولاً محققان پس از جمع‌آوری داده‌ها با یک یا چند منبع بالقوه تورش مواجه می‌شوند. برخی از این تورش‌ها در هنگام طراحی مطالعه شناسایی شده‌اند اما جلوگیری از آن‌ها بسیار دشوار است و برخی دیگر هنگامی شناسایی می‌شوند که زمان برای اتخاذ راهکار جهت جلوگیری از وقوع آن‌ها بسیار دیر است. در هر حالت، یکی از روش‌های متداول برای کنترل تورش‌ها، جمع‌آوری اطلاعات اضافی برای برآورد اندازه اثر تورش‌های بالقوه

است. برای مثال، فرض کنید محقق متوجه می‌شود که گروه شاهد بستری شده در بیمارستان به‌طور مناسبی نماینده جامعه هدف افراد بدون عارضه سکته قلبی نیست، زیرا آن‌ها مصرف قهوه را به علت ابتلا به بیماری‌های مزمن کاهش داده‌اند. اندازه این تورش ناشی از نحوه نمونه‌گیری، به‌وسیله تقسیم گروه کنترل به دو زیرگروه بیمارانی که ملزم هستند به علت نوع خاص بیماری (مثلاً زخم معده) مصرف قهوه را کاهش دهند و بیمارانی که چنین شرایطی ندارند قابل برآورد هستند. اگر هر دو زیرگروه تقریباً به‌اندازه افراد گروه مورد قهوه مصرف کنند، تورش ناشی از نحوه نمونه‌گیری اثر کمی بر یافته‌های تحقیق می‌گذارد. هم‌چنین هنگامی که محققان در کیفیت جمع‌آوری اطلاعات مربوط به مقدار مصرف قهوه دچار تردید هستند، یکی از روش‌ها برای بررسی این تورش، انتخاب زیرمجموعه‌ای از موردها و شاهد‌های مطالعه و بررسی مجدد مقدار مصرف قهوه آن‌ها به‌وسیله یک پرسشگر بی‌اطلاع از گروه (مورد یا شاهد) بیماران است. میزان توافق نتایج با اطلاعات قبلی اندازه اثر تورش ناشی از پایین بودن کیفیت داده‌ها را مشخص می‌سازد. یک روش متداول دیگر، مقایسه یافته‌های تحقیق با نتایج مطالعات مشابه است. اگر نتایج با یکدیگر همسو و مشابه باشند، احتمالاً روابط مشاهده‌شده به‌طور جدی تحت تأثیر تورش‌های بالقوه قرار نگرفته‌اند. این موضوع خصوصاً هنگامی که مطالعات دیگر با روش‌های دیگری انجام گرفته باشند مناسب است. زیرا در این صورت منابع تورش مشترک در مطالعات به حداقل می‌رسند. به‌طور کلی در موارد زیادی تورش‌های بالقوه یکی از مشکلات مهم تحقیق هستند. در فصل ۸ مطالب کامل‌تری درباره شناسایی، ارزیابی اثر و تعدیل نتایج در حضور تورش‌های مؤثر بر مطالعه بیان شده است.

۵-۴ روابط حقیقی به‌جز رابطه سببی

هنگامی که از تصنعی نبودن روابط مشاهده‌شده اطمینان حاصل می‌شود، به‌جز رابطه سببی دو گزینه دیگر یعنی رابطه اثر- علت (معلول-علت) و رابطه مخدوش شدن امکان وقوع دارند.

۵-۴-۱ رابطه اثر- علت یا معلول- علت

همیشه این امکان وجود دارد که آنچه محققان به‌عنوان علت تصور می‌کنند پیامد یا عارضه‌ای باشد که در اثر آنچه معلول خوانده می‌شده رخ داده است. به‌بیان‌دیگر جهت واقعی رابطه سببی عکس آنچه محققان فرض نموده‌اند، باشد. این موضوع در مطالعات مقطعی و مورد-شاهدی مسئله مهمی است. برای مثال فرض کنید در مطالعه‌ای سطوح بالای پروتئین^{۱۴} CRP در اثر وقوع عارضه سکته قلبی کاهش پیدا کند و به وضعیت عادی بازگردد در این صورت سکته قلبی باعث بالا رفتن سطح پروتئین CRP می‌شود نه عکس آن.

بروز پدیده اثر- علت در مطالعات هم‌گروهی بسیار نادر است، زیرا میزان مواجهه با عوامل خطر پیش از شروع بیماری بارها اندازه‌گیری می‌شوند. باین‌حال وقوع آن در برخی بیماری‌ها که دارای دوره نهفتگی طولانی و بدون علائم قابل تشخیص هستند، محتمل است. یکی از مثال‌های خوب در این زمینه، رابطه مشاهده‌شده بین سطوح پایین کلسترول و افزایش مرگ‌ومیر ناشی از سرطان است که در اکثر مطالعات هم‌گروهی مشاهده‌شده

۱. C-Reactive Protein

است. رابطه مشاهده شده با طولانی تر شدن زمان پیگیری قوی تر می شود. این موضوع باعث شکل گیری این فرضیه شد که دوره نهفتگی سرطان در برخی اشخاص باعث کاهش سطح کلسترول می شود، که مثالی از پدیده اثر - علت است، زیرا مطالعات بعدی نشان داد مرگ و میر ناشی از سرطان با پایین آمدن سطح کلسترول همراه است [۱] و [۲] و [۳].

۵-۴-۲ مخدوش شدن

یکی دیگر از گزینه ها، هنگامی که یک رابطه واقعی مشاهده می شود، متأثر بودن رابطه مشاهده شده به وسیله یک عامل خارجی است. نحوه تشخیص عوامل مخدوش کننده و راه های کنترل آن ها در فصل های ۶ و ۷ مورد بررسی قرار می گیرند. کنترل و یا احتمالاً حذف اثر متغیرهای مخدوش کننده در دو مرحله طراحی و تجزیه و تحلیل داده ها امکان پذیر است. در مرحله طراحی، محدودسازی، همسان سازی و در مطالعات تجربی تصادفی کردن، روش های متداول برای حذف اثر متغیرهای مخدوش کننده هستند. در جدول ۵-۳ مزایا و معایب این روش ها بررسی شده اند.

جدول ۵-۳: کنترل و یا حذف اثر مخدوش کنندگی در مرحله طراحی مطالعات

راهبرد	مزایا	معایب
تصادفی کردن	اگر حجم نمونه به اندازه کافی بزرگ باشد انتساب تصادفی باعث تعدیل تمام متغیرهای تأثیرگذار (چه آن ها که توسط محققان اندازه گیری می شوند چه آن هایی که اندازه گیری نمی شوند و یا احیاناً محققان از وجود آنان اطلاعی ندارند) می شود.	به جز مطالعات تجربی در سایر مطالعات قابل به کارگیری نیست. برای اجرای این روش باید ملاحظات زیادی از قبیل مسائل اخلاقی، هزینه و زمان را در نظر داشت.
محدود کردن	فهم و اجرای آن ساده است.	باعث محدود شدن تعمیم پذیری نتایج می شود.
همسان سازی	اثر برخی متغیرهای زمینه ای نظیر سن و جنس (و به همراه آن ها توزیع پارامترهایی که با آن عوامل مرتبط هستند) تا حد زیادی در گروه های تحت مطالعه همانند می شوند. با تعدیل نسبی مخدوش کننده ها برای موردها و شاهد ها در هر رده دقت (توان آماری) مطالعه افزایش می یابد.	ممکن است پرهزینه و وقت گیر باشد. برخی مواقع افزایش حجم نمونه کارا تر است. به علت آنکه همسان سازی قبل از انجام مطالعه صورت می گیرد و مخدوش کننده ها به طور قطعی مشخص نیستند، ممکن است اثرات سو بر تحلیل ها و استنباط ها داشته باشد. برای مخدوش کننده هایی که همسان سازی بر اساس آن ها صورت گرفته شده است امکان محاسبه خطر نسبی و یا مدل سازی آماری وجود ندارد. تحلیل این مطالعات نیاز به روش های ویژه دارد. امکان وقوع پدیده بیش همسان سازی وجود دارد (همسان سازی برای عواملی که واقعاً مخدوش کننده نیستند).

طبقه‌بندی و تحلیل چند متغیره (مدل‌سازی) داده‌ها روش‌های مرسوم برای تعدیل اثر متغیر مخدوش‌کننده در مرحله تجزیه و تحلیل نتایج هستند. جدول ۴-۵ به بررسی مزایا و معایب این روش‌ها می‌پردازد.

جدول ۴-۵: تعدیل اثر مخدوش‌کنندگی در مرحله تجزیه و تحلیل اطلاعات

راهبرد	مزایا	معایب
طبقه‌بندی	فهم و اجرای آن ساده است.	ممکن است تعداد طبقات نسبت به حجم نمونه بسیار زیاد شود. همه متغیرهای مورد بررسی باید به صورت اسمی باشند این موضوع در مورد متغیرهای پیوسته (نظیر سن، فشارخون) می‌تواند باعث از دست رفتن قسمتی از اطلاعات و یا طبقه‌بندی نامناسب شود.
تحلیل چند متغیره	امکان تعدیل هم‌زمان چند مخدوش‌کننده وجود دارد. امکان بررسی متغیرهای پیوسته وجود دارد. این شیوه‌ها بسیار متنوع و انعطاف‌پذیر هستند.	ممکن است مدل انتخاب‌شده برای بررسی رابطه متغیر پاسخ و مخدوش‌کننده‌های مطالعه مناسب نباشد در این صورت مخدوش‌کننده‌ها به درستی کنترل و تعدیل نمی‌شوند. در صورت انتخاب مدل نامناسب و یا بی‌دقتی در مدل‌سازی، برآوردهای نادرستی از اندازه اثر عامل خطر بیماری محاسبه می‌شود. ممکن است به کارگیری و فهم نتایج دشوار باشد.

استفاده از روش محدود کردن برای کنترل اثر مخدوش‌کنندگی هنگامی که هدف اصلی مطالعه بررسی برخی از زیرگروه‌های جامعه است بسیار کارا و مفید است. یکی از تصمیم‌های مهم در هر مطالعه، تصمیم‌گیری درباره انجام همسان‌سازی است. معمولاً استفاده از روش همسان‌سازی برای کنترل متغیرهای زمینه‌ای نظیر سن، جنس و نژاد که دارای اثر متقابل با سایر متغیرها نیستند، بسیار مؤثر است. این روش در مواردی که حجم نمونه نسبت به تعداد متغیرهای مخدوش‌کننده در مطالعه کم است نیز می‌تواند مفید واقع شود. یکی دیگر از موارد مهم، توجه به این موضوع است که آیا همسان کردن متغیرها آسان‌تر است یا اندازه‌گیری آن‌ها، اندازه‌گیری برخی متغیرها می‌تواند دشوار و یا حتی ناممکن باشد. تصمیم‌گیری درباره تحلیل نتایج به روش طبقه‌بندی یا تحلیل چند متغیره را می‌توان تا زمان تکمیل داده‌ها به تأخیر انداخت. در این صورت می‌توان ابتدا مخدوش‌کننده بودن عوامل را بررسی کرده و سپس راجع به نحوه تعدیل داده‌ها تصمیم گرفت [۴].

۵-۵ استنباط رابطه اثر - علت بر اساس قراین

در بخش‌های قبلی روش‌های شناسایی و ارزیابی روابط سببی سایر روابط بررسی شدند. در این قسمت به نحوه ارزیابی رابطه سببی بر اساس شواهد و قراین موجود می‌پردازیم. به‌طور کلی موارد اصلی برای تأیید سببی بودن یک رابطه عبارت‌اند از: رابطه زمانی، ثبات (تکرارپذیری و سازگاری) یافته‌ها، توجیه بیولوژیکی. در کنار این عوامل، شواهدی دیگر نظیر رابطه دوز- پاسخ، قدرت همبستگی، قطع مواجهه، اختصاصی بودن همبستگی و

توضیحات دیگری که می‌شود ارائه کرد، نیز می‌تواند در استنباط رابطه سببی مؤثر باشد. [۴-۸]

می‌دانیم اگر عامل در ایجاد بیماری نقش داشته باشد باید مواجهه با آن قبل از ظهور نشانه‌های بیماری صورت گرفته باشد. در اغلب اوقات نشان دادن رابطه زمانی در مطالعات هم‌گروهی آینده‌نگر ساده‌تر از مطالعات مورد- شاهدهی و یا هم‌گروهی گذشته‌نگر است. رابطه زمانی مواجهه و ایجاد بیماری نه تنها از نظر ترتیب وقوع مهم است، بلکه از نظر فاصله زمانی بین مواجهه و شروع نشانه‌های بالینی بیماری نیز اهمیت دارد. هنگامی که تکرار مطالعه، با روش‌ها و طرح‌های مختلف باعث مشاهده نتایج سازگار می‌شود امکان مشاهده روابط تصادفی یا تحریف شده بر اثر تورش‌ها ضعیف می‌شود. روابط حقیقی یعنی سببی، اثر- علت یا مخدوش شدن معمولاً دارای ثبات هستند. برای مثال اگر مصرف سیگار و نوشیدن قهوه رابطه آماری با جهت مثبت در جامعه داشته باشد، انتظار داریم همواره در مطالعات مختلف وجود رابطه بین نوشیدن قهوه و بروز سکتة قلبی تأیید شود. یکی از عوامل مهم برای وجود رابطه سببی، وجود مقبولیت زیستی است. اگر یک ساختار سببی که از لحاظ علمی موجه باشد برای رابطه مشاهده شده وجود داشته باشد، دلیل و مدرک محکمی برای سببی بودن رابطه فراهم می‌شود. وقتی یک عامل باعث ایجاد بیماری می‌شود، انتظار داریم وقتی مواجهه با آن عامل کاهش یافته و یا به‌طور کامل قطع می‌شود خطر ابتلا به بیماری نیز کاسته شده و یا از بین برود.

وجود رابطه قوی‌تر، شاهد محکم‌تری به نفع مشاهده رابطه مستقیم است. روابط مشاهده شده بر اثر وجود عوامل مخدوش‌کننده به علت اینکه غیرمستقیم هستند عموماً از روابط مستقیم نظیر رابطه سببی ضعیف‌تر هستند. علاوه بر این، روابط قوی‌تر بیشتر اوقات منجر به مشاهده مقدار احتمال معنی‌دار می‌شوند که این موضوع احتمال تصادفی بودن رابطه را کاهش می‌دهد.

وجود رابطه دوز- پاسخ یکی از شواهد مثبت برای وجود رابطه سببی است. برای مثال در سیگاری‌های معمولی شیوع سرطان ریه نسبت به افراد عادی بالاتر است و در سیگاری‌های حرفه‌ای شیوع سرطان از هر دو گروه بیشتر است. هرگاه که امکان‌پذیر باشد، باید متغیرهای پیش‌بینی‌کننده پیامد را به‌صورت پیوسته یا در چندطبقه اندازه‌گیری نمود تا امکان مشاهده رابطه دوز- پاسخ فراهم شود. باین‌حال امکان مشاهده اثر دوز- پاسخ در روابط سببی و یا روابط تحریف شده به‌وسیله مخدوش‌ها نیز امکان‌پذیر است. برای مثال اگر افرادی که بیشتر سیگار می‌کشند بیشتر هم قهوه مصرف کنند امکان دارد اثر دوز- پاسخ در رابطه بین مصرف قهوه و وقوع سکتة قلبی هم مشاهده شود.

اگر همبستگی سببی باشد، انتظار داریم یافته‌های مشاهده‌شده با یافته‌های سایر مطالعات همخوانی داشته باشد. به‌عنوان مثال اگر سیگار کشیدن رابطه سببی با ابتلا به سرطان ریه داشته باشد باید این رابطه هم در زن‌ها و هم در مرد‌ها در مطالعات جداگانه مشاهده شود.

برای فهم بهتر نحوه انجام استنباط سببی به بررسی رابطه بین هلیکوباکتریلوری و ابتلا به زخم معده که در کتاب اپیدمیولوژی تألیف گوردیس به آن پرداخته شده را بررسی می‌کنیم [۵].

مطالعاتی که از سال ۱۹۸۲ در زمینه عفونت با هلیکوباکتریلوری، یک باکتری گرم منفی شروع شد، نشان داد که این عفونت با گاستریت‌های فعال مزمن در ارتباط است. مطالعات بعدی نشان داد که این باکتری در ایجاد زخم دستگاه گوارش نقش دارد. در جدول ۵-۵ شواهد و قراین ممکن درباره سببی بودن همبستگی مشاهده شده گردآوری شده است. همان‌طور که مشاهده می‌شود شواهد موجود سببی بودن رابطه را تقویت می‌کند؛ اما شواهد موجود هنوز برای انجام استنباط سببی کافی نیست.

جدول ۵-۵: ارزیابی شواهد ارائه شده برای وجود رابطه سببی بین عفونت با هلیکوباکتریپیلوری و زخم دئودنوم

رابطه زمانی
هلیکوباکتریپیلوری به شکل واضح با گاستریت مزمن مرتبط است. حدود ۱۱٪ از بیماران مبتلا به گاستریت مزمن در طول ۱۰ سال به زخم دئودنوم دچار می‌شوند. در یک مطالعه که ۴۵۴ بیمار آندوسکوپی شده بودند، پس از ۱۰ سال ۳۴ نفر از ۳۲۱ نفر که حامل هلیکوباکتریپیلوری بوده‌اند (۱۱٪) به زخم دئودنوم مبتلا شده‌اند در حالی که از ۱۳۳ نفری که به این باکتری آلوده نبودند تنها یک نفر به زخم دئودنوم مبتلا شده است (۰/۸ درصد).
قدرت رابطه
هلیکوباکتریپیلوری از ۹۰٪ از بیماران مبتلا به زخم دئودنوم جدا شده است. حداقل در یک جامعه که گزارشی از ابتلا به زخم روده نداشته‌اند (از یک قبیله بومی در قسمت شمالی استرالیا که از جوامع دیگر جدا بودند)، هلیکوباکتریپیلوری هرگز جدا نشده است.
رابطه دوز - پاسخ
تراکم هلیکوباکتریپیلوری در افراد مبتلا به زخم دئودنوم در هر میلی‌متر مربع بیشتر از بیمارانی است که مبتلا به این زخم نیستند (به قدرت رابطه نیز توجه شود).
تکرار پذیری یافته‌ها
بسیاری از مشاهدات مربوط به این یافته تکرارپذیر بوده و در واقع تکرار شده است.
توجیه بیولوژیکی
باکتری‌ها به سلول‌های ابتدایی روده باریک متصل شده و همراه آن‌ها داخل محوطه روده می‌شوند. هلیکوباکتریپیلوری واسطه‌های التهابی را نیز القا می‌کند. مخاط‌های آلوده به هلیکوباکتریپیلوری ضعیف شده و نسبت به اثرات زیان‌بخش اسید معده حساس‌تر می‌شوند.
همخوانی با سایر اطلاعات
شیوع عفونت هلیکوباکتریپیلوری نزد مردها و زن‌ها یکسان است. بروز زخم روده که در گذشته گفته می‌شد نزد مردها بیشتر از زن‌ها است، در سال‌های اخیر یکسان گزارش می‌شود.
قطع مواجهه
ریشه‌کنی عفونت هلیکوباکتریپیلوری، با همان سرعت درمان با آنتاگونیست‌های هیستامینیک، منجر به بهبود زخم روده می‌شود. ریشه‌کنی عفونت هلیکوباکتریپیلوری با استفاده هم‌زمان از سه دارو، باعث می‌شود عود زخم روده حتی پس از مدت‌های طولانی وجود نداشته باشد در صورتی که درمان با آنتاگونیست‌های هیستامینیکی بین ۶۰ تا ۸۰ درصد عود دارد.
اختصاصی بودن همبستگی
شیوع عفونت هلیکوباکتریپیلوری نزد مبتلایان به زخم قسمت اول روده باریک ۹۰ تا ۱۰۰ درصد است. این عفونت نزد بعضی از بیماران مبتلا به زخم معده و افراد سالم هم گزارش شده است.
بررسی توضیحات دیگر
نتایج به‌دست‌آمده از مطالعات نشان می‌دهد که سیگاری بودن بیماران آلوده به هلیکوباکتریپیلوری خطر ایجاد زخم روده در آن‌ها را افزایش می‌دهد ولی این عامل نزد کسانی که این عفونت در آن‌ها ریشه‌کن شده است مخاطره را افزایش نمی‌دهد.

منابع

1. Jacobs D, Blackburn H, Higgins M, et al. Report of the conference on low blood cholesterol: mortality associations. *Circulation*, 1992, 86:1046-60.
2. Pekkanen J, Nissinen A, Punsar S, Karuonen MJ. Short-and long-term association of serum cholesterol with mortality: the 25-Year follow-up of the Finnish cohort of the seven countries. *Am J Epidemiol*, 1992, 135:1251-8.
3. Sharp SJ, Pocock SJ. Time trends in serum cholesterol before cancer death. *Epidemiology*, 1997, 8:132-6.
4. Hulley SB, et al. *Designing Clinical research: An epidemiologic approach*. (2th ed.). Lippincott Williams & Wilkins, 2001.
5. Gordis L. *Epidemiology*. (2 nd ed.) WB Saunders. London 2000.
6. Elwood JM. The Diagnosis of Causation. In: *Causal Relationships in Medicine. A Practical System for Critical Appraisal*. New York: Oxford University Press, 1988: 163-182.
7. Soori H. Children indoor and outdoor play pattern in Ahwaz city: implications for injury prevention. *Eastern Meditr Health J*. 2006;12(3/4): 19-28.

۸. نوح جاه ص، سوری ح. بررسی رابطه مصرف قرص‌های ترکیبی با دو قلو زایی در حاملگی بعدی. *مجله علمی پژوهشی فیض*. شماره ۳۴، تابستان ۱۳۸۴: ۵۷-۵۳.



ارزیابی نقش مخدوش کننده در مطالعات تحلیلی

هنگامی که به بررسی وجود ارتباط بین یک عامل خطر و بیماری می‌پردازیم، معمولاً عامل سومی را می‌توان یافت که دارای اثرات مهمی بر رابطه مشاهده شده بین عامل خطر و بیماری باشد. اگر این عامل (حداقل به‌طور نسبی) قابلیت توجیه رابطه مشاهده شده را داشته باشد، پدیده مخدوش‌کنندگی بروز می‌کند. برای مثال، ممکن است رابطه‌ای بین تعداد فرزندان و شیوع سرطان پستان در نمونه‌ای از مادران مشاهده شود که به‌وسیله سن مادران قابل توجیه باشد. مادران مسن‌تر فرزندان بیشتری دارند و همچنین افزایش سن، باعث افزایش شانس ابتلا به سرطان پستان می‌شود. در این صورت سن عامل سومی وجود دارد که رابطه مشاهده شده بین تعداد فرزندان و شیوع سرطان پستان را توجیه می‌کند. جدول (۱-۶) داده‌های فرضی یک عامل خطر و پیامد آن (بیمار/ سالم) را نمایش می‌دهد.

جدول ۱-۶: وضعیت عامل خطر و وجود یا عدم وجود بیماری در یک مثال فرضی

وضعیت بیماری			عامل خطر
خطر	سالم	بیمار	
۰/۷۳۶۴	۲۹	۸۱	در معرض
۰/۱۳۳۳	۱۸۲	۲۸	بدون تماس
	۵/۵۲		خطر نسبی

همان‌طور که ملاحظه می‌شود خطر نسبی (وجود عامل خطر در برابر عدم وجود آن) برابر ۵/۵۲ و این رابطه از لحاظ آماری معنی‌دار است ($P < 0.001$). در این صورت ممکن است نتیجه‌گیری شود که عامل خطر واقعاً مؤثر بر وضعیت بیماری است.

حال فرض کنید اطلاعات مربوط به متغیر سومی که بالقوه مخدوش‌کننده است، نیز جمع‌آوری شده است. این عامل را با C نمایش می‌دهیم. فرض کنید وجود یا عدم وجود ویژگی C اندازه‌گیری و این اطلاعات در جدول (۲-۶) آورده شده است.

جدول ۶-۲: وضعیت عامل خطر و بیماری به تفکیک مخدوش کننده فرضی در داده‌های مربوط به جدول ۶-۱

وضعیت عامل خطر			وجود مخدوش کننده			عدم وجود مخدوش کننده		
خطر	بیمار	سالم	خطر	بیمار	سالم	خطر	بیمار	سالم
در معرض	۱	۹	۰/۱۰۰	۸۰	۲۰	۰/۸۰۰		
بدون تماس	۲۰	۱۸۰	۰/۱۰۰	۸	۲	۰/۸۰۰		
خطر نسبی	۱			۱				

در این جدول ارتباط بین عامل خطر و وضعیت بیماری به‌طور جداگانه برای افرادی که با مخدوش کننده C مواجهه داشته و آن‌ها که بدون مواجهه بوده محاسبه شده است. در مورد افراد دارای مواجهه، خطر برای هر دو گروه در معرض و بدون تماس برابر ۰/۱ است در گروه بدون مواجهه نیز خطر برای هر دو گروه مساوی ۰/۸ است. بنابراین خطر نسبی در هر دو سطح متغیر مخدوش کننده برابر می‌شود. نگاه عمیق‌تر به جدول امکان توجیه این پدیده را میسر می‌کند. متغیر مخدوش کننده و عامل خطر به هم وابسته هستند و عامل مخدوش کننده به‌تنهایی یک عامل خطر برای بیماری به حساب می‌آید زیرا خطر نسبی آن برابر

$$\frac{(۸۰ + ۸) / (۸۰ + ۸ + ۲۰ + ۲)}{(۱ + ۲۰) / (۱ + ۲۰ + ۹ + ۱۸۰)} = ۸$$

است. در این صورت هنگامی که تصور می‌شد عامل خطر بر بیماری تأثیر داشته، در واقع اثر مخدوش کننده بر بیمار مشاهده شده است.

جدول ۶-۳ نتایج یک مطالعه فرضی دیگر را نمایش می‌دهد. در این مسئله، قبل از آنکه اثر مخدوش کننده را منظور کنیم، عامل خطر بالقوه بی‌اثر است.

جدول ۶-۳: وضعیت عامل خطر و بیماری در داده‌های فرضی

وضعیت بیماری			عامل خطر
خطر	سالم	بیمار	
در معرض	۴۲۰	۲۴۰	۰/۳۶۳۶
بدون تماس	۳۵۰	۲۰۰	۰/۳۶۳۶
خطر نسبی	۱/۰۰		

در جدول ۶-۴ خطر و خطر نسبی ابتدا به بیماری برای همان اطلاعات، با توجه به یک مخدوش کننده فرضی نشان داده شده است.

جدول ۶-۴: جدول توافقی سه طرفه مربوط به عامل خطر، وضعیت بیماری و مواجهه با یک مخدوش کننده فرضی.

عدم وجود مخدوش کننده		وجود مخدوش کننده			وضعیت عامل خطر	
خطر	سالم	بیمار	خطر	سالم		بیمار
۰/۹۵۴۵	۵	۱۰۵	۰/۳۴۵۵	۴۱۵	۱۳۵	در معرض
۰/۳۹۰۰	۳۰۵	۱۹۵	۰/۱۰۰۰	۴۵	۵	بدون تماس
۲/۴۵			۲/۴۵			خطر نسبی

همان طور که ملاحظه می شود در سطوح مختلف متغیر مخدوش کننده خطرهای نسبی مشاهده شده از ۱ بزرگ تر است. خطر نسبی در هر دو رده متغیر مخدوش کننده (وجود یا عدم وجود مخدوش کننده) برابر ۲/۴۵ با سطح معنی داری ($P = ۰/۰۳$) و ($P < ۰/۰۰۱$) است.

در این مثال وجود عامل مخدوش کننده و عدم تماس با عامل خطر دارای وابستگی مثبت هستند. نظیر مثال قبل در این داده ها نیز عامل مخدوش کننده به طور مجزا یک عامل خطر به حساب می آید. خطر نسبی مواجهه با مخدوش کننده نسبت به عدم مواجهه برابر است با

$$\frac{(۱۰۵ + ۱۹۵)/(۱۰۵ + ۱۹۵ + ۵ + ۳۰۵)}{(۱۳۵ + ۵)/(۱۳۵ + ۴۱۵ + ۴۵ + ۵)} = ۲/۱۱$$

هنگامی که مخدوش کنندگی کامل وجود داشته باشد خطرهای نسبی (و یا هر برآوردی از نسبت شانس برای ابتلا به بیماری) برای سطوح مختلف، (نه الزاماً دو سطح) متغیر مخدوش کننده یکسان است. این مقدار مشترک هنگامی که متغیر مخدوش کننده در تحلیل نتایج در نظر گرفته می شود، متفاوت است. مثال های قبل هر دو نمایش دهنده مخدوش کنندگی کامل هستند. مخدوش کنندگی کامل به ندرت در داده ها دیده می شود. آنچه در واقعیت مشاهده می شود درجاتی از مخدوش کنندگی است که باعث برآورد بیش از اندازه واقعی یا کمتر از اندازه واقعی خطر می شود. با این حال اگر برآورد خطر در سطوح مختلف متغیر مخدوش کننده مشابه باشد و یا برآوردها در هر رده با برآورد خطر کلی تقریباً یکسان باشد، می توان استدلال کرد که عامل مخدوش کننده اثر چندانی ندارد. جدول زیر از داده های واقعی مربوط به شش سال پیگیری در مطالعه سلامت و بیماری های عروق تاجی قلب (CHD) در اسکاتلند به دست آمده است [۱].

جدول ۶-۵: وضعیت مسکن و وقوع CHD در مردان

وقوع CHD			وضعیت مسکن
خطر	ندارد	دارد	
۰/۰۴۴۶	۱۸۲۱	۸۵	اجاره ای
۰/۰۳۱۱	۲۴۰۰	۷۷	تملیکی
۱/۴۳			خطر نسبی

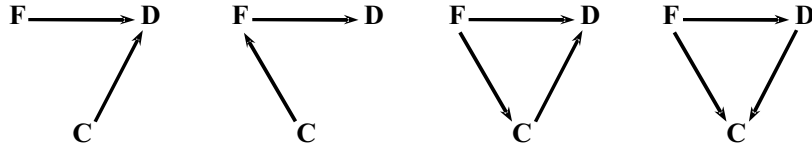
این اطلاعات مربوط به وجود علائم CHD و وضعیت مسکن است. همان طور که ملاحظه می شود، شانس وقوع

CHD در افرادی که دارای مسکن اجاره‌ای هستند بیشتر است. سؤالی که پس از مشاهده این رابطه به ذهن می‌رسد این است که آیا شیوه زندگی نامناسب در مردانی که دارای مسکن اجاره‌ای هستند می‌تواند این رابطه را توجیه کند یا خیر. به‌طور خاص ۵۷٪ از مردانی که خانه اجاره‌ای داشته‌اند و ۳۷٪ از مردانی که خانه تملیکی داشته‌اند سیگار می‌کشند. می‌دانیم که سیگار کشیدن یکی از عوامل خطر در وقوع CHD است. در جدول (۶-۶)، جدول توافقی سه‌طرفه برای نمایش هم‌زمان وضعیت مسکن، وقوع CHD و مصرف سیگار نمایش داده شده است.

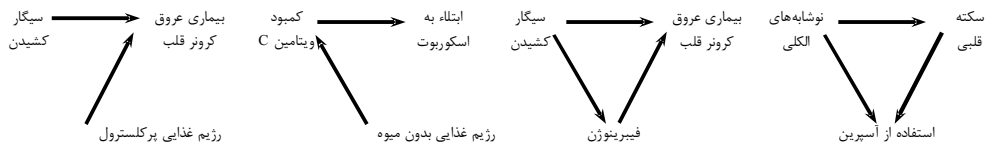
جدول ۶-۶: داده‌های به تفکیک استعمال و عدم استعمال سیگار آورده شده است.

وضعیت مسکن	عدم مصرف سیگار			مصرف سیگار		
	وقوع CHD	عدم وقوع CHD	خطر	وقوع CHD	عدم وقوع CHD	خطر
اجاره‌ای	۳۳	۹۲۳	۰/۰۳۴۵	۸۹۸	۰/۰۵۴۷	۹۲
تملیکی	۴۸	۱۷۲۲	۰/۰۲۷۱	۶۷۸	۰/۰۴۱۰	۲۹
خطر نسبی	۱/۲۷			۱/۳۳		

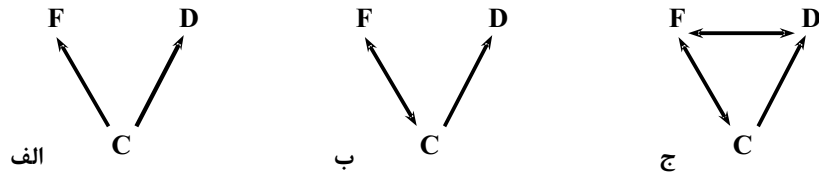
همان‌طور که از جدول بالا ملاحظه می‌شود، زندگی در خانه اجاره‌ای حتی پس از تعدیل مخدوش‌کننده مصرف سیگار هنوز هم یک عامل خطر به حساب می‌آید. هرچند که میزان خطر نسبی پس از تعدیل اندکی کاهش یافته است؛ اما به علت ناچیز بودن این کاهش می‌توان استنباط کرد اثر مخدوش‌کنندگی سیگار کشیدن کوچک است. مخدوش شدن در داده‌های واقعی به علت وجود رابطه بین متغیرها و یا نحوه جمع‌آوری داده‌ها ظاهر می‌شود. برای مثال اگر مصرف مواد مخدر و استفاده از نوشابه‌های الکلی با همدیگر وابسته باشند در این صورت اثر بیماری مرتبط با مصرف نوشابه‌های الکلی با اثر مصرف مواد مخدر مخدوش می‌شود. مثال دیگری در نظر بگیرید که برای بررسی کارایی یک داروی پیشگیری‌کننده از یک بیماری، مطالعه‌ای تجربی طراحی می‌شود و افراد تحت مطالعه به‌طور تصادفی داروی مؤثر یا دارونما دریافت می‌کنند. فرض کنید در گروه دریافت‌کننده دارو تصادفاً نسبت مردها بیشتر از زن‌ها و در گروه دریافت‌کننده دارونما نسبت زن‌ها بیشتر از مردها باشد. در این صورت اگر شیوع بیماری موردنظر در مردها بیشتر از زن‌ها باشد، به احتمال بسیار زیاد اثر پیشگیری‌کنندگی دارو کمتر از مقدار واقعی برآورد می‌شود. این اتفاق به علت مخدوش شدن نتایج با اثر جنس رخ می‌دهد. این نوع مخدوش شدن به علت نوع طراحی مطالعه رخ می‌دهد در بدترین حالت، تمام افراد گروه دارونما زن و افراد داروی مؤثر مرد هستند. در این حالت اثر داروی مؤثر و تأثیر جنس غیرقابل تشخیص می‌شود. نمودارهای ۶-۱ تا ۶-۴ موقعیت‌های مختلفی که عامل C می‌تواند مخدوش‌کننده یا تنها یک پیامد باشد را نمایش می‌دهد در این نمودارها پیکان‌ها رابطه بین دو عامل تحت بررسی را صرف‌نظر از سایر روابط بین عوامل موردبحث، مشخص می‌کند. پیکان‌های دوطرفه نشان‌دهنده رابطه غیرسببی است و پیکان‌های یک‌طرفه جهت رابطه سببی را معین می‌کند [۲] و [۳] و [۴].



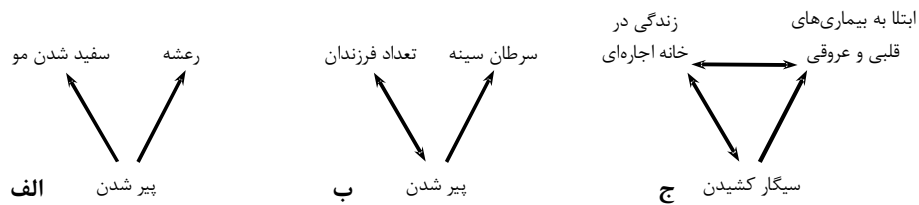
نمودار ۱-۶: برخی از موقعیت‌ها که C برای رابطه F-D مخدوش کننده است



نمودار ۲-۶: مثال‌هایی برای نمودار ۱-۶

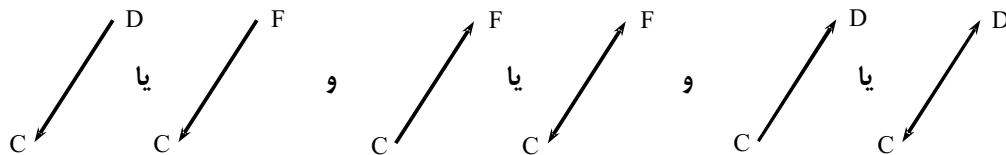


نمودار ۳-۶: برخی از موقعیت‌ها که C برای رابطه F-D مخدوش کننده نیست



نمودار ۴-۶: مثال‌هایی برای نمودار ۳-۶

به‌طور کلی برای تصمیم‌گیری راجع به این موضوع که یک عامل می‌تواند مخدوش کننده بالقوه به حساب آید یا خیر، رسم این نمودارها بسیار مفید است. شرایط لازم برای مخدوش کننده بودن در نمودار ۵-۶ نمایش داده شده است.



نمودار ۶-۵: شرایط لازم برای مخدوش کننده بودن C در رابطه F-D

هر متغیری که شرایط موجود در این نمودار را داشته باشد احتمالاً باید مخدوش کننده تلقی شود. هنگامی که از برقرار بودن یکی از روابط مطمئن نیستیم، بهترین روش ممکن فرض بر وجود ارتباط در هنگام طراحی و جمع آوری داده‌ها و ارزیابی این مفروضات پس از جمع آوری داده‌ها است. به همین علت معمولاً باید تمام عوامل خطر شناخته شده برای یک بیماری را مخدوش کننده بالقوه در نظر گرفت. پس از تکمیل مراحل اجرایی مطالعه روش‌های آماری متنوعی برای ارزیابی اثر متغیرهای مخدوش کننده بالقوه وجود دارد. این موضوع به‌طور کامل در فصل بعد مورد بررسی قرار می‌گیرد.

۶-۱ ویژگی‌های عامل مخدوش کننده

در حالت کلی، عامل مخدوش کننده باید هم با عامل مواجهه و هم با بیماری تحت مطالعه وابستگی داشته باشد. با این حال این ویژگی به‌طور کامل عامل مخدوش کننده را مشخص نمی‌کند. مواردی وجود دارد که متغیری واجد این همبستگی‌ها است اما مخدوش کننده محسوب نمی‌شود. این پدیده به روش‌های مختلف امکان وقوع دارد. شایع‌ترین طریق رخ دادن این پدیده در مواقعی است که مواجهه مورد مطالعه دارای عوارض باشد. در این صورت عوارض به علت وابستگی مواجهه و بیماری، با بیماری نیز همبسته می‌شود. برای مثال فرض کنید فراوانی مصرف نوشابه با میزان مصرف پیتزا وابسته باشد و مصرف نوشابه یکی از عوامل خطر در ایجاد سرطان مثانه باشد. در این صورت آیا مصرف پیتزا یک مخدوش کننده به حساب می‌آید؟ در ابتدا شاید پاسخ مثبت به نظر برسد زیرا مصرف پیتزا هم با مصرف نوشابه و هم با ایجاد سرطان مثانه مرتبط است؛ اما اگر مصرف پیتزا تنها به صورت یک طرفه و به علت وابستگی با مصرف نوشابه با ایجاد سرطان مثانه در ارتباط باشد یک عامل مخدوش کننده محسوب نمی‌شود. مخدوش کننده باید به تنهایی نیز یک عامل مؤثر بر وقوع بیماری باشد. یعنی اگر مصرف پیتزا یکی از عوامل مؤثر بر سرطان مثانه در بین افرادی که نوشابه مصرف نمی‌کنند باشد، آنگاه می‌توان این عامل را به عنوان مخدوش کننده در نظر گرفت.

۶-۱-۱ معیارهای لازم برای مخدوش کننده بودن یک متغیر

سه معیار زیر باید برای هر متغیر صادق باشد تا بتوان آن را به عنوان مخدوش کننده در نظر گرفت.

- ۱- هر عامل مخدوش کننده باید یک عامل خطر برای بیماری باشد. هر عامل مخدوش کننده بالقوه باید یکی از علت‌های اساسی برای بیماری و یا یک پیش‌آگهی مرتبط با یکی از علل اصلی بیماری باشد. وابستگی موجود بین مخدوش کننده بالقوه و بیماری باید مستقیم باشد. وابستگی‌هایی از نوع غیرمستقیم که ناشی از رابطه با

یکی از مواجهه‌های ایجادکننده بیماری نشئت گرفته باشد، سبب مخدوش‌کنندگی نمی‌شوند. استفاده از داده‌ها برای مشاهده رابطه بین مخدوش‌کننده بالقوه و بیماری می‌تواند راهنمای مناسبی باشد. رابطه مشاهده‌شده از ساختار داده‌ها ماهیت متغیرها را از نظر مخدوش‌کننده بودن مشخص نمی‌کند، بلکه وجود یک رابطه قطعی و اثبات‌شده از نظر تئوری، ماهیت مخدوش‌کنندگی را تعیین می‌کند. در مطالعاتی با حجم نمونه بزرگ که با خطاهای نمونه‌گیری کوچک‌تری مواجه هستیم انتظار داریم رابطه به‌دست‌آمده از داده‌ها به واقعیت نزدیک‌تر باشد. در صورتی که در مطالعات محدود، داده‌ها ممکن است به‌اندازه کافی قابل‌اعتماد نباشند.

۲- عامل مخدوش‌کننده باید در جمعیت مبنا (جمعیت در معرض خطر که بیماران را نیز شامل می‌شود) نیز با عامل مواجهه وابسته باشد. وابستگی بین مخدوش‌کننده بالقوه و عامل مواجهه نباید به‌صورت ثانویه و از راه وابستگی بین مواجهه و بیماری برقرار شود. به‌عنوان مثال در مطالعات هم‌گروهی وابستگی بین عامل مخدوش‌کننده بالقوه و مواجهه باید در ابتدای مطالعه و در هنگامی که هنوز بیماری ایجاد نشده است وجود داشته باشد. به همین علت در مطالعه هم‌گروهی، وابستگی بین مواجهه و مخدوش‌کننده باید به‌وسیله اطلاعات موجود و نه به‌وسیله اطلاعات نظری قبلی، بررسی شود. هنگامی که مطالعات تجربی طراحی می‌شود، گاهی اوقات به علت تخصیص تصادفی درمان‌ها در افراد، اثرات عوامل مخدوش‌کننده در مطالعه نادیده گرفته می‌شود؛ اما تصادفی کردن یک فرآیند احتمالی است که امکان دارد به‌طور کامل وابستگی‌های بین مواجهه و عوامل خطر خارجی را، مخصوصاً اگر حجم کلی افراد تحت مطالعه کوچک باشد، خنثی نسازد. به‌طور کلی می‌توان گفت مخدوش‌کننده‌ها ممکن است در مطالعات تصادفی نیز وجود داشته باشد هرچند که اثرات آن‌ها در مقایسه با مطالعات غیرتصادفی کوچک‌تر است. اثر مخدوش‌کنندگی در مطالعات تصادفی با حجم نمونه بزرگ و طراحی مناسب قابل‌صرف‌نظر هستند. در مطالعات مورد-شاهدی وابستگی مذکور باید در جامعه مبنا، که موارد را شامل می‌شود، وجود داشته باشد. اگر تعداد اعضای گروه کنترل به‌اندازه کافی زیاد و خطای انتخاب وجود نداشته باشد، برآورد قابل‌قبولی از وابستگی بین عامل مخدوش‌کننده بالقوه و مواجهه در جمعیت مبنا، قابل‌محاسبه است. در غیر این صورت تمام تلاش باید متوجه یافتن درجات وابستگی بین مخدوش‌کننده بالقوه و مواجهه در جامعه مبنا شود که شاهدهای انتخاب‌شده در مطالعه تنها جایگزینی از این جامعه هستند. متأسفانه اطلاعات قابل‌اعتماد تکمیلی درباره وابستگی‌های احتمالی بین عوامل خطر در جامعه مبنا به‌ندرت در دسترس است. به همین علت در مطالعات مورد-شاهدی معمولاً مجبور هستیم، با استفاده از داده‌های موجود برآوردی از وابستگی بین مواجهه و مخدوش‌کننده ارائه کنیم.

۳- عامل مخدوش‌کننده نباید تحت تأثیر رابطه مواجهه با بیماری قرار گیرد. به‌طور خاص عامل مخدوش‌کننده نمی‌تواند یک حلقه میانی از مسیر سببی بین مواجهه و بیماری باشد. به‌عنوان مثال الگوی زیر را در نظر بگیرید.

بیماری قلبی و عروقی → پرفشاری خون → سیگار کشیدن

آیا پرفشاری خون را می‌توان به‌عنوان یک عامل مخدوش‌کننده در نظر گرفت. به‌طورقطع پرفشاری خون یک عامل خطر برای بیماری قلبی-عروقی است و می‌دانیم که با مواجهه موردنظر (سیگار کشیدن) همبسته است، زیرا بیماری به‌وسیله سیگار کشیدن ایجاد می‌شود؛ اما پرفشاری خون در افرادی که مواجهه (سیگار کشیدن) ندارند نیز یک عامل خطر محسوب می‌شود، زیرا این عارضه به عللی غیر از سیگار کشیدن نیز ایجاد می‌شود. با این حال این عامل را نمی‌توان به‌طور مطلق یک مخدوش‌کننده در نظر گرفت، زیرا پرفشاری خون عارضه‌ای است که حد واسط بین سیگار کشیدن و ابتلا به بیماری قلبی-عروقی واقع شده است. عوامل خطری که به‌عنوان یک عارضه حد واسط در حلقه علت و معلول واقع می‌شوند، مخدوش‌کننده به حساب نمی‌آیند و نباید به‌عنوان یک عامل خطر خارجی در نظر گرفته شود.

این نکته، مقدم بودن عامل مخدوش‌کننده بر مواجهه و بیماری را نشان می‌دهد. محقق باید بررسی کند آیا یک مکانیسم سببی وجود دارد و از طریق عامل مواجهه یا بیماری، دستیابی به عامل مخدوش‌کننده امکان‌پذیر است؟ یا اینکه عاملی که به‌عنوان مخدوش‌کننده خوانده می‌شود، تنها یک حلقه میانی است. اگر عامل موردبحث تنها یک حلقه میانی باشد روش‌های تجزیه و تحلیل مجزایی برای تعیین ماهیت این عمل لازم است. هرچند که با وجود سه شرط بالا بیشتر اوقات قادر به تعریف و تشخیص مخدوش‌کنندگی هستیم؛ اما به دلایل محتوایی و تکنیکی محدود شدن به این سه ویژگی در تشخیص عامل مخدوش‌کننده ممکن است منجر به اشتباه شود. از نظر محتوایی اساس مخدوش‌کنندگی بر مخلوط شدن و یا تحریف شدن اثرات خارجی با اثر عامل موردبررسی در مطالعه استوار است. دو مشخصه اول، نتیجه منطقی این تعریف است. یعنی شرط‌های لازم برای مخدوش‌کننده بودن یک عامل. ویژگی سوم شرایطی را که اثر مخدوش‌کننده بالقوه به‌طور مستقیم نمی‌تواند اعمال شود را مستثنا می‌کند. از نظر تکنیکی این امکان وجود دارد که یک عامل هر سه ویژگی مذکور را دارا باشد، اما با اثر عامل مواجهه مخلوط نشود. این امر در مواقعی رخ می‌دهد که مخدوش‌کننده بالقوه هیچ اثر یا عارضه اضافی بر بیماری در بین گروه مواجهه نداشته باشد و تنها یک وابستگی بین این عامل، مواجهه و اثر آن بر بیماری وجود دارد. به‌عنوان مثال چنین حالتی می‌تواند زمانی اتفاق بیافتد که تعدادی مخدوش‌کننده بالقوه وجود دارد که هر کدام به‌تنهایی بر میزان بروز بیماری تأثیرگذار است؛ اما یکی از این عوامل به‌طور کامل در حضور سایر مخدوش‌کننده‌ها بی‌اثر می‌شود و میزان بروز اضافی در گروه مواجهه مرتبط با سایر این مخدوش‌کننده‌ها است.

۶-۱-۲ مخدوش شدن بر اثر ترکیب اثر مواجهه‌ها

در ساده‌ترین حالت، مخدوش شدن می‌تواند به‌صورت ترکیبی از اثرات سطوح مختلف مواجهه نمایان شود. به‌طور خاص، اثری که ظاهراً به‌عنوان اثر مواجهه مشاهده شده است، می‌تواند اثر یک عامل خارجی باشد که با اثر مواجهه موردنظر ترکیب شده باشد. تحریف استنباط‌ها که به‌وسیله عامل مخدوش‌کننده ایجاد می‌شود می‌تواند بسیار زیاد باشد و ممکن است منجر به برآورد بیش از اندازه یا کمتر از اندازه اثر مواجهه اصلی شود. این موضوع بستگی به جهت وابستگی بین عامل مخدوش‌کننده با مواجهه و بیماری دارد.

۶-۱-۳ مخدوش کننده جانشین

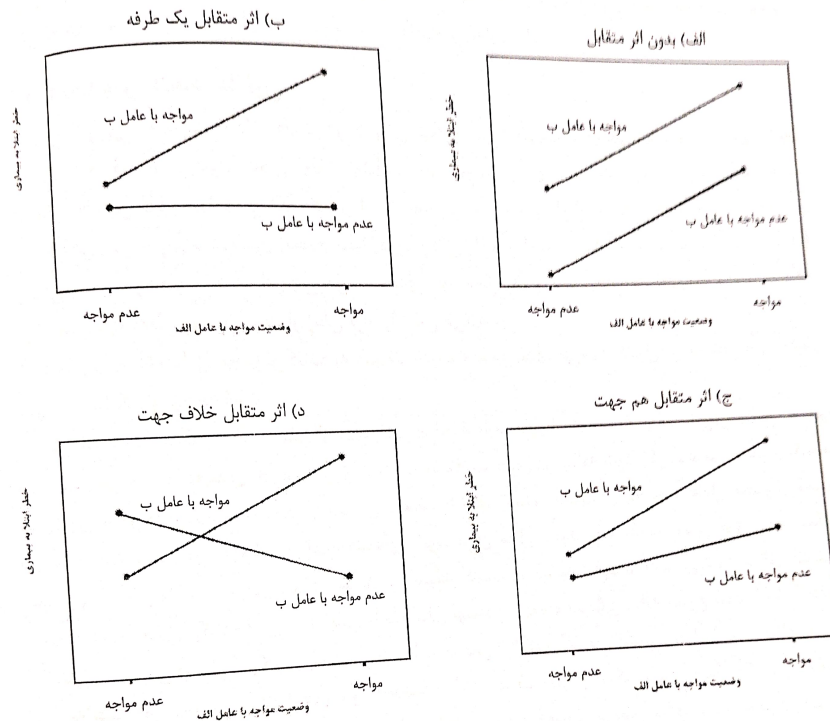
هر عامل خارجی که باعث ایجاد تفاوت در فراوانی بیماری بین گروه‌های مواجهه و بدون مواجهه شود، مخدوش کننده خوانده می‌شود. گاهی اوقات مشاهده این عوامل به‌تنهایی امکان‌پذیر نیست و تنها برخی متغیرهای جانشین قابل مشاهده و اندازه‌گیری است. به‌عنوان مثال، یکی از این عوامل جانشین، سن تقویمی است. افزایش سن با پدیده پیر شدن (تجمع سلول‌های جهش یافته و آسیب یافته که منجر به بیماری می‌شود) وابستگی بسیار نیرومندی دارد؛ اما افزایش سن به‌خودی‌خود عامل پاتولوژیک ایجاد این تغییرات نیست. در واقع سن تقویمی تنها مقدار عمر طی شده از زمان تولد را نشان می‌دهد.

هنگامی که یک عامل مخدوش کننده به‌صورت نادرست طبقه‌بندی می‌شود کنترل آن در تجزیه و تحلیل نتایج بسیار مشکل و پیچیده است.

سو طبقه‌بندی هنگامی که مستقل از طبقه‌بندی مواجهه یا بیماری است معمولاً باعث تورش نتایج به نفع فرضیه صفر می‌شود. طبقه‌بندی نادرست می‌تواند باعث مشکلات جدی در مطالعه شود. اگر اثر عامل مخدوش کننده بسیار قوی و در مقابل رابطه مواجهه و بیماری ضعیف یا صفر باشد، طبقه‌بندی نادرست عامل مخدوش کننده می‌تواند باعث بروز نتایج فوق‌العاده گمراه‌کننده‌ای شود. برای مثال وجود یک رابطه سببی قوی بین سیگار کشیدن و سرطان مثانه به همراه وابستگی شدید بین سیگار کشیدن و نوشیدن قهوه، باعث می‌شود سیگار کشیدن به‌عنوان یک مخدوش کننده قوی در ارزیابی رابطه میان نوشیدن قهوه و سرطان مثانه مطرح شود؛ اما از آنجایی که کنترل مخدوش‌کنندگی سیگار، وابسته به اطلاعات صحیح از الگوی سیگار کشیدن افراد تحت مطالعه است. این اطلاعات نمی‌تواند عاری از خطای سو طبقه‌بندی باشد. صرف‌نظر از اینکه الگوی سیگار کشیدن چگونه اندازه‌گیری می‌شود، می‌توان نتیجه گرفت در ارزیابی رابطه نوشیدن قهوه و سرطان مثانه اثرات مخدوش‌کنندگی در تحلیل نتایج به‌طور کامل قابل کنترل نیست [۶] و [۷].

۶-۲ اثر متقابل

هنگامی که اثر یک عامل خطر مرتبط با بیماری در سطوح مختلف یک عامل خطر دوم متفاوت باشد، بین دو عامل خطر مورد بررسی، اثر متقابل وجود دارد [۹ و ۸]. به همین ترتیب هنگامی که اثر متقابل بین دو عامل خطر در سطوح مختلف عامل خطر دیگر وجود نداشته باشد، اثر هر کدام از عوامل خطر یکسان (همگن) می‌شود. یکی از مثال‌های معروف در این زمینه، اثر سیگار کشیدن بر شانس وقوع سکته قلبی و میزان مرگ‌ومیر در شش‌ماهی اول پس از وقوع اولین سکته قلبی است. از آنجا که سیگار کشیدن یکی از عوامل شناخته شده در وقوع اولین سکته قلبی است، این موضوع باعث می‌شود که بین سیگار کشیدن و سابقه وقوع سکته قلبی (دارد/ ندارد) اثر متقابل وجود داشته باشد.



نمودار ۶-۶: حالت‌های مختلف وجود اثر متقابل بین عوامل خطر الف و ب

برای ساده‌تر شدن موضوع حاضر تصور می‌کنیم عوامل خطر مورد مطالعه تنها دو سطح داشته باشند. این دو سطح را با نام‌های در معرض خطر و آن‌ها که در معرض خطر نیستند مشخص می‌کنیم. نمودار ۶-۶ حالت‌های مختلف اثر متقابل و حالتی که اثر متقابل وجود ندارد را نمایش می‌دهد. امکان بروز سه نوع مختلف اثر متقابل وجود دارد: خلاف جهت، هم‌جهت و یک‌طرفه.

همان‌طور که ملاحظه می‌گردد نداشتن اثر متقابل باعث ایجاد خطوط موازی در نمودار اثرهای متقابل می‌شود؛ اما در عمل مشاهده چنین خطوطی به ندرت اتفاق می‌افتد و تغییرات نمونه‌ای در شانس‌های محاسبه شده باعث بروز تغییرات جزئی و یا حتی تغییرات اساسی می‌شود. در این صورت این تغییرات باعث متقاطع شدن خطوط نمودار می‌شود درحالی‌که واقعاً اثر متقابل وجود ندارد می‌شود. به همین علت، آزمودن داده‌ها برای تشخیص وجود اثر متقابل بسیار منطقی به نظر می‌رسد.

منابع

1. Woodward M. Epidemiology: Study design and data analysis. Chapman & Hall: London, 1999.
2. Schlesselman JJ. Case-control studies: Design, conduct, analysis. New York: Oxford University Press, 1982.
3. Kahn HA, Sempos CT. Statistical methods in epidemiology. Oxford University Press: New York, 1989.
4. Rothman KJ, Sander G. Modern epidemiology (2nd ed.). Lippincott Williams & Wilkins: Philadelphia, 1998.
5. Szklo M, Nieto FJ. Epidemiology. Beyond the basics. Aspen, Maryland, 2000.
6. Sackett DL. Bias in analytic research. J Chronic Dis, 1979. 32:51-63.
7. Flegul KM, Keyl PM, Nieto FJ. Differential misclassification arising from nondifferential errors in exposure measurement. Am J Epidemiol, 1991. 134:1233-44.
۸. نوح‌جاه ص، سوری ح. بررسی رابطه مصرف قرص‌های ترکیبی با دوقلو زایی و حاملگی بعدی. مجله علمی پژوهشی فیض. شماره ۳۴، تابستان ۱۳۸۴: ۵۷-۵۳
9. Soori H. Children indoor and outdoor play pattern in Ahwaz city: implications for injury prevention. Eastern Meditr Health J. 2006, 12(3/4): 19-28.

روش‌های مهار مخدوش‌کنندگی در مطالعات پزشکی



در مطالعات پزشکی برای مهار اثر مخدوش‌کنندگی روش‌های متنوعی وجود دارد که می‌توان آن‌ها را به‌تنهایی یا با هم به کار برد. به‌طور کلی روش‌های کنترل مخدوش‌کنندگی به دو رده روش‌های مهار هنگام طراحی و روش‌های مهار هنگام تجزیه و تحلیل داده‌ها تقسیم می‌شود. تصادفی کردن و محدود کردن در روش کنترل و یا حذف مخدوش شدن هنگام طراحی مطالعه هستند. همسان کردن روشی است که هم در مرحله طراحی مطالعه و هم در مرحله تجزیه و تحلیل به کار می‌رود. طبقه‌بندی و تجزیه و تحلیل چند متغیره روش‌های کنترل مخدوش‌کنندگی هنگام تجزیه و تحلیل داده‌ها هستند.

۲-۱ روش‌های مهار مخدوش‌کنندگی هنگام طراحی مطالعه

برای مهار مخدوش شدن هنگام طراحی مطالعه سه روش، تصادفی کردن، محدود کردن و همسان کردن قابل استفاده است. تصادفی کردن فقط در مطالعات مداخله‌ای به کار گرفته می‌شود، ولی محدود کردن و همسان کردن در طراحی همه مطالعه‌های تحلیلی قابل استفاده است.

۲-۱-۱ تصادفی کردن

تصادفی کردن موجه‌ترین روش برای تعدیل مخدوش‌کنندگی است. قدرت منحصربه‌فرد تصادفی کردن به علت پخش شدن تقریباً یکسان عوامل خطر و مواجهه‌های مختلف در گروه‌ها تحت بررسی است. تصادفی شدن در صورتی که حجم مطالعه به اندازه کافی بزرگ باشد، توزیع یکسان داده‌های تمام متغیرهای مخدوش‌کننده متداول شناخته شده و حتی عوامل خطر ناشناخته را در گروه‌های مقایسه تضمین می‌کند. در نمونه‌های کوچک‌تر و یا حتی در نمونه‌های بزرگ که شانس باعث نامتعادل شدن توزیع عوامل خطر در گروه‌ها می‌شود، امکان مهار این عوامل با به‌کارگیری روش‌های تعدیل مناسب در مرحله تجزیه و تحلیل داده‌ها، وجود دارد.

۲-۱-۲ محدود کردن

همان‌طور که قبلاً اشاره شد در صورت یکسانی عامل بالقوه مخدوش‌کننده در گروه‌های مقایسه، اثر مخدوش شدن خنثی می‌شود. یکی از روش‌های مهار مخدوش‌کنندگی نیز بر همین اساس، محدود کردن معیارهای ورود

افراد در مطالعه است تا افراد مورد مطالعه تنها در یک یا چند رده خاص از عامل مخدوش کننده بالقوه قرار بگیرند. به عنوان مثال اگر نژاد و جنس دو عامل بالقوه مخدوش کننده باشند.

۷-۱-۳ همسان سازی

در مطالعات مورد-شاهدی همسان سازی به انتخاب یکسان موردها و شاهد‌ها از لحاظ برخی از متغیرهای مخدوش کننده گفته می‌شود. همسان کردن و محدود ساختن دو راهکار متداول در کنترل مخدوش کننده‌ها هستند که در مرحله نمونه‌گیری اعمال می‌شوند. در این روش‌ها تنها موردها و شاهد‌هایی با یکدیگر مقایسه می‌شوند که از لحاظ مخدوش کننده‌های خاصی همانند باشند. همسان سازی به علت اینکه امکان حضور نمونه‌هایی از تمام سطوح متغیرهای مخدوش کننده در مطالعه وجود دارد، حالت تعمیم یافته‌ای از روش محدود کردن به حساب می‌آید. همسان کردن معمولاً برای گروه‌های دو عضوی (زوج) انجام می‌گیرد. برای مثال در بررسی میزان مصرف قهوه و وقوع سکته قلبی، هر کدام از موردها (بیماران دارای عارضه سکته قلبی) می‌توانند با یک یا چند نفر از شاهد‌ها از لحاظ میزان مصرف سیگار همسان شود. (مثلاً سیگار نمی‌کشد، ۱۰ نخ سیگار در روز و بیشتر از ۱۰ نخ سیگار در روز) در این صورت مصرف قهوه در هر کدام از موردها، تنها با شاهد (های) همسان شده مقایسه می‌شود.

روش دیگر همسان سازی، همسان سازی به صورت گروهی است. در این روش برای هر سطح از متغیر مخدوش کننده (سیگار کشیدن در مثال قبل) تعدادی از موردها و شاهد‌های همسان برای مقایسه وجود دارند. در مثال قبل اگر نسبت موردها به شاهد‌ها ۲ به ۱ باشد و ۱۰ مورد بیش از ۲۰ نخ سیگار در روز مصرف کنند، لازم است ۴۰ نفر شاهد با همان مقدار مصرف سیگار در روز انتخاب شوند تا با گروه مورد ۲۰ نفری همسان باشند. با توجه به شیوه انجام دو روش همسان سازی فردی و گروهی، مشخص است که همسان سازی گروهی ساده‌تر از همسان سازی فردی است؛ اما شباهت افراد در گروه‌های مقایسه‌ای در همسان سازی فردی بیشتر است. نکته دیگر این است که همسان سازی فردی به روش‌های آماری خاص نیاز دارد اما تحلیل نتایج در روش همسان سازی گروهی تفاوتی با روش‌های معمول ندارد.

تعداد متغیرهایی که برای همسان شدن استفاده می‌شوند باید کمترین تعداد لازم باشند. زیاد شدن متغیرهای همسان شده باعث سخت‌تر شدن معیارها برای انتخاب افراد شاهد مناسب می‌شود و مطالعه در عمل ناکارآمد می‌شود.

کاربرد عمده همسان سازی در مطالعه مورد-شاهدی است اما در مطالعات دیگر هم کاربرد دارد.

مزایا و معایب همسان سازی

برخی مزایای همسان سازی عبارت‌اند از:

- ♦ همسان سازی یک روش کارا در جلوگیری از مخدوش شدن به وسیله عوامل زمینه‌ای نظیر سن و جنس است این عوامل دارای اثر قوی و تعیین کننده‌ای در بسیاری از بیماری‌ها هستند. معمولاً همسان سازی

برای عواملی صورت می‌گیرد که اثر متقابل با عوامل دیگر نداشته و در الگوی سببی نقش واسطه‌ای ایفا نمی‌کنند.

- ♦ همسان‌سازی در کنترل مخدوش‌کننده‌هایی که کمیت آن‌ها قابل اندازه‌گیری و کنترل به‌وسیله سایر روش‌ها نیست، مناسب است. برای مثال همسان کردن فرزندان یک خانواده و یا در حالت ایده‌آل دو قلوها می‌تواند باعث کنترل تمام عوامل ژنتیکی و محیطی غیرقابل اندازه‌گیری شود.
 - ♦ همسان‌سازی به علت متعادل کردن تعداد موردها و شاهد‌ها در هر سطح عامل مخدوش‌کننده باعث افزایش دقت مقایسه‌ها (افزایش توان) می‌شود. این موضوع هنگامی که موردها نادر هستند یا دسترسی به آن‌ها با محدودیت همراه است بسیار مهم است.
 - ♦ همسان‌سازی می‌تواند به‌عنوان یک روش برای انجام نمونه‌گیری در دسترس به کار گرفته شود. برای مثال در مطالعه‌ای درباره سندرم شوک ناشی از مسمومیت، از بیماران مورد مطالعه خواسته شد برخی دوستان خود را به‌عنوان گروه کنترل معرفی کنند اما این نوع نمونه‌گیری خطر بیش همسان‌سازی را افزایش می‌دهد [۱].
- برخی معایب این روش عبارت‌اند از:
- ♦ در مطالعاتی که از همسان‌سازی استفاده می‌کنند امکان بررسی اثر متغیرهای همسان شده وجود ندارد. برای مثال اگر سن افراد در دو گروه مورد مقایسه همسان شود، دیگر امکان بررسی اثر سن بر پیامد مورد نظر وجود ندارد. به‌بیان‌دیگر متغیرهایی باید برای همسان‌سازی انتخاب شوند که اثر مخدوش‌کنندگی آنان به‌خوبی شناخته شده باشد.
 - ♦ گاهی اوقات همسان‌سازی به زمان و هزینه اضافی برای پیدا کردن گروه‌های همسان هر مورد، نیاز دارد. برای مثال در مطالعات مورد-شاهدی هر قدر معیارهای همسان‌سازی بیشتر باشد امکان دسترسی به منابع بیشتری از شاهد‌ها برای پیدا کردن افراد مناسب کمتر است و موردهایی که شاهد‌های مناسب برای آن‌ها پیدا نشود از مطالعه حذف می‌شوند.
 - ♦ از آنجاکه همسان‌سازی یک راهکار نمونه‌گیری محسوب می‌شود تصمیم درباره همسان‌سازی باید در آغاز مطالعه اتخاذ شود. این موضوع تحلیل ابتدایی میزان اثر مخدوش‌کننده‌هایی که همسان شدن برای آن‌ها انجام شده را الزامی می‌کند.
 - ♦ اگر متغیر همسان شده یک عامل واسطه‌ای در زنجیره سببی باشد، مشکلات جدی در استنباط‌ها به وجود می‌آید. برای مثال اگر تمایل به بررسی اثرات مصرف الکل بر سکت قلبی داشته باشیم و مطالعه را بر اساس همسان‌سازی مقدار سرم HDL انجام دهیم، از میزان اثر واسطه‌ای الکل در افزایش HDL بی‌اطلاع خواهیم بود.
 - ♦ تجزیه و تحلیل داده‌های همسان شده شیوه‌های آماری خاص خود را می‌طلبد و استفاده از روش‌های معمولی آماری برای داده‌های زوج شده نادرست است. این روش‌ها بسیار تخصصی هستند و در بسیاری از نرم‌افزارهای آماری وجود ندارند.
 - ♦ یکی دیگر از معایب همسان‌سازی، احتمال بیش همسان‌سازی است. این موضوع زمانی اتفاق می‌افتد که

همسان‌سازی برای عواملی صورت گیرد که واقعاً مخدوش‌کننده نیستند. بیش همسان‌سازی باعث کاهش توان آماری مطالعات می‌شود. در نتیجه یک رابطه حقیقی به‌سختی از لحاظ آماری معنی‌دار می‌شود. مثلاً در مورد مطالعه سندرم شوک ناشی از مسمومیت استفاده از همسان‌سازی ممکن است باعث همسان شدن ناخواسته عواملی نظیر منطقه مسکونی موردها و شاهد‌ها، مارک محصولات غذایی مصرف‌شده و عوامل دیگر نظیر این‌ها شود. با این حال خاطرنشان کردن این نکته مهم است که بیش همسان‌سازی سبب تورش برآورد خطر نسبی نمی‌شود (البته اگر از شیوه‌های ویژه تحلیل داده‌های همسان شده استفاده شود) بلکه باعث کاهش سطح معنی‌داری می‌شود. به همین علت هنگامی که یک رابطه معنی‌دار از یک مطالعه همسان شده به دست می‌آید، بیش همسان‌سازی نتایج را تحت تأثیر قرار نمی‌دهد [۱].

۲-۷ روش‌های مهار مخدوش‌کنندگی در هنگام تجزیه و تحلیل داده‌ها

در مطالعات تحلیلی هدف ارزیابی رابطه بین بیماری با عواملی نظیر مواجهه‌های محیطی، متغیرهای دموگرافیک، شاخص‌های بهداشتی، صفات ژنتیکی و غیره است. حتی اگر هدف مطالعه بررسی یک رابطه ساده بین بیماری و یک مواجهه باشد عوامل تأثیرگذار دیگری وجود دارند که این رابطه را تحت تأثیر قرار می‌دهند. طبقه‌بندی و تحلیل چند متغیره (مدل‌سازی) روش‌های آماری هستند که برای کنترل اثر این عوامل تأثیرگذار (مخدوش‌کننده)، محاسبه اثر مواجهه پس از تعدیل مخدوش‌کننده‌ها و پیدا کردن الگوهای وابستگی بین بیماری و متغیرهای پیش‌بینی‌کننده کاربرد فراوان دارند.

۲-۷-۱ روش طبقه‌بندی

ساده‌ترین روش برای تحلیل مطالعاتی که دارای منابع مخدوش‌کننده بالقوه هستند طبقه‌بندی آنها است. طبقه‌بندی شامل روش‌های ساده و قابل‌فهمی برای ارزیابی وجود و میزان تأثیر عامل مخدوش‌کننده است. جدول ۱-۷ یک مثال فرضی از مطالعه‌ای مورد-شاهدی را نشان می‌دهد که به بررسی رابطه جنس و آلودگی به بیماری مالاریا می‌پردازد.

همان‌طور که در این جدول ملاحظه می‌شود در تحلیل خام، خطر ابتلای مردها به مالاریا ($OR = 1/71$) بیشتر از زن‌ها است. اگر این رابطه یک رابطه سببی باشد، وابستگی مشاهده شده نباید تحت تأثیر عوامل دیگری که با جنس و بیماری در ارتباط هستند قرار گیرد. یکی از این عوامل شغل است. افرادی که دائماً در محیط‌های باز بسر می‌برند (مانند کشاورزان و کارگران) احتمال مواجهه بیشتری با پشه ناقل مالاریا (آنوفل) دارند. بنابراین اگر نسبت مردهایی که در محیط‌های باز کار می‌کنند از زن‌ها بیشتر باشد رابطه به‌دست‌آمده به‌وسیله نوع شغل قابل توجیه است. بر اساس جدول ۱-۷ برآوردهای طبقه‌بندی‌شده از برآوردهای خام متفاوت هستند. در واقع نسبت‌های خطر طبقه‌بندی‌شده در هر رده بسیار نزدیک به ۱ هستند. به کمک طبقه‌بندی نتایج، بر اساس مخدوش‌کننده‌های بالقوه امکان کنترل اثر این عوامل فراهم می‌شود. در مثال بالا امکان محاسبه وابستگی بین جنس (مواجهه) و مالاریا (بیماری)، به‌طور جداگانه برای دو رده شغلی وجود دارد. نسبت‌های خطر محاسبه‌شده در هر رده تقریباً مساوی و متفاوت از برآورد خام هستند، در نتیجه فرضیه مخدوش‌کنندگی رده شغلی برای رابطه

بین جنس و مالاریا تقویت می‌شود. همچنین با توجه به نزدیک بودن نسبت‌های خطر به دست‌آمده از دو رده شغلی این فرضیه که رابطه‌ای حقیقی بین جنس و مالاریا وجود ندارد، تقویت می‌شود [۲].

جدول ۷-۱: تحلیل بر مبنای روش طبقه‌بندی برای ارزیابی رابطه جنس و بیماری مالاریا

- تحلیل خام

مجموع	بیمار	سالم	
۱۵۶	۸۸	۶۸	مرد
۱۴۴	۶۲	۸۲	زن
۳۰۰	۱۵۰	۱۵۰	مجموع

$$۱/۷۱ = \text{نسبت خطر}$$

- تحلیل طبقه‌بندی شده بر اساس شغل

افرادی که به خاطر مسائل شغلی اکثر مواقع در فضای باز هستند.

مجموع	بیمار	سالم	
۶۸	۵۳	۱۵	مرد
۱۳	۱۰	۳	زن
۸۱	۶۳	۱۸	مجموع

$$۱/۰۶ = \text{نسبت خطر}$$

افرادی که به خاطر مسائل شغلی اکثر مواقع در فضای سرپوشیده هستند.

مجموع	بیمار	سالم	
۸۸	۳۵	۵۳	مرد
۱۳۱	۵۲	۷۹	زن
۲۱۹	۸۷	۱۳۲	مجموع

$$۱/۰۰ = \text{نسبت خطر}$$

به کمک طبقه‌بندی قادر به تشخیص اثرهای متقابل نیز هستیم. در مثال قبل نسبت‌های خطر مربوط به دوطبقه بسیار نزدیک به هم هستند. محاسبه نسبت شانس تعدیل شده بر اساس رده شغلی امکان‌پذیر است. نحوه محاسبه این نسبت شانس که به روش میانگین موزن معروف است در همین فصل توضیح داده می‌شود.

جدول ۷-۲ مثال دیگری از مطالعه مورد-شاهدی را نشان می‌دهد که در آن به بررسی تأثیر قرص‌های خوراکی ضدبارداری (OCP) در افزایش خطر سکته قلبی پرداخته شده است.

جدول ۷-۲: تحلیل طبقه‌بندی رابطه قرص‌های ضدبارداری خوراکی (OCP) و سکته قلبی در زن‌ها.

- تحلیل خام:

سکته کرده	سکته نکرده	
۲۹	۱۳۵	OCP مصرف کرده
۲۰۵	۱۶۰۷	OCP مصرف نکرده
۲۳۴	۱۷۴۲	

$$۱/۷ = \text{نسبت خطر}$$

- تحلیل طبقه‌بندی شده بر اساس سن:

♦ سن ۲۵-۲۹

سکته کرده	سکته نکرده	
۴	۶۲	OCP مصرف کرده
۲	۲۲۴	OCP مصرف نکرده

$$۷/۲ = \text{نسبت خطر}$$

♦ سن ۳۰-۳۴

سکته کرده	سکته نکرده	
۹	۳۳	OCP مصرف کرده
۱۲	۳۹۰	OCP مصرف نکرده

$$۸/۹ = \text{نسبت خطر}$$

♦ سن ۳۵-۳۹

سکته کرده	سکته نکرده	
۴	۲۶	OCP مصرف کرده
۳۳	۳۳۰	OCP مصرف نکرده

$$۱/۵ = \text{نسبت خطر}$$

♦ سن ۴۰-۴۴

سکته کرده	سکته نکرده	
۶	۹	OCP مصرف کرده
۶۵	۳۶۲	OCP مصرف نکرده

$$۳/۷ = \text{نسبت خطر}$$

♦ سن ۴۵-۴۹

سکته کرده	سکته نکرده	
۶	۵	OCP مصرف کرده
۹۳	۳۰۱	OCP مصرف نکرده

$$۳/۹ = \text{نسبت خطر}$$

همان‌طور که در قسمت اول جدول نشان داده شده، خطر سکته قلبی در زنان مصرف‌کننده OCP تقریباً ۷۰٪ بیشتر از زنانی است که OCP مصرف نمی‌کنند. با این حال احتمال مخدوش شدن نتایج به‌وسیله سن مورد توجه محققان قرار گرفت. سن یکی از عوامل شناخته‌شده‌ای است که به‌طور مستقیم خطر سکته قلبی را افزایش می‌دهد و رابطه‌ای معکوس با مواجهه مورد بررسی در این مطالعه دارد [۳].

تحلیل طبقه‌بندی‌شده نتایج برحسب سن در قسمت دوم جدول ۷-۲ آورده شده است به‌جز رده سنی ۳۹-۳۵ سال در تمام رده‌ها نسبت‌های خطر، بسیار بزرگ‌تر از برآورد خام نسبت خطر است. این موضوع فرضیه مخدوش‌کننده بودن سن با اثر معکوس را تقویت می‌کند.

۷-۲-۲ پیش‌فرض‌ها

تحلیل به روش طبقه‌بندی تقریباً بدون پیش‌فرض است. تنها شرط این روش معنی‌دار بودن رده‌ها از لحاظ علمی و منطقی است. این موضوع مترادف همگن بودن داده‌ها برحسب عامل مخدوش‌کننده است. در جدول ۷-۱ فرض یکنواخت بودن وابستگی جنس و بیماری مالاریا در دو رده شغلی مورد پذیرش واقع شده است و در جدول ۷-۲ فرض همگن بودن وابستگی بین مصرف قرص OCP و سکته قلبی رده‌های سنی پذیرفته شده است. اگر این فرض در هر کدام از دو مثال قلبی مناسب نباشد رده‌های دقیق‌تری (مثلاً تفکیک دقیق‌تر و جزئی‌تر شغل‌ها و یا بازه‌های سنی کوتاه‌تر) برای تحلیل به روش طبقه‌بندی مورد نیاز است.

تعدیل شاخص‌ها نیاز به مفروضات بیشتری دارد. به‌طور کلی تمام روش‌های تعدیل شاخص‌ها بر اساس پذیرفتن مدل‌های آماری، که وابستگی بین متغیرهای تحت بررسی را خلاصه می‌کند، استوار است. بعضی مواقع مدل‌های آماری مفروض نظیر مدل‌هایی که تعدیل کردن، بر اساس روش طبقه‌بندی انجام می‌شود (مانند روش مانند - هانزل) ساده هستند. در این روش میانگین موزون نسبت‌های شانس اختصاصی هر رده به‌عنوان خلاصه‌ای از برآورد نسبت شانس کلی محاسبه می‌شود. همان‌طور که قبلاً اشاره شد همگن بودن داده‌ها (عدم وجود اثر متقابل) شرط اصلی این نوع تحلیل است. این روش‌های ساده هنگامی که تعداد مخدوش‌کننده‌ها کم است و طبقه‌بندی کردن متغیرها باعث از بین رفتن اطلاعات نمی‌شود بسیار کارا هستند. در حالت‌های پیچیده‌تر روش‌های تعدیل چند متغیره بر اساس مدل‌های رگرسیونی انجام می‌شوند. در این روش‌ها برای بررسی روابط، عوامل متعدد به‌طور هم‌زمان تعدیل می‌شوند. در این مدل‌ها هنگامی که اثر یک عامل بر بیماری بررسی می‌شود به‌طور هم‌زمان تمام عوامل مخدوش‌کننده دیگر تحت کنترل قرار دارند. در این مدل‌ها امکان بررسی تأثیر مخدوش‌کننده‌ها در حالت پیوسته (بدون طبقه‌بندی) نیز وجود دارد.

بسیاری از شاخص‌های بهداشتی نظیر میزان بروز بیماری و مرگ با متغیرهای زمینه‌ای مانند سن و جنس ارتباط دارند. برای مثال جوامع پیرتر دارای میزان مرگ خام بالاتری از جوامعی با ترکیب جمعیتی جوان‌تر هستند. در این موارد استفاده از میزان‌های خام گمراه‌کننده است. این موضوع برای شاخص‌های دیگر و مخدوش‌کننده بالقوه مرتبط با آن‌ها صدق می‌کند. این قسمت به معرفی روش‌های مبتنی بر طبقه‌بندی برای مهار متغیرهای مخدوش‌کننده اصلی و قابل‌مقایسه ساختن گروه‌های مورد مقایسه با یکدیگر اختصاص دارد.

۷-۲-۳ استاندارد کردن به روش مستقیم

استاندارد کردن مستقیم روشی است که معمولاً برای تعدیل سن در مقایسه میزان‌های مرگ و بیماری بین نواحی مختلف یا دوره‌های زمانی متفاوت به کار می‌رود؛ اما کاربرد متداول به معنی این نیست که این روش تنها برای تعدیل سن استفاده می‌شود. استفاده آسان مدل‌های آماری به‌وسیله نرم‌افزار، کاربرد این روش را در تحلیل‌ها محدود ساخته است، اما به علت منطق ساده و قابل‌فهم این روش هنوز هم به‌عنوان یکی از روش‌های اصلی تحلیل داده‌ها محسوب می‌شود.

جدول ۷-۳ نحوه عملکرد این روش را برای مقایسه میزان بروز دو گروه الف و ب (مثلاً مواجهه یافته و بدون مواجهه با مخدوش‌کننده بالقوه موردنظر) نشان داده شده است.

- ♦ برای هر رده مخدوش‌کننده، میزان بروز در دو گروه محاسبه می‌شود (ستون‌های ۴ و ۷).
- ♦ اعضای جمعیت مرجع در هر رده از متغیر مخدوش‌کننده مشخص می‌شود (ستون ۸).
- ♦ تعداد مورد‌های منتظره از جمعیت مرجع با ضرب میزان‌های بروز هر رده از گروه‌های الف و ب در تعداد اعضای جامعه مرجع محاسبه می‌شود (ستون‌های ۹ و ۱۰).
- ♦ نسبت مورد‌های هر گروه از تعداد کل جمعیت مرجع، میزان بروز تعدیل‌یافته یا استاندارد شده هر گروه را معین می‌کند.

میزان استاندارد نشان‌دهنده این موضوع است که اگر دو گروه الف و ب ساختاری کاملاً مشابه با ساختار جمعیت مرجع از لحاظ مخدوش‌کننده موردبررسی داشتند، (مثلاً توزیع سنی یکسان با جمعیت مرجع داشتند) دارای میزان‌های بروز I_A^* و I_B^* می‌شوند.

رابطه‌ای که بر اساس آن میزان‌های استاندارد شده محاسبه می‌شوند نوعی میانگین وزنی از میزان‌های اختصاصی هر گروه است. وزن‌های موردنظر در این روش تعداد افراد جامعه مرجع در هر رده هستند. به کمک میزان‌های استاندارد شده می‌توان خطر قابل انتساب تعدیل‌یافته یا خطر نسبی تعدیل‌یافته را به‌صورت زیر محاسبه کرد:

$$AR = I_A^* - I_B^* \quad \text{خطر قابل انتساب تعدیل یافته}$$

$$RR = \frac{I_A^*}{I_B^*} \quad \text{خطر نسبی تعدیل یافته}$$

جدول ۷-۳: استفاده از روش استاندارد مستقیم برای مقایسه میزان بروز (I) در دو گروه مورد مطالعه

جمعیت مرجع			گروه ب			گروه الف			متغیر مخدوش‌کننده بالقوه
موردهای منتظره با استفاده از بروز گروه ب	موردهای منتظره با استفاده از بروز گروه الف	تعداد	بروز	مورد	تعداد	بروز	مورد	تعداد	(۱)
$(10) = (7) \times (8)$	$(9) = (4) \times (8)$	(۸)	$(7) = (6) / (5)$	(۶)	(۵)	$(4) = (3) / (2)$	(۳)	(۲)	(۱)
$IB_1 \times W_1$	$IA_1 \times W_1$	W_1	IB_1	XB_1	DB_1	IA_1	XA_1	DA_1	رده ۱
$IB_2 \times W_2$	$IA_2 \times W_2$	W_2	IB_2	XB_2	DB_2	IA_2	XA_2	DA_2	رده ۲
$IB_3 \times W_3$	$IA_3 \times W_3$	W_3	IB_3	XB_3	DB_3	IA_3	XA_3	DA_3	رده ۳
-	-	-	-	-	-	-	-	-	-
$IB_K \times W_K$	$IA_K \times W_K$	W_K	IB_K	XB_K	DB_K	IA_K	XA_K	DA_K	رده K

در این جدول DA_i و DB_i تعداد نمونه در رده نام از گروه الف و ب، XA_i و XB_i تعداد وقایع رخ داده در گروه الف و ب و IA_i و IB_i میزان بروز در گروه الف و ب است.

در جدول ۷-۴ استفاده از روش استاندارد کردن مستقیم در محاسبه میزان‌های تعدیل یافته وقوع بیماری گرفتگی عروق کرونر (CHD) نمایش داده شده است. در این مطالعه تمام جمعیت ۶۴-۲۵ ساله مناطق شمالی گلاسکو مورد بررسی قرار گرفتند. در جدول ۷-۴ تعداد موارد بروز CHD به تفکیک سن و رده اجتماعی اقتصادی (کاملاً محروم = ۱ تا کاملاً برخوردار = ۴) نشان داده شده است [۴].

جدول ۷-۴: وقوع بیماری گرفتگی عروق کرونر به تفکیک گروه‌های سنی و رده‌های اجتماعی-اقتصادی در مناطق شمالی گلاسکو و جمعیت مرجع انگلستان در محدوده سنی ۶۴-۲۵

گروه‌های اقتصادی اجتماعی									
جمعیت مرجع	۴		۳		۲		۱		داده‌های سنی
	جمعیت	تعداد وقایع	جمعیت	تعداد وقایع	جمعیت	تعداد وقایع	جمعیت	تعداد وقایع	
۱۲۹۸۲۹۰۰	۴۴۴۰	۰	۴۳۵۱	۰	۴۹۷۲	۰	۴۷۸۴	۰	۲۵-۲۹
۱۰۷۷۲۱۲۰	۳۶۸۵	۱	۳۲۳۲	۰	۴۰۴۵	۰	۴۲۱۰	۰	۳۰-۳۴
۸۰۸۷۹۲۰	۲۹۶۶	۵	۲۴۳۸	۱	۳۰۹۴	۴	۳۳۹۶	۱	۳۵-۳۹
۷۶۱۹۵۰۰	۲۷۶۳	۱۰	۲۲۴۱	۹	۲۶۵۵	۷	۳۲۲۶	۶	۴۰-۴۴
۶۹۲۱۸۶۰	۲۳۸۸	۱۵	۲۳۶۰	۱۷	۲۳۴۳	۱۳	۲۳۹۱	۷	۴۵-۴۹
۶۹۷۵۵۰۴۰	۲۵۶۶	۲۴	۲۷۰۸	۱۹	۲۳۹۴	۱۱	۲۱۵۶	۱۶	۵۰-۵۴
۷۱۹۵۱۴۰	۲۳۸۷	۲۸	۲۹۶۸	۴۳	۲۵۹۷	۲۸	۲۱۸۲	۱۷	۵۵-۵۹
۷۱۳۰۱۶۰	۲۳۸۰	۵۶	۲۸۰۲	۵۳	۲۶۶۷	۴۴	۲۰۵۴	۲۵	۶۰-۶۵
۵۷۹۸۴۶۴۰	۲۳۵۷۵	۱۳۹	۲۳۱۰۰	۱۴۲	۲۴۷۶۷	۱۰۷	۲۴۳۹۹	۷۲	مجموع
	۵/۹۰		۶/۱۵		۴/۳۲		۲/۹۵		میزان (در هزار)

۷-۲-۴ روش استاندارد غیرمستقیم

استاندارد کردن غیرمستقیم هم به‌طور متداول برای حذف اثر سن در مطالعات مربوط به مرگ بیماری‌زایی استفاده می‌شود. در این روش تعداد مورد انتظار برای وقایع تحت بررسی (مثلاً مرگ) در هر رده از مخدوش‌کننده بالقوه (نظیر مواجهه‌های شغلی متفاوت) با استفاده از میزان‌های جامعه مرجع (میزان‌های استاندارد) محاسبه می‌شود.

نسبت مجموع وقایع مورد انتظار در هر گروه برآوردی از خطر نسبی یا نسبت شانس تعدیل شده ارائه می‌دهد. جدول ۷-۵ نحوه محاسبه این برآوردها را نمایش می‌دهد. هنگامی که برای تحلیل داده‌های مرگ از این برآورد استفاده می‌شود، با عنوان نسبت مرگ استاندارد (SMR) شناخته می‌شود. واژه‌های میزان بروز استاندارد (SIR) و نسبت شیوع استاندارد (SPR) نیز متداول هستند. استاندارد کردن غیرمستقیم، هنگامی که تعداد وقایع در هر کدام از رده‌های مواجهه کوچک است و یا اصلاً در برخی رده‌ها هیچ واقعه‌ای اتفاق نیفتاده است، بسیار مناسب است. در این شرایط روش استاندارد مستقیم کارا نیست.

یکی از نکات مهم در استفاده از این روش به کار بردن جمعیت مرجع مناسب است. جمعیت مرجع معمولاً جمعیتی جدا از افراد تحت بررسی است، هرچند استفاده از داده‌های ادغام‌شده به‌عنوان جمعیت مرجع نیز متداول است؛ اما توجه به این موضوع که هنگامی که مطالعات مختلف جمعیت‌های مرجع متفاوتی داشته باشند احتمالاً میزان‌های استاندارد شده حاصله قابل مقایسه نیست، بسیار مهم است.

جدول ۷-۵: استفاده از استاندارد کردن غیرمستقیم برای محاسبه میزان مرگ‌ومیر استاندارد شده با جمعیت مرجع

جمعیت مطالعه الف		جمعیت مرجع		
مخدوش‌کننده بالقوه	تعداد	مرگ‌های مشاهده شده	میزان مرگ‌ومیر	مرگ‌های مورد انتظار هنگامی که جمعیت الف دارای میزان مرگ‌ومیر مشابه جمعیت مرجع باشد
(۱)	(۲)	(۳)	(۴)	(۵) = (۴) × (۲)
رده ۱	nA_1	XA_1	M_1	$M_1 \times nA_1$
رده ۲	nA_2	XA_2	M_2	$M_2 \times nA_2$
رده ۳	nA_3	XA_3	M_3	$M_3 \times nA_3$
-	-	-	-	-
رده K	nA_k	XA_k	M_k	$M_k \times nA_k$
مجموع		$\sum_i XA_i$		$\sum_i [M_i \times nA_i]$

$$SMR = \frac{\text{مرگ‌های مشاهده شده}}{\text{مرگ‌های مورد انتظار}} = \frac{\sum_i XA_i}{\sum_i [M_i \times nA_i]}$$

نسبت مرگ‌ومیر استاندارد

۷-۲-۵ روش مانتل - هانزل برای برآورد شاخص‌های تعدیل یافته

هنگامی که مایل به بیان خلاصه نسبت شانس یا خطر نسبی تعدیل شده برای یک مخدوش‌کننده اسمی باشیم، معمولاً از روش مانتل - هانزل استفاده می‌شود [۵-۷]. در جدول ۶-۷ نمادگذاری‌های روش مانتل هانزل نمایش داده شده است. بر اساس این نمادگذاری نسبت شانس تعدیل شده برای داده‌های اسمی، OR_{MH} به این صورت است:

$$OR_{MH} = \frac{\sum_i \frac{a_i d_i}{N_i}}{\sum_i \frac{b_i c_i}{N_i}}$$

جدول ۶-۷: نمادگذاری‌های به‌کاررفته برای محاسبه نسبت شانس تعدیل‌یافته به روش مانتل-هانزل

- رده نام مخدوش‌کننده

مجموع	موردها	شاهدها	
M_{1i}	a_i	b_i	دارای مواجهه
M_{2i}	c_i	d_i	بدون مواجهه
N_i	n_{1i}	n_{2i}	مجموع

محاسبه OR_{MH} بسیار ساده است. در جدول (۷-۱) شانس‌های خام ابتدا به مالاریا به تفکیک محاسبه شد که نشان‌دهنده خطر ابتلای بالاتر مردها است؛ اما محاسبه شانس‌های طبقه‌بندی‌شده بر اساس شغل بیانگر عدم وجود وابستگی بین میزان ابتلا و جنس بود. استفاده از نسبت شانس تعدیل‌یافته با روش مانتل - هانزل، برآورد خلاصه‌ای از مقدار وابستگی بین جنس و میزان ابتلا را به دست می‌دهد. نحوه محاسبه OR_{MH} به این صورت است:

$$OR_{MH} = \frac{\sum_i \frac{a_i d_i}{N_i}}{\sum_i \frac{b_i c_i}{N_i}} = \frac{\frac{53 \times 3}{81} + \frac{35 + 79}{219}}{\frac{10 \times 15}{81} + \frac{52 \times 53}{219}} = 1/01$$

برای مثال نحوه محاسبه OR_{MH} برای رابطه بین وقوع سکته قلبی و مصرف OCP در جدول (۷-۲) به این صورت محاسبه می‌شود:

$$OR_{MH} = \frac{\sum_i \frac{a_i d_i}{N_i}}{\sum_i \frac{b_i c_i}{N_i}} = \frac{\frac{4 \times 224}{292} + \frac{9 \times 390}{444} + \frac{4 \times 330}{393} + \frac{6 \times 362}{442} + \frac{6 \times 301}{405}}{\frac{2 \times 62}{292} + \frac{12 \times 33}{444} + \frac{33 \times 26}{393} + \frac{65 \times 9}{442} + \frac{93 \times 5}{405}} = 4/0$$

۷-۲-۶ روش مانتل - هانزل برای تحلیل مطالعات همسان شده (روش مک - نماز)

نسبت شانس در مطالعات مورد-شاهدی همسان شده، از تقسیم تعداد زوج‌هایی که در آن‌ها موردها دارای مواجهه و شاهد‌ها بدون مواجهه هستند بر تعداد زوج‌هایی که در آن‌ها موردها بدون مواجهه و شاهد‌ها دارای مواجهه هستند، محاسبه می‌شود. بدین ترتیب در جداول توافقی 2×2 نظیر آنچه در جدول ۷-۷ مشاهده می‌شود نسبت شانس برابر $\frac{b}{c}$ است.

جدول ۷-۷: نحوه نمایش نتایج در مطالعات زوج شده

شاهد‌ها			موردها
بدون مواجهه	دارای مواجهه	دارای مواجهه	
b	a	دارای مواجهه	موردها
d	c	بدون مواجهه	

$$\text{نسبت شانس زوج شده} = \frac{b}{c}$$

۷-۲-۷ محدودیت‌های روش‌های تعدیل مبتنی بر طبقه‌بندی

هرکدام از روش‌هایی که در قسمت‌های قبلی برای کنترل متغیرهای مخدوش‌کننده موردبررسی قرار گرفت می‌تواند برای کنترل چند متغیر مخدوش‌کننده نیز به کار رود. برای این کار باید تمام ترکیب‌های ممکن از مجموعه متغیرهای مخدوش‌کننده را به‌عنوان رده‌های موردنظر تفکیک کرد. جدول ۷-۸ مثال فرضی از این حالت را نمایش می‌دهد. این روش تحلیل برای تعدیل چند متغیره دارای محدودیت‌ها و معایب زیر است.

- ♦ شاخص‌های تعدیل‌یافته محاسبه‌شده تنها برای برآورد رابطه بین مواجهه و پاسخ موردنظر صورت می‌گیرد. برای مثال برای ارزیابی رابطه میان مصرف OCP و سکنه قلبی ممکن است مخدوش‌کننده‌های سن و میزان تحصیلات تعدیل شوند. اگر لازم باشد رابطه میان سطح تحصیلات نیز با سکنه قلبی ارزیابی شود، تمام مراحل باید از ابتدا تکرار شود.
- ♦ این روش‌ها تنها برای تعدیل مخدوش‌کننده‌های اسمی (نظیر جنس و نژاد) کاربرد دارند. هنگامی که با متغیرهای پیوسته سروکار داریم طبقه‌بندی آن‌ها اجتناب‌ناپذیر است، نحوه طبقه‌بندی متغیرها می‌تواند سلیقه‌ای باشد و به نتایج متفاوتی منجر شود.
- ♦ هنگامی که تعداد مخدوش‌کننده‌ها زیاد باشد، تعداد طبقه‌ها زیاد می‌شود و ممکن است داده‌ها طوری پراکنده شوند که نتایج غیرقابل‌اعتماد شود.

جدول ۷-۸: مثال فرضی از نحوه انجام تعدیل مستقیم برای کنترل هم‌زمان جنس، نژاد و میزان تحصیلات

جنس	نژاد	میزان تحصیلات (سال)	طبقه	میزان‌های مربوط به گروه A	میزان‌های مربوط به گروه B	وزن‌های به‌دست‌آمده از جمعیت مرجع
مرد	سیاه	< ۱۲	۱	-	-	-
		≥ ۱۲	۲	-	-	-
	سفید	< ۱۲	۳	-	-	-
		≥ ۱۲	۴	-	-	-
زن	سیاه	< ۱۲	۵	-	-	-
		≥ ۱۲	۶	-	-	-
	سفید	< ۱۲	۷	-	-	-
		≥ ۱۲	۸	-	-	-

۷-۲-۸ مدل رگرسیون برای مهار هم‌زمان مخدوش‌کننده‌ها

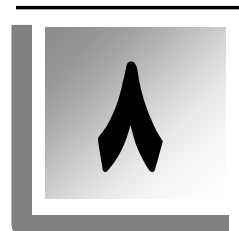
با توجه به محدودیت‌ها و مشکلات ناشی از روش‌های مبتنی بر طبقه‌بندی، روش رگرسیون چندگانه معمولاً کارآمدتر است. در این روش اثر هم‌زمان تمام مواجهه‌ها در حالی که برای مخدوش‌کننده‌های بالقوه تعدیل شده است، به صورت یک مدل ارائه می‌شود. در این روش امکان بررسی متغیرهای پیوسته نیز وجود دارد [۷]. یکی دیگر از مزایای این روش‌ها امکان پیش‌بینی برآوردهای موردنظر، با استفاده از مدل ساخته‌شده برای متغیرها است.

بحث و بررسی بیشتر این مبحث به فصل هشتم که به تحلیل داده‌ها اختصاص دارد، موکول می‌شود.

منابع

1. Rothman KJ. Epidemiology: An introduction. Oxford University Press, New York 2002.
2. Szklo M, Nieto FJ. Epidemiology: Beyond the basics. Aspen, Maryland, 2000.
3. Shapiro S, Slone D, Rosenberg L, et al. Oral-contraceptive use in relation to myocardial infraction. Lancet, 1979. 1:743-7.
4. Woodward M. Epidemiology: Study design and data analysis. Chapman & Hall/CRC; London, 1999.
۵. گروه مؤلفان. روش‌شناسی پژوهش‌های کاربردی در علوم پزشکی، نشر وزارت بهداشت، درمان و آموزش پزشکی، معاونت تحقیقات و فن‌آوری ۱۳۸۳.
6. Bhopal RS. Concept of epidemiology. Oxford University Press, London. 2002.
۷. سوری ح، رحیمی م، محسنی ح. بررسی ارتباط بین استرس‌های شغلی و حوادث ناشی از کار: یک مطالعه مورد شاهدی. مجله تخصصی اپیدمیولوژی. جلد ۱ شماره ۲، زمستان ۱۳۸۵: ۵۸-۵۳.

تحلیل داده‌ها



به‌طور معمول هر تحقیق شامل داده‌های جمع‌آوری شده از متغیرهای مختلف نظیر متغیرهای زمینه‌ای، متغیرهای مخدوش‌کننده، پارامترهای مربوط به پیامد یا رخداد موردنظر، اندازه‌های مربوط به پیگیری افراد تحت مطالعه و بروز عوارض احتمالی است. قبل از تجزیه و تحلیل اطلاعات، طراحی برنامه مدون برای تحلیل، بررسی کیفیت داده‌ها، انجام بررسی‌های مقدماتی نظیر نمودارها، آمارهای توصیفی و جداول توافقی، حذف داده‌های کم‌اهمیت، تجزیه و تحلیل یک متغیره و درنهایت مدل‌سازی (تحلیل هم‌زمان متغیرها) الزامی است. در غیر این صورت اطمینان کافی در کسب نتایج معتبر و تکرارپذیر، وجود ندارد.

۸-۱ طراحی برنامه مدون برای تجزیه و تحلیل

معمولاً در هر پروپوزال تحقیقاتی قسمتی برای توضیح راجع به نحوه تجزیه و تحلیل نتایج وجود دارد. بیان جزئیات تجزیه و تحلیل در این قسمت برای هدایت طرح بسیار مفید و راهگشا است. نکات مهم، بسته به نوع مطالعه متفاوت است. در کارآزمایی‌های تصادفی، به علت ماهیت دقیق این مطالعات، برنامه تجزیه و تحلیل با جزئیات دقیق قابل ذکر است. این موضوع می‌تواند شامل نحوه بررسی پیامدهای اصلی و جانبی، روش‌های آماری مورد استفاده و دستورالعمل‌هایی راجع به انجام تعدیل برای متغیرهای پایه باشد.

در مطالعات مشاهده‌ای که جنبه توصیفی و بیان روابط بین پارامترها مهم است، معمولاً فرضیه‌های اصلی بسیار مفید و جهت بخش است. در قسمت دیگری از برنامه تجزیه و تحلیل باید فهرستی از متغیرهای مستقل، پاسخ و مخدوش‌کننده‌های بالقوه که برای پاسخگویی به سؤالات، جمع‌آوری آن لازم است، تهیه شود.

۸-۲ بررسی کیفیت داده‌ها

بررسی صحت داده‌ها و کنترل کیفیت مجموعه داده‌ها یا بانک‌های اطلاعاتی قبل از انجام تجزیه و تحلیل بسیار مهم است. گام اول بررسی فرم‌های اولیه برای تعیین داده‌های گم شده یا موارد خارج شده از مطالعه است. مرحله بعدی جستجو در داده‌ها برای یافتن داده‌های دورافتاده (پرت) است. در داده‌های پیوسته این موضوع با رسم توزیع پراکندگی داده‌ها و در متغیرهای رسته‌ای با تهیه جداول توافقی و ارزیابی فراوانی‌های مربوط به هر خانه آغاز می‌شود. در مواردی که حجم داده‌ها و تعداد متغیرها وسیع است، اهمیت این موضوع دوچندان می‌شود.

گام بعدی، ارزیابی سازگاری اطلاعات است. برای مثال اگر نوع سرطان و جنس ثبت شده باشند، مشاهده خانم‌هایی با سرطان پروستات ناسازگار است. در مورد متغیرهای پیوسته رسم نمودار پراکنش بسیار سودمند است. برای مثال رسم نمودارهای وزن و قد یا وزن و سن به یافتن موارد ناسازگار کمک می‌کند. برخی دیگر از داده‌های دورافتاده با این روش شناسایی می‌شوند.

خطاهای احتمالی، هر جا که امکان‌پذیر باشد باید به‌وسیله کنترل داده‌های ثبت شده با مقادیر اصلی برطرف شوند. در برخی موارد امکان تصحیح خطاها وجود دارد. در بیشتر موارد لازم است کدهای خاصی برای ثبت موارد خطا، داده‌های گم شده، دورافتاده یا تکمیل‌نشده معرفی شوند. برای موارد گم‌شده روش‌های آماری مخصوصی وجود دارد که استفاده از این روش‌ها نیاز به مشاوره با متخصصان آمار دارد.

۸-۳ آمارهای توصیفی

پس از اتمام مرحله بررسی کیفیت داده‌ها، می‌توان به توصیف اطلاعات پرداخت. بهتر است از رسم نمودارهای فراوانی و پراکنش شروع کنیم. این موضوع علاوه بر کنترل نهایی صحت داده‌ها، باعث فراهم آوردن درک شهودی از آنچه در تحقیق اتفاق افتاده می‌شود. بررسی مجزای زیرگروه‌های موردبررسی (فرضاً بیماران و افراد سالم) در بسیاری از موارد الزامی است.

برای ارزیابی هر سؤال تحقیق، تعیین متغیر پاسخ و متغیرهای مواجهه ضروری است. در مطالعات مشاهده‌ای، مهار مخدوش‌کننده‌ها نقش اساسی در صحت مطالعات ایفا می‌کند. به این منظور لازم است متغیرهای مخدوش‌کننده بالقوه و زمینه‌ای قبل از شروع مطالعه مشخص شوند.

۸-۳-۱ میانگین

میانگین یک مجموعه اعداد، معمولاً میانگین حسابی اعداد است که به‌اختصار میانگین گفته می‌شود. میانگین به طریق جمع اعداد و تقسیم آن‌ها به تعدادشان محاسبه می‌شود:

$$\bar{X} = \frac{\sum x}{n}$$

میانگین: \bar{X}

در این رابطه X نشان‌دهنده مقادیر عددی از متغیر موردنظر و Σ علامت مجموع n عدد با یکدیگر است.

۸-۳-۲ واریانس

در اکثر تحلیل‌های آماری، مناسب‌ترین معیار برای نشان دادن پراکندگی داده‌ها واریانس یا انحراف معیار است. واریانس به‌صورت زیر محاسبه می‌شود.

$$S^2 = \frac{\sum (x - \bar{x})^2}{(n-1)}$$

واریانس: S^2

۳-۳-۸ انحراف معیار

یکی از مشکلات تفسیر واریانس این است که واحد آن مربع واحد اصلی داده‌ها است. برای مثال اگر واحد مشاهدات وزن، گرم باشد واحد واریانس گرم به توان ۲ است. در اکثر موارد مایلیم با جذر گرفتن از واریانس واحد را به مقیاس اصلی برگردانیم این معیار جدید انحراف معیار نام دارد. انحراف معیار به صورت زیر محاسبه می‌شود.

$$S = \sqrt{\frac{\sum (x - \bar{x})^2}{(n-1)}} \quad \text{انحراف معیار (sd)}$$

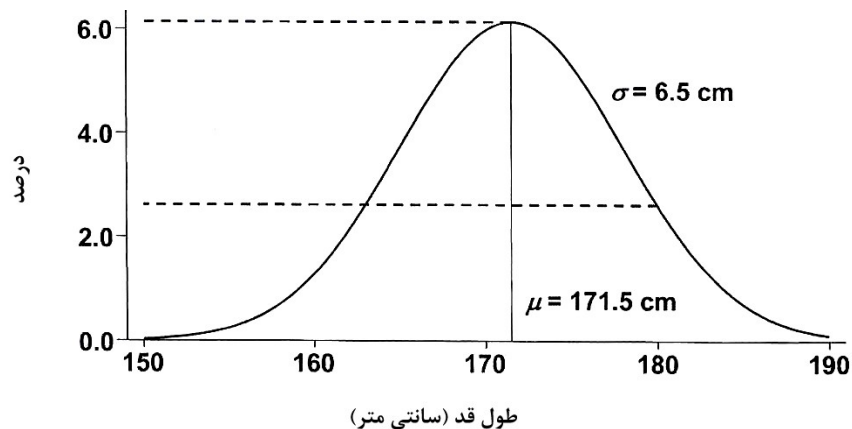
و یا

$$S = \sqrt{\frac{\sum x^2 - \frac{(\sum x)^2}{n}}{(n-1)}} \quad \text{انحراف معیار}$$

۴-۳-۸ توزیع نرمال

در موارد زیادی توزیع فراوانی داده‌ها به صورت توزیعی متقارن و زنگوله‌ای شکل مناسب به نظر می‌رسد. در عمل چنین توزیع تجربی را با توزیع نرمال برابر در نظر می‌گیرند. شکل ۸-۱ تصویری از توزیع نرمال که از اندازه‌گیری قد مردان بالغ حاصل شده است را نمایش می‌دهد.

پیدا کردن توزیع‌های فراوانی که شبیه توزیع نرمال نیستند کار مشکلی نیست. توزیع درآمد، طول عمر و رشد یک تومور مثال‌هایی از این قبیل هستند. در چنین مواردی معمولاً انجام تبدیل (مثلاً تبدیل به وسیله لگاریتم گرفتن از داده‌ها)، به نرمال شدن داده‌ها کمک می‌کند.



شکل ۸-۱: شکل نرمال تقریبی برای توزیع قد مردان بالغ

اهمیت و کاربرد فراوان توزیع نرمال تنها به علت شبیه بودن تقریبی توزیع‌های فراوانی تجربی به شکل توزیع نرمال، بلکه به علت برخی ویژگی‌های مطلوب این توزیع از لحاظ ریاضی است. این توزیع نقش بسیار مهمی در تحلیل آماری ایفا می‌کند. یکی از این ویژگی‌های مهم از قضیه حد مرکزی نتیجه گرفته می‌شود. با توجه به این قضیه، بدون توجه به شکل اولیه توزیع فراوانی تجربی، اگر حجم نمونه به اندازه کافی بزرگ باشد، توزیع فراوانی میانگین نمونه معادل توزیع نرمال می‌شود.

توزیع نرمال به وسیله دانستن میانگین (μ) و انحراف معیار (σ) جامعه مشخص می‌شود. رابطه ریاضی این توزیع به شکل زیر است:

$$Y = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

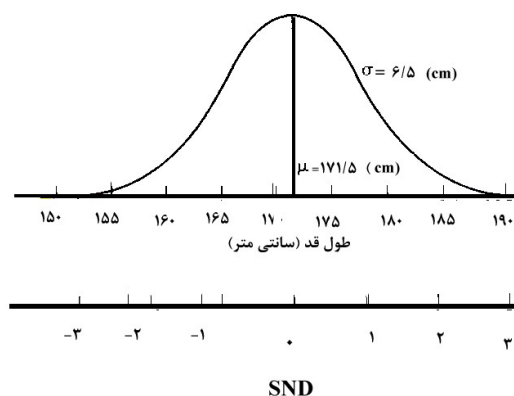
۸-۳-۵ توزیع نرمال استاندارد

اگر توزیع متغیرها نرمال باشد، تغییرات در واحد متغیر باعث عوض شدن توزیع نمی‌شود. بنابراین برای مثال، اگر قد با واحد سانتی‌متر، اینچ یا متر اندازه گرفته شود باز هم توزیع متغیر قد نرمال است. تغییر در میانگین باعث جابجایی منحنی نرمال در محور افقی شده و تغییر انحراف معیار عرض منحنی را تغییر می‌دهد.

با توجه به این ویژگی می‌توان تمام توزیع‌های نرمال را به توزیع نرمال استاندارد که دارای میانگین صفر و انحراف معیار ۱ باشد، تبدیل کرد. این موضوع با تفریق داده‌ها از میانگین نمونه‌ای و تقسیم آن‌ها بر انحراف معیار نمونه حاصل می‌شود. این رابطه را می‌توان به صورت

$$SND: Z = \frac{x - \mu}{\sigma}$$

نمایش داد. در این رابطه x متغیر اصلی با میانگین (μ) و انحراف معیار (σ) و Z متغیر استاندارد شده (SND) است. این موضوع برای توزیع نرمال شکل ۸-۱ انجام شده است. توزیع نرمال استاندارد حاصل در شکل ۸-۲ نمایش داده شده است.



شکل ۸-۲: رابطه میان توزیع نرمال در مقیاس اصلی و توزیع نرمال استاندارد شده

۸-۳-۶ فاصله اطمینان

شاخص‌هایی از قبیل میانگین (μ) و انحراف معیار (σ) پارامترهای جامعه تحت بررسی هستند. به‌طور کلی هر شاخص از جامعه را یک پارامتر می‌نامند. پارامترها مقدار ثابت اما نامعلوم دارند. آنچه در عمل در پژوهش‌های کاربردی رخ می‌دهد، انتخاب نمونه‌ای مناسب از جامعه و برآورد پارامترهای موردنظر است. به این ترتیب \bar{X} و S برآوردهای نقطه‌ای از پارامترهای جامعه یعنی μ و σ هستند. نوع دیگری از برآوردها، برآوردهای فاصله‌ای هستند. برآوردهای فاصله‌ای روشی برای نشان دادن میزان دقت موجود در برآورد پارامترها است. به این منظور ابتدا باید خطای معیار هر برآورد محاسبه شود. رابطه کلی برای محاسبه فاصله اطمینان به صورت زیر است:

$$\text{خطای معیار برآورد} \times (\text{ضریب اطمینان}) \pm (\text{برآورد نقطه‌ای}) = \text{فاصله اطمینان}$$

برآورد فاصله‌ای برای پارامتر میانگین جامعه به صورت زیر محاسبه می‌شود:

$$\bar{X} \pm t_{n-1} \frac{S}{\sqrt{n}}$$

در این رابطه، t_{n-1} ضریب اطمینان موردنظر در سطح ثابت و از پیش تعیین شده (مثلاً ۹۵٪)، از توزیع t -استیودنت با درجه آزادی $n-1$ است. توزیع t را می‌توان جایگزینی برای توزیع نرمال هنگامی که حجم نمونه کم است در نظر گرفت.

۸-۳-۷ جداول توافقی و اندازه‌گیری اثر مواجهه

مفیدترین روش برای خلاصه کردن رابطه بین دو متغیر رسته‌ای تنظیم جدولی نظیر جدول ۸-۱ است. به این جدول، جدول توافقی 2×2 گفته می‌شود.

جدول ۸-۱: جدول توافقی 2×2 به همراه نماد گذاری‌های مربوط به نمایش رابطه بین یک متغیر پاسخ دوحالتی و متغیر مواجهه دوسطحی.

پیامد			
مواجهه	پیامد رخ داده D	پیامد رخ نداده H	مجموع
گروه ۱ (مواجهه دارد)	d_1	h_1	n_1
گروه ۲ (مواجهه ندارد)	d_2	h_2	n_2
مجموع	D	H	n

محاسبه شاخص‌های بیماری‌زایی، مواجهه، خطر، شانس و نظیر این‌ها برای چنین جدولی بسیار ساده‌تر است. برای مثال، اگر در جدول ۸-۱ مایل به پیدا کردن نسبت بیماری در هر گروه مواجهه و به صورت کلی باشیم، این نسبت‌ها به صورت:

$$P_1 = \frac{d_1}{n_1}$$

نسبت بیماری در گروه مواجهه

$$P_2 = \frac{d_2}{n_2}$$

نسبت بیماری در گروه بدون مواجهه

$$P = \frac{d}{n}$$

نسبت بیماری در کل افراد

قابل محاسبه هستند.

به همین صورت شیوع

$$\text{شیوع} = \frac{\text{تعداد افراد بیمار در یک مقطع زمانی}}{\text{کل جمعیت}}$$

و بروز

$$\text{بروز تجمعی} = \frac{\text{تعداد موارد جدید ابتلا در یک دوره زمانی}}{\text{تعداد افراد عاری از بیماری جامعه}}$$

از این جداول قابل محاسبه است.

۸-۳-۸ مقایسه دو نسبت با یکدیگر

در موارد زیادی مایل هستیم دو نسبت از دو جامعه، یا زیرگروه‌های یک جامعه را با یکدیگر مقایسه کنیم. این نسبت‌ها می‌توانند مقایسه تعداد موارد بهبودیافته ناشی از داروی جدید نسبت به داروی استاندارد، نسبت‌های بیماری‌زایی در دو اقلیم جغرافیایی متفاوت و غیره باشند. در این قسمت روش‌های آماری مختلف انجام این مقایسه‌ها را معرفی می‌کنیم.

- ♦ تفاضل خطر (خطر قابل انتساب)
- ♦ خطر نسبی
- ♦ نسبت شانس

نحوه محاسبه هر کدام از این شاخص‌ها در جدول ۸-۲ نشان داده شده است.

جدول ۸-۲: شاخص‌های مشهور برای مقایسه دو نسبت

شاخص مقایسه	رابطه برای محاسبه
تفاضل خطر (AT)	$P_1 - P_2$
خطر نسبی (RR)	$\frac{d_1}{d_1 + h_1} \frac{P_2}{P_1}$ و یا $\frac{d_1}{d_2 - h_2}$
نسبت شانس (OR)	$\frac{d_1}{h_1} \frac{d_2 \times h_1}{d_1 \times h_2}$ و یا $\frac{odds_1}{odds_2} = \frac{d_1}{d_2} \frac{h_2}{h_1}$

فاصله اطمینان ۹۵٪ برای تفاضل خطر به صورت زیر محاسبه می‌شود:

$$AT: P_1 - P_2 \pm 1/96 \sqrt{\frac{P_1(1-P_1)}{n_1} + \frac{P_2(1-P_2)}{n_2}}$$

در این رابطه P_1 برآورد نسبت بیماری در گروه مواجهه و P_2 برآورد نسبت بیماری در گروه بدون مواجهه است. فاصله اطمینان ۹۵٪ برای خطر نسبی به صورت زیر محاسبه می‌شود:

$$\log(RR): \log\left(\frac{P_1}{P_2}\right) \pm 1/96 \text{ se}(\log RR)$$

که در این رابطه $\text{se}(\log RR)$ به صورت:

$$\text{se}(\log RR) = \sqrt{\frac{1}{d_1} - \frac{1}{d_1 + h_1} + \frac{1}{d_2} - \frac{1}{d_2 + h_2}}$$

محاسبه می‌شود.

در این صورت کران بالا و پایین فاصله اطمینان ۹۵٪ به صورت زیر محاسبه می‌شود:

$$L_{\log} = \log RR - 1/96 \text{ se}(\log RR)$$

$$U_{\log} = \log RR + 1/96 \text{ se}(\log RR)$$

که در تمام روابط فوق \log ، لگاریتم در مبنای عدد نپر است. برای پیدا کردن فاصله اطمینان در واحد اصلی لازم

است کران‌های به دست آمده را به توان عدد e برسانیم. یعنی:

$$L = \exp(L_{\log})$$

$$U = \exp(U_{\log})$$

فاصله اطمینان ۹۵٪ برای نسبت شانس به صورت زیر محاسبه می‌شود:

$$\log(OR) : \log\left(\frac{\text{odds}_1}{\text{odds}_2}\right) \pm 1/96 \text{ se}(\log OR)$$

که در این رابطه $\text{se}(\log OR)$ به صورت:

$$\text{se}(\log OR) = \sqrt{\frac{1}{h_1} + \frac{1}{d_1} + \frac{1}{h_2} + \frac{1}{d_2}}$$

محاسبه می‌شود.

در این صورت کران بالا و پایین فاصله اطمینان ۹۵٪ به صورت زیر محاسبه می‌شود:

$$L_{\log} = \log OR - 1/96 \text{ se}(\log OR)$$

$$U_{\log} = \log OR + 1/96 \text{ se}(\log OR)$$

نظیر فاصله اطمینان برای RR لازم است کران‌های به دست آمده به توان عدد e رسانده شود. یعنی:

$$L = \exp(L_{\log})$$

$$U = \exp(U_{\log})$$

مثال: فرض کنید متغیر سیگار کشیدن به صورت سیگاری و غیرسیگاری بر اساس جنس در جدول ۸-۳ جمع‌آوری شده باشد.

جدول ۸-۳: وضعیت سیگار کشیدن بر اساس جنس

وضعیت سیگار کشیدن			
جنس	سیگاری	غیرسیگاری	مجموع
مرد	۲,۲۷۹	۲,۲۴۱	۴,۵۲۰
زن	۱,۵۶۲	۲,۵۵۹	۴,۱۶۱
مجموع	۳,۸۴۱	۴,۸۴۰	۸,۶۸۱

در این صورت نسبت سیگاری‌ها در مردها و زن‌ها برابر:

$$P = \frac{2279}{4520} = 0.5042$$

$$P_1 = \frac{1562}{4161} = 0.3754$$

است و فاصله اطمینان ۹۵٪ برای تفاضل خطر سیگاری بودن در مردها به زن‌ها به صورت زیر محاسبه می‌شود:

$$AT: 0.5042 - 0.3754 \pm 1/96 \sqrt{\frac{0.5042(1-0.5042)}{4520} + \frac{0.3754(1-0.3754)}{4161}}$$

که برابر 0.1288 ± 0.0207 و یا $(0.108, 0.150)$ خواهد بود.

مثال: جدول ۸-۴ از یک مطالعه چندمرکزی با عنوان The Pooling Project برای بررسی عوامل خطر در بیماری‌های قلبی-عروقی در مردها در شهرهای مختلف ایالات متحده انجام شده است، به دست آمده است [۱].

جدول ۸-۴: وضعیت سیگار کشیدن و ایجاد بیماری‌های قلبی و عروقی در طول پیگیری

ایجاد بیماری قلبی و عروقی در طول پیگیری			
مجموع	خیر	بله	وضعیت سیگار کشیدن در ابتدا مطالعه
۱۰۳۴۲	۱۰۱۷۶	۱۶۶	بله
۵۶۳	۵۱۳	۵۰	خیر
۱۰۹۰۵	۱۰۶۸۹	۲۱۶	مجموع

در این مطالعه برآورد خطر ابتلا به بیماری برای افراد سیگاری $P_1 = \frac{166}{1342} = 0.1237$ و در افراد غیرسیگاری

برابر $P = \frac{50}{563} = 0.0888$ است. در این صورت خطر نسبی ابتلا به بیماری‌های قلبی و عروقی در سیگاری‌ها

نسبت به غیرسیگاری‌ها برابر $RR = \frac{0.1237}{0.0888} = 1.393$ محاسبه می‌شود و خطای معیار برای لگاریتم خطر

نسبی برابر:

$$\sqrt{\frac{1}{166} - \frac{1}{1342} + \frac{1}{50} - \frac{1}{563}} = 0.1533$$

محاسبه می‌شود و داریم:

$$L_{\log} = \log(1/393) - 1/96 \times 0.1533 = 0.0310$$

$$U_{\log} = \log(1/393) + 1/96 \times 0.1533 = 0.06319$$

و در آخر کران‌های بالا و پایین فاصله اطمینان ۹۵٪ برای RR به صورت زیر است:

$$L = \exp(-0.310) = 1/0.31$$

$$U = \exp(-0.6319) = 1/881$$

در نتیجه محققان ۹۵٪ اطمینان دارند که فاصله اطمینان (۱/۰۳, ۱/۸۸) خطر نسبی حقیقی جامعه را در بر دارد. مثال: برای داده‌های جدول ۸-۵ فاصله اطمینان ۹۵٪ برای OR به صورت زیر است: شانس ایجاد بیماری قلبی در سیگاری‌ها نسبت به غیرسیگاری‌ها:

$$OR = \frac{\frac{166}{1176}}{\frac{50}{513}} = 1/448$$

خطای معیار این برآورد به صورت:

$$\sqrt{\frac{1}{166} + \frac{1}{1176} + \frac{1}{50} + \frac{1}{513}}$$

محاسبه می‌شود و در نهایت فاصله اطمینان ۹۵٪ برای OR به صورت (۱/۰۴, ۲/۰۲) خواهد بود.

۸-۴ مدل سازی

در این بخش به معرفی مختصر چهار مدل مشهور برای تحلیل چند متغیره می‌پردازیم. رگرسیون خطی، که در تحلیل پیامدهای پیوسته (نظیر فشارخون و وزن) کاربرد دارد. رگرسیون پواسن برای تعدیل میزان‌ها هنگام وجود چند مخدوش‌کننده بالقوه، به کار می‌رود. رگرسیون لجستیک که برای پیامدهای رسته‌ای (نظیر بروز بیماری و وقوع مرگ) به کار می‌رود و رگرسیون کاکس که در تحلیل داده‌های بقا استفاده می‌شود. روابط و جزییات محاسبه، تفسیر و ارزیابی این مدل‌ها در کتاب‌های آماری زیادی مورد بررسی قرار گرفته است. در این بخش از این روابط صرف نظر شده و فقط به بررسی کاربردهای آن اکتفا شده است. جدول ۸-۵ نحوه نمادگذاری این مدل‌ها را نشان می‌دهد.

جدول ۸-۵: نحوه نمادگذاری و تفسیر ضرایب رگرسیونی در مدل‌های رگرسیون چندگانه

مدل رگرسیون	نمادگذاری	تفسیر ضرایب (b)
- خطی	$Y = b_0 + b_1X_1 + b_2X_2 + \dots + b_kX_k$	هنگامی که تمام مخدوش‌کننده‌های موردنظر در مطالعه هم‌زمان تحت کنترل هستند به ازای هر یک واحد افزایش در متغیر X_1 ، میانگین پیامد موردبررسی (Y) به اندازه b_1 افزایش می‌یابد.
- پواسن	$\log(\text{rate}) = b_0 + b_1X_1 + b_2X_2 + \dots + b_kX_k$	هنگامی که تمام مخدوش‌کننده‌های موردنظر در مطالعه هم‌زمان تحت کنترل هستند به ازای هر یک واحد افزایش در متغیر X_1 ، لگاریتم میزان پیامد موردبررسی (Y) به اندازه b_1 افزایش می‌یابد.
- لجستیک	$\log(\text{odds}) = b_0 + b_1X_1 + b_2X_2 + \dots + b_kX_k$	هنگامی که تمام مخدوش‌کننده‌های موردنظر در مطالعه هم‌زمان تحت کنترل هستند به ازای هر یک واحد افزایش در متغیر X_1 ، لگاریتم شانس پیامد موردبررسی (Y) به اندازه b_1 افزایش می‌یابد.
- کاکس	$\log(\text{hazard}) = b_0 + b_1X_1 + b_2X_2 + \dots + b_kX_k$	هنگامی که تمام مخدوش‌کننده‌های موردنظر در مطالعه هم‌زمان تحت کنترل هستند به ازای هر یک واحد افزایش در متغیر X_1 ، لگاریتم مخاطره پیامد موردبررسی (Y) به اندازه b_1 افزایش می‌یابد.

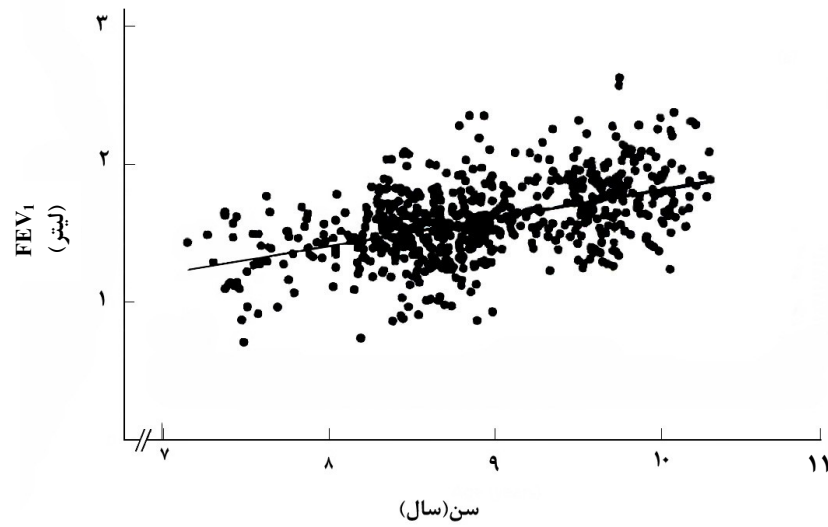
۸-۴-۱ رگرسیون خطی

موارد زیادی وجود دارد که محققان مایلند ارتباط هم‌زمان یک پیامد با مقیاس پیوسته را با تعدادی متغیر مستقل بررسی نمایند. این موضوع با به‌کارگیری روش رگرسیون خطی چندگانه امکان‌پذیر است. کاربردهای اصلی رگرسیون چندگانه عبارت‌اند از، برآورد اندازه اثر یک مواجهه هنگامی که اثر سایر عوامل مهار شده است. برای مثال اگر دو گروه مواجهه از نظر عواملی نظیر سن، جنس، طبقه اجتماعی که مؤثر بر پیامد موردبررسی هستند، متفاوت باشد، تعدیل اثر این عوامل برای قابل‌مقایسه‌کردن گروه‌های مواجهه بسیار مهم است. دلیل دیگر استفاده از مدل‌ها افزایش دقت برآورد اثر مواجهه به علت کم شدن انحراف معیار برآوردها است. برای مثال در یک مطالعه به‌منظور بررسی حجم ریه کودکان ۷ تا ۱۰ ساله، حداکثر میزان هوای تنفس شده در ۱ ثانیه (FEV_1)^{۱۵} به‌وسیله اسپیرومتر^{۱۶} اندازه‌گیری شد. متغیرهای سن، قد و هرگونه بیماری تنفسی در یک سال اخیر به‌عنوان عوامل مرتبط با پیامد موردنظر اندازه‌گیری شدند. نمودارهای ۸-۱ الف و ب، نحوه پراکنش FEV_1 با سن و قد را برای ۶۳۰ کودک نمایش می‌دهد [۲].

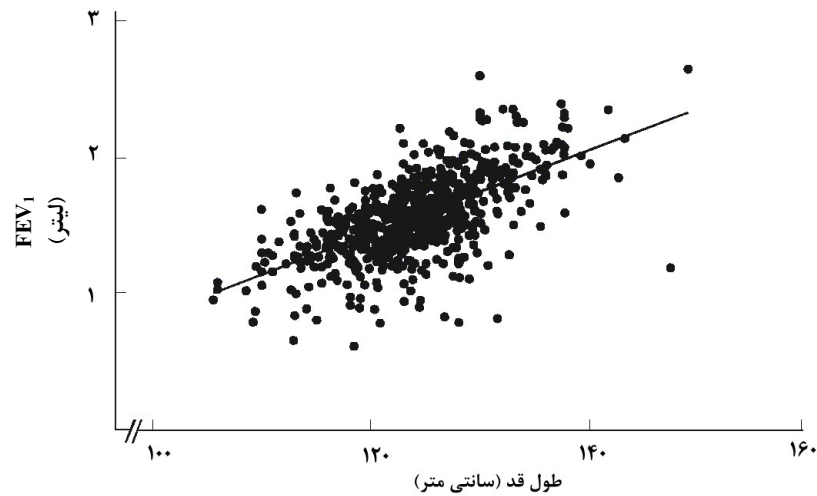
۱. Forced Expiratory Volume in 1 Second

۲. Spirometer

(الف)



(ب)



نمودار ۸-۱: نمودار پراکندگی FEV_1 به همراه سن (الف) و قد ۶۳۶ کودک (ب)

همان‌طور که از نمودارها ملاحظه می‌شود، امکان وجود ارتباط خطی نیرومند بین FEV_1 ، سن و قد بسیار محتمل است. این موضوع به وسیله آزمون ضرایب رگرسیونی تأیید می‌شود. جدول ۴-۸ نتایج برازش مدل رگرسیون

خطی به این داده‌ها را نمایش می‌دهد.

جدول ۸-۶: خروجی کامپیوتری برآورد ضرایب رگرسیونی چندگانه برای اندازه‌گیری FEV₁

فاصله اطمینان ۹۵٪	مقدار احتمال	انحراف معیار	ضریب	FEV ₁
-۲/۶۶۴۵ تا ۱/۹۵۲۹	۰/۰۰۰	۰/۰۸۱۲	-۲/۳۰۸۷	مقدار ثابت
۰/۰۲۱۴ تا ۰/۰۲۳۸۵	۰/۰۰۰	۰/۰۱۵۷	۰/۰۸۹۷	سن
۰/۰۵۸۸ تا ۰/۱۲۰۶	۰/۰۰۰	۰/۰۰۱۸	۰/۰۲۵۰	قد

با توجه به جدول ۸-۶ بهترین مدل خطی برای این داده‌ها به صورت زیر است:

$$FEV_1 = -2/3087 + 0/0897 \times \text{سن} + 0/0250 \times \text{قد}$$

۸-۴-۲ رگرسیون پوآسن

در این بخش به معرفی مدل رگرسیون پوآسن برای مقایسه میزان‌ها می‌پردازیم. نکته مهم این مدل محاسبه برآورد در مقیاس لگاریتمی است. به همین دلیل باید برای تحلیل، ضرایب را به مقیاس واقعی تبدیل کرد. در این مدل میزان پیامد موردبررسی در گروه‌های مطالعه پس از تعدیل برای عوامل مخدوش‌کننده محاسبه می‌شود. جدول ۸-۷ حالت ساده‌ای از این مدل برای مقایسه دو گروه را نمایش می‌دهد.

جدول ۸-۷: معادله‌های مدل رگرسیون پوآسن برای مقایسه دو گروه مواجهه

گروه مواجهه	میزان پیامد	لگاریتم میزان
دارای مواجهه	میزان مینا × نسبت میزان مواجهه	لگاریتم (میزان مینا) + لگاریتم (نسبت میزان مواجهه)
بدون مواجهه	میزان مینا	لگاریتم (میزان مینا)

میزان مینا، میزان ابتلا در یکی از گروه‌های مواجهه است که میزان‌های حاصل از گروه‌های دیگر با آن مقایسه می‌شوند.

نسبت میزان مواجهه، اثر مواجهه بر میزان ابتلا بیماری را بیان می‌کند و به صورت زیر محاسبه می‌شود.

$$\text{نسبت میزان مواجهه} = \frac{\text{نرخ ابتلا در گروه دارای مواجهه}}{\text{نرخ ابتلا در گروه بدون مواجهه}}$$

مثال زیر کاربرد رگرسیون پوآسن در بررسی عوامل خطر ابتلا به بیماری‌های قلبی و عروقی را نمایش می‌دهد. ۱۷۸۶ مرد در مطالعه‌ای به منظور بررسی عوامل خطر در ابتلا به بیماری‌های قلبی - عروقی موردبررسی قرار گرفتند. افراد تحت مطالعه در زمان اولین معاینه دارای محدوده سنی ۴۳ تا ۶۱ سال بودند و تا ۱۹ سال مورد پیگیری قرار گرفتند. اولین معاینه در بین سال‌های ۱۹۷۹-۱۹۸۳ میلادی صورت پذیرفته است. اطلاعات پیامد

موردبررسی (سکته قلبی و یا مرگ ناشی از بیماری قلبی) در سال ۱۹۹۹ میلادی جمع‌آوری شد. اطلاعات بیشتر درباره این مطالعه، در سایت اینترنتی www.epi.bris.ac.uk/mrc-caarphilly موجود است (تاریخ دسترسی ۸۶/۱/۲۰).

جدول ۸-۸، ده سطر اول بانک اطلاعاتی مربوط به این مطالعه را نمایش می‌دهد.

جدول ۸-۸: ده سطر اول بانک اطلاعاتی مطالعه بررسی عوامل خطر بیماری‌های قلبی و عروقی (Caarphillystudy)

شماره شناسایی	تاریخ تولد	تاریخ اولین معاینه	تاریخ خروج از مطالعه	سال‌های پیگیری	سکته قلبی	وضعیت استعمال دخانیات
۱	20/May/1929	17/Jun/1982	31/Dec/1998	۱۶/۴۵	۰	۱
۲	9/Jul/1930	10/Jan/1983	24/Oct/1998	۱۵/۹۵	۰	۰
۳	6/Feb/1929	23/Dec/1982	26/Nov/1998	۱۵/۹۳	۰	۱
۴	24/May/1931	7/Jul/1983	22/Nov/1984	۱۳/۳۸	۱	۰
۵	9/Feb/1934	3/Sep/1980	19/Dec/1998	۱۸/۲۹	۰	۰
۶	14/May/1930	17/Nov/1981	31/Dec/1998	۱۷/۱۲	۰	۰
۷	13/May/1924	30/Oct/1980	27/Dec/1998	۱۸/۱۶	۰	۱
۸	23/May/1924	24/Apr/1980	24/Jan/1980	۵/۷۵	۱	۱
۹	11/Jun/1890	11/Jun/1980	12/Dec/1998	۱۵/۱۸	۱	۰
۱۰	12/May/1929	17/Nov/1979	20/Jan/1995	۱۸/۵	۱	۰

وضعیت فعلی استعمال دخانیات یک متغیر اسمی است که برای سیگاری‌ها ۱ و برای غیر سیگاری‌ها صفر به کاررفته است. در مورد سکته قلبی هم کد ۱ نشانه مرگ ناشی از بیماری قلبی یا سکته قلبی و کد صفر برای غیره به کاررفته است.

در مجموع ۹۹۰ مرد سیگاری در مطالعه وجود داشت و ۷۹۶ مرد غیرسیگاری که هرگز سیگار نکشیده بودند یا سیگار را ترک کرده بودند. جدول ۸-۹، میزان‌های بروز سکته قلبی در این دو گروه را نمایش می‌دهد. نسبت

میزان دو گروه هرگز سیگار نکشیده به سیگاری‌ها برابر $\frac{16}{98} = 1/7$ است.

جدول ۸-۹: میزان‌های سکته قلبی در دو گروه سیگاری و غیرسیگاری.

وضعیت فعلی استعمال سیگار	تعداد سکته‌های قلبی	شغل - سال در معرض خطر	میزان در هر ۱۰۰۰ شخص - سال
بله (مواجهه)	$d_1 = 230$	$T_1 = 13,978$	$\lambda_1 = \frac{230}{13,978} = 16/98$
خیر (بدون مواجهه)	$d_0 = 118$	$T_0 = 12,183$	$\lambda_0 = \frac{118}{12,183} = 9/68$

مجموع	$d = ۳۴۸$	$T = ۲۶,۱۶۱$	$\lambda = \frac{۳۴۸}{۲۶/۱۶۱} = ۱۳/۳۰$
-------	-----------	--------------	--

جدول ۸-۱۰ خروجی کامپیوتری از تحلیل این اطلاعات با روش رگرسیون پواسن را نمایش می‌دهد.

جدول ۸-۱۰: خروجی رگرسیون پواسن برای محاسبه میزان سگته قلبی در دو گروه سیگاری و غیرسیگاری

نسبت میزان	Z	$P > z $	فاصله اطمینان ۹۵٪
سیگاری بودن	۴/۶۸۰	۰/۰۰۰	۱/۳۶۱ و ۲/۱۲۱
ضریب ثابت	-۵۰/۳۷	۰/۰۰۰	۰/۰۰۸۰۹ و ۰/۰۱۱۶

در این صورت خواهیم داشت:

$$\text{وضعیت سیگار کشیدن (۱/۷)} \times \exp(۰/۰۰۹۶۹) = \text{میزان ابتلا به سگته قلبی}$$

۸-۴-۳ رگرسیون لجستیک

مشهورترین مدل برای تحلیل چند متغیره داده‌های کیفی مدل رگرسیون لجستیک است. این مدل نیز مانند رگرسیون پواسن در مقیاس لگاریتمی محاسبه می‌شود و لگاریتم نسبت شانس مواجهه در هر کدام از گروه‌های مطالعه نسبت به یک گروه مبنا (و تعدیل هم‌زمان برای مخدوش‌کننده‌های بالقوه) محاسبه می‌شود. نسبت شانس مواجهه را می‌توان به صورت زیر تعریف کرد.

$$\text{نسبت شانس مواجهه} = \frac{\text{شانس ابتلا در گروه مواجهه}}{\text{شانس ابتلا در گروه بدون مواجهه}}$$

در جدول ۸-۱۱ این روابط در مبنای لگاریتمی نظیر آنچه در مدل لجستیک به کار می‌رود برای مقایسه دو گروه مواجهه نشان داده شده است [۳].

جدول ۸-۱۱: معادله‌های مدل رگرسیون لجستیک برای مقایسه دو گروه مواجهه

گروه مواجهه	شانس پیامد	لگاریتم شانس
مواجهه	شانس مبنا \times شانس مواجهه	لگاریتم (شانس مبنا) + لگاریتم (شانس مواجهه)
بدون مواجهه	شانس مبنا	لگاریتم (شانس مبنا)

مطالعه‌ای در کشور سیرالیون درباره بیماری کوری رودخانه^{۱۷} انجام پذیرفت. در این مطالعه افراد بر اساس این که در روستاهای واقع در ساوانا (مراغ) و یا مناطق جنگلی زندگی می‌کردند طبقه‌بندی شدند. افراد این مطالعه

۱. River blindness

برحسب و عدم وجود میکروفیلاریای انکوسرکا ولولوس^{۱۸} (mf) در زیر پوست مورد بررسی قرار گرفتند. جدول ۸-۱۲ شیوع این پیامد در دو گروه را نمایش می‌دهد [۴].

جدول ۸-۱۲: تعداد و درصد افراد آلوده به کوری رودخانه به تفکیک مکان سکونت

محل سکونت	آلودگی به mf		مجموع	شانس آلودگی
	دارد	ندارد		
جنگل استوایی	$d_1 = 541$ (% ۷۱/۷)	$h_1 = 213$ (% ۲۸/۳)	۷۵۴	$\frac{541}{754} = 2/540$
ساوانا	$d_2 = 281$ (% ۵۱/۳)	$h_2 = 267$ (% ۴۸/۷)	۵۴۸	$\frac{281}{548} = 1/0.52$
مجموع	۸۲۲	۴۸۰	۱۳۰۲	

در این مطالعه مایل هستیم اثر سن را بر رابطه شیوع بیماری و محل سکونت بررسی کنیم. جدول ۸-۱۳ رابطه بین سن و آلودگی به mf را نشان می‌دهد.

جدول ۸-۱۳: رابطه بین گروه‌های سنی و mf

گروه‌های سنی	آلودگی به mf		نسبت شانس در مقایسه با گروه مبنا
	دارد	ندارد	
۵-۹	۴۶	۱۵۶	$\frac{46}{156} = 0.295$
۱۰-۱۹	۹۹	۱۱۹	$\frac{99}{119} = 0.832$
۲۰-۳۹	۲۹۹	۱۲۵	$\frac{299}{125} = 2.392$
≥ 40	۳۷۸	۸	$\frac{378}{8} = 4.725$
مجموع	۸۲۲	۴۸۰	

جدول ۸-۱۴ خروجی کامپیوتری این مدل را نمایش می‌دهد.

جدول ۸-۱۴: خروجی کامپیوتری برای مدل آلودگی به mf با محل سکونت و گروه‌های سنی

فاصله اطمینان %۹۵	$P > z $	Z	نسبت شانس	محل سکونت*
۲/۳۵۴-۴/۰۳۸	۰/۰۰۰	۸/۱۸۱	۳/۰۸۳	محل سکونت*
۱/۶۸۲-۴/۰۱۶	۰/۰۰۰	۴/۳۰۱	۲/۵۹۹	گروه سنی (۱)
۶/۴۹۳-۱۴/۶۹	۰/۰۰۰	۱۰/۹۴۴	۹/۷۶۵	گروه سنی (۲)
۱۱/۵۶-۲۶/۹۳	۰/۰۰۰	۱۳/۲۹۵	۱۷/۶۴	گروه سنی (۳)
۰/۱۰۰-۰/۲۷	۰/۰۰۰	-۹/۷۴۱	-۰/۱۴۷	مقدار ثابت

* محل سکونت (۱= جنگل استوایی، ۰= ساوانا)

گروه سنی: (۰-۵=۹ سال، ۱=۱۰ تا ۱۹ سال، ۲=۲۰ تا ۳۹ سال، ۳=بیشتر از ۴۰ سال)

۸-۴-۴ تحلیل داده‌های بقا

روش‌های تحلیل و مدل‌بندی معرفی شده در قسمت‌های قبلی بر این اساس استوار هستند که میزان ابتلا به بیماری در طول مطالعه ثابت باقی می‌ماند؛ اما در برخی از مطالعات طولی که بر مبنای پیگیری افراد تحت مطالعه طراحی می‌شوند، ممکن است پذیرش این فرض مناسب نباشد. برای مثال:

- ♦ خطر مرگ پس از انجام عمل جراحی قلب بسیار زیاد است و پس از سپری شدن مرحله نقاهت با گذشت زمان کاهش می‌یابد.
 - ♦ میزان عود تومور پس از تشخیص و درمان با گذشت زمان تغییر پیدا می‌کند.
- روش‌های معروف به تحلیل بقا، به‌منظور بررسی چنین وقایعی بدون پذیرش فرض ثابت بودن میزان ابتلا در طول زمان طراحی شده‌اند. دو شاخص اصلی در تحلیل این داده‌ها عبارت‌اند از:
- مخاطره که با $h(t)$ نمایش داده می‌شود و بیانگر میزان وقوع لحظه‌ای در زمان t است. تابع بقا که با $s(t)$ نمایش داده می‌شود و نشان‌دهنده منحنی بقا افراد تحت مطالعه است.
- روش‌های جدول عمر، کاپلان-مایر و نلسون-آلن معروف‌ترین روش‌های توصیفی برای برآورد و توصیف داده‌های بقا هستند.

مثال: در یک مطالعه کارآزمایی تصادفی شده برای تعیین مقدار اثربخشی داروی آزاتیوپرین^{۱۹} بر بیماری سیروز صفراوی در برابر دارونما، ۲۸۴ بیمار وارد مطالعه شده و به مدت ۱۲ سال پیگیری شدند. از این ۲۸۴ بیمار، ۲۴۸ بیمار دارای اطلاعات تکمیل شده همه متغیرهای مبنا بودند. از این تعداد ۳۱ بیمار دارای کِلستاز درونی^{۲۰} (مشخصه‌ای که بیانگر حادثه شدن بیماری است) بودند که ۲۴ نفر از آن‌ها فوت کردند و ۷ بیمار دیگر پیگیری ناتمام داشتند. اطلاعات این بیماران در جدول ۸-۱۵ نمایش داده شده است [۵].

۱. Azathioprine

۲. Central cholestasis

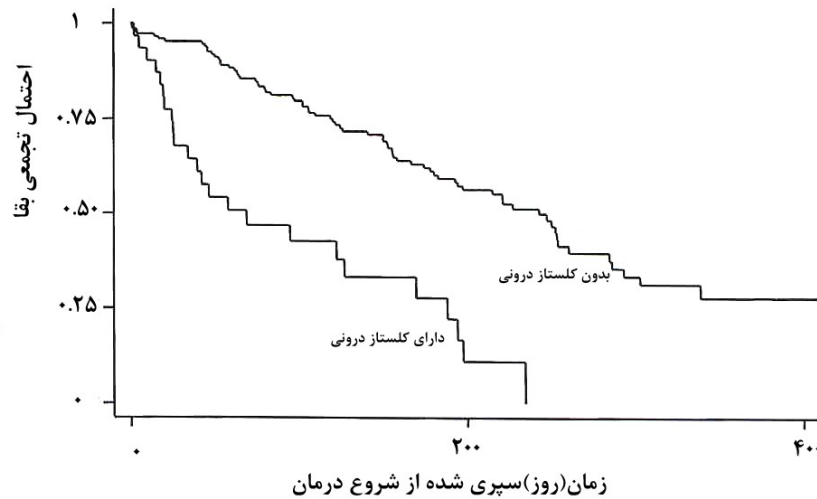
جدول ۸-۱۵: نحوه محاسبه برآورد کاپلان- مایر از تابع بقا برای ۳۱ بیمار مبتلا به سیروز صفراوی دارای کلستاز درونی

زمان (روز)	تعداد در معرض خطر در زمان وقوع سبب	تعداد مرگ‌ها در زمان t	تعداد افراد سانسور شده در پیگیری تا زمان t	خطر مرگ $r_t = d_t / n_t$	احتمال بقا $S_t = 1 - r_t$	تابع بقا $S^*(S(t)) = S(\text{percentage})$
۱۹	۳۱	۱	۰	۰/۰۲۲۳	۰/۸۷۶۶	۰/۸۷۶۶
۶۸	۲۰	۱	۰	۰/۰۲۲۳	۰/۸۵۴۳	۰/۸۵۴۳
۶۹	۲۹	۱	۰	۰/۰۳۴۵	۰/۸۲۹۸	۰/۸۲۹۸
۱۵۰	۲۸	۱	۰	۰/۰۳۵۷	۰/۸۰۴۱	۰/۸۰۴۱
۱۷۷	۲۷	۱	۰	۰/۰۳۷۰	۰/۷۷۷۱	۰/۷۷۷۱
۱۹۳	۲۶	۱	۰	۰/۰۳۸۵	۰/۷۴۸۶	۰/۷۴۸۶
۲۰۱	۲۵	۱	۰	۰/۰۴۰۰	۰/۷۰۸۶	۰/۷۰۸۶
۲۴۵	۲۴	۱	۰	۰/۰۴۱۷	۰/۶۶۶۹	۰/۶۶۶۹
۲۵۱	۲۳	۱	۰	۰/۰۴۳۵	۰/۶۲۱۹	۰/۶۲۱۹
۲۵۶	۲۲	۱	۰	۰/۰۴۵۵	۰/۵۷۷۴	۰/۵۷۷۴
۲۰۲	۲۱	۰	۱	۰	۰/۵۷۷۴	۰/۵۷۷۴
۳۹۱	۲۰	۰	۱	۰/۰۵۰۰	۰/۵۲۲۵	۰/۵۲۲۵
۳۹۵	۱۹	۱	۰	۰/۰۵۲۶	۰/۴۶۹۹	۰/۴۶۹۹
۴۲۱	۱۸	۱	۰	۰/۰۵۵۶	۰/۴۱۴۳	۰/۴۱۴۳
۴۶۴	۱۷	۱	۰	۰/۰۵۸۸	۰/۳۵۹۴	۰/۳۵۹۴
۵۷۸	۱۶	۱	۰	۰/۰۶۲۵	۰/۳۰۶۹	۰/۳۰۶۹
۵۸۳	۱۵	۰	۱	۰	۰/۳۰۶۹	۰/۳۰۶۹
۵۸۶	۱۴	۰	۱	۰	۰/۳۰۶۹	۰/۳۰۶۹
۶۸۸	۱۳	۱	۰	۰/۰۷۶۹	۰/۲۳۰۱	۰/۲۳۰۱
۸۲۸	۱۲	۱	۱	۰	۰/۲۳۰۱	۰/۲۳۰۱
۹۴۷	۱۱	۱	۰	۰/۰۹۰۹	۰/۱۳۹۲	۰/۱۳۹۲
۱۱۵۹	۱۰	۰	۱	۰	۰/۱۳۹۲	۰/۱۳۹۲
۱۲۱۹	۹	۰	۱	۰/۱۱۱۱	۰/۰۸۸۱	۰/۰۸۸۱
۱۲۶۸	۸	۰	۱	۰/۱۲۵۰	۰/۰۷۵۰	۰/۰۷۵۰
۱۲۹۳	۷	۰	۱	۰	۰/۰۷۵۰	۰/۰۷۵۰
۱۶۹۳	۶	۰	۱	۰/۱۶۶۷	۰/۰۱۸۳	۰/۰۱۸۳
۱۸۸۱	۵	۰	۱	۰/۲۰۰۰	۰/۰۱۰۰	۰/۰۱۰۰
۱۹۴۰	۴	۰	۱	۰/۲۵۰۰	۰/۰۰۷۵	۰/۰۰۷۵
۱۹۷۵	۳	۰	۱	۰/۳۳۳۳	۰/۰۰۶۶	۰/۰۰۶۶
۲۳۳۸	۲	۰	۱	۰	۰/۱۱۰۵	۰/۰۱۱۰۵
۲۳۴۳	۱	۰	۱	۰	۰	۰



جدول ۸-۱۵: نحوه محاسبه برآورد کاپلان- مایر از تابع بقا برای ۳۱ بیمار مبتلا به سیروز صفراوی دارای کلستاز درونی

اولین مرگ در روز ۱۹ اتفاق افتاده است. پس خطر مرگ در ۱۹ روز اول برابر $r_{19} = \frac{1}{31} = 0.0323$ است. احتمال بقا تا این روز به صورت $S_{19} = 1 - 0.0323 = 0.9677$ و تابع بقا به صورت $S_{19} = S_{19} = 0.9766$ است. مرگ بعدی در روز ۴۸ اتفاق افتاده است. در این زمان ۳۰ بیمار در معرض خطر قرار دارند و خطر مرگ در این دوره برابر $r_{48} = \frac{1}{30} = 0.0333$ است. احتمال زنده ماندن تا این روز برابر $S_{48} = 1 - 0.0333 = 0.9667$ و تابع بقا برابر $S(48) = S_{19} \times S_{48} = 0.9355$ است. مقادیر این شاخص‌ها به طور کامل در جدول ۸-۱۵ نمایش داده شده است. در نمودار ۸-۲ برآورد کاپلان-مایر تابع بقا برای دو گروه دارای کِلستاز درونی و بدون کِلستاز درونی نمایش داده شده است.



نمودار ۸-۲: برآورد کاپلان-مایر از تابع بقا $S(t)$ برای دو گروه از بیماران مبتلا به سیروز صفراوی

این گروه‌ها را می‌توان به وسیله انجام آزمون لگ-رتبه از لحاظ آماری مقایسه کرد. از مفروضات اصلی در بسیاری از روش‌های توصیفی و تحلیلی بقا، متناسب بودن مخاطره در طول زمان است. به عبارت دیگر فرض می‌کنیم در تمام دوره مطالعه این رابطه

$$\frac{h_1(t)}{h_2(t)} = \text{ثابت}$$

برقرار است.

۸-۴-۱ آزمون لگ-رتبه

برای مقایسه زمان‌های بقا در دو گروه مختلف، روش متداول مقایسه نسبت‌های مخاطره در دو گروه است. آزمون لگ-رتبه نیز بر همین اساس بنا شده است. در این روش تعداد موارد مرگ اتفاق افتاده در هر دوره با تعداد موارد مرگ مورد انتظار، هنگامی که نسبت مخاطره‌ها برابر یک است (دو گروه دارای زمان‌های بقا یکسان هستند) مقایسه می‌شوند.

برای مقایسه هم‌زمان میزان‌های مخاطره که برای متغیرهای مخدوش‌کننده تعدیل شده‌اند مدل‌های پارامتری و نیمه پارامتری بسیار متنوعی پیشنهاد شده‌اند که مشهورترین آن‌ها، مدل‌های رگرسیون کاکس و وایبول هستند.

در این قسمت به معرفی، مرور و بررسی روش‌های پیشرفته آماری پرداخته می‌شود. بررسی دقیق و کامل این روش‌ها و همچنین روش‌های متداول دیگر در تحلیل داده‌ها خارج از حوصله این کتاب است و خوانندگان محترم می‌توانند به منابعی نظیر [۶] تا [۱۹] مراجعه کنند.

۸-۵ انتخاب روش مناسب برای تحلیل داده‌ها

هدف عمده این فصل معرفی روش‌های آماری موردنیاز برای تجزیه و تحلیل اثر گروهی از متغیرها به متغیر پاسخ است که می‌توانند شامل متغیرهای مخدوش‌کننده همراه و مواجهه یا درمان باشند. جدول ۸-۱۶ را به‌طور خلاصه بر اساس هدف مطالعه و نوع متغیرهای موردبررسی بیان می‌کند.

جدول ۸-۱۶: خلاصه‌ای از روش‌های تحلیل داده‌ها

متغیرهای پیوسته
<p>یک گروه مواجهه</p> <p>میانگین پیامد، فاصله اطمینان برای میانگین حجم نمونه زیاد: استفاده از روش‌های مبتنی بر توزیع نرمال حجم نمونه کم: استفاده از روش‌های مبتنی بر توزیع t نرمال نبودن داده‌ها: استفاده از تبدیل برای نرمال کردن داده‌ها، استفاده از روش‌های ناپارامتری (روش‌های مبتنی بر رتبه‌ها، روش بوتسترپ)</p>
<p>دو گروه مواجهه</p> <p>تفاضل بین میانگین‌ها، فاصله اطمینان برای تفاضل میانگین‌ها حجم نمونه زیاد: استفاده از روش‌های مبتنی بر توزیع نرمال حجم نمونه کم: استفاده از روش‌های مبتنی بر توزیع t نرمال نبودن داده‌ها: استفاده از تبدیل برای نرمال کردن داده‌ها نامساوی بودن انحراف معیار دو گروه: استفاده از آزمون فیشر-بهرن، آزمون ولش روش‌های ناپارامتری (روش‌های مبنی بر رتبه‌ها، روش بوتسترپ، روش‌های استوار برآورد) اندازه‌های زوج شده: آزمون‌های مربوط به داده‌های زوجی</p>
<p>بیش از دو گروه مواجهه</p>

<p>آنالیز واریانس یک طرفه رگرسیون خطی ساده متغیرهای مواجهه پیوسته یا رتبه‌ای رگرسیون خطی ساده ضریب همبستگی پیرسون ضریب همبستگی رتبه‌ای اسپیرمن، ضریب همبستگی کندال بررسی روابط غیرخطی وجود متغیرهای مخدوش کننده یا همراه رگرسیون چندگانه آنالیز واریانس دو یا چند طرفه استفاده از روش‌های استانداردسازی مستقیم</p>
متغیرهای رسته‌ای با دو رده
یک گروه مواجهه
نسبت، فاصله اطمینان برای نسبت آزمون فرض برای بررسی مقدار نسبت شانس، فاصله اطمینان برای شانس
دو گروه مواجهه
معیارهای برآورد اثر مواجهه: تفاضل نسبت‌ها نسبت خطر نسبت شانس آزمون کای-دو برای جداول توافقی 2×2 آزمون‌های دقیق برای نمونه‌های کم
بیش از دو گروه مواجهه
آزمون کای-دو برای جداول توافقی 2×2 رگرسیون لجستیک متغیرهای مواجهه پیوسته یا رتبه‌ای آزمون کای-دو برای بررسی روند رگرسیون لجستیک وجود متغیرهای مخدوش کننده یا همراه روش مانتل-هانزل رگرسیون لجستیک رگرسیون پواسن روش استانداردسازی مستقیم داده‌های همسان شده یا زوج شده مقایسه نسبت‌ها نسبت شانس و آزمون مک-نمار

رگرسیون لجستیک شرطی
متغیرهای رسته‌ای با بیش از دو رده
آزمون کای-دو برای جداول $r \times c$ رگرسیون لجستیک چندحالتی رگرسیون لجستیک رتبه‌ای
میزان‌ها و زمان بقا
یک گروه مواجهه
میزان، فاصله اطمینان برای میزان تحلیل بقا: جدول عمر برآوردهای کاپلان-مایر برای منحنی بقا
دو گروه مواجهه
روش‌های مبتنی بر میزان‌ها: تفاضل میزان‌ها نسبت میزان‌ها رگرسیون پواسن تحلیلی بقا: نسبت مخاطره‌ها (آزمون لگ-رتبه) رگرسیون کاکس بررسی متناسب بودن مخاطره‌ها
بیش از دو گروه مواجهه
رگرسیون پواسن رگرسیون کاکس متغیرهای مواجهه پیوسته یا رتبه‌ای رگرسیون پواسن رگرسیون کاکس وجود متغیرهای مخدوش‌کننده یا همراه: روش مانتل-هانزل رگرسیون پواسن روش استانداردسازی مستقیم روش استانداردسازی غیرمستقیم

۶-۸ تحلیل داده‌ها بر اساس نوع طرح مطالعه

هرکدام از انواع مطالعات که تابع حال مورد بررسی قرار گرفته‌اند دارای اهداف و ساختار متفاوتی هستند که این ویژگی‌ها باید در تجزیه و تحلیل داده‌ها مورد توجه واقع شود. صرف نظر از داده‌های توصیفی که از ماهیت پیچیده‌ای در تحلیل داده‌ها برخوردار نیستند، در این بخش به بررسی کارآزمایی‌های بالینی تصادفی شده و مطالعات مشاهده‌ای می‌پردازیم.

۱-۶-۸ کارآزمایی‌های بالینی تصادفی شده

نکات مهم برای انتخاب روش آماری در تحلیل این نوع داده‌ها عبارت‌اند از:

- ♦ نوع پاسخ مورد بررسی
- ♦ نیاز به تعدیل برای متغیرهای مینا
- ♦ نیاز به انجام مقایسه بین زیرگروه‌های مورد بررسی
- ♦ داده‌های طولی

دو نکته کلیدی در تحلیل کارآزمایی‌های تصادفی علاوه بر موارد قبلی وجود داده‌های خوشه‌ای و انجام طرح‌های متقاطع است.

جدول ۸-۱۷ نحوه انتخاب روش مناسب در تحلیل این داده‌ها را نشان می‌دهد.

۲-۶-۸ مطالعات طولی و متقاطع

نکات مهم برای انتخاب روش آماری مناسب برای تحلیل این مطالعات عبارت‌اند از:

- ♦ نوع طرح نمونه‌گیری
 - ♦ تعداد انجام اندازه‌گیری‌ها (یک مرتبه، به‌طور پیوسته یا به دفعات مکرر در طول زمان)
 - ♦ نوع متغیر پاسخ و مواجهه
- جدول ۸-۱۸ نحوه انتخاب روش مناسب در تحلیل این داده‌ها را نمایش می‌دهد.

جدول ۸-۱۷: روش تحلیل داده‌ها در مطالعات کارآزمایی تصادفی

جدول ۸-۱۷: روش تحلیل داده‌ها در مطالعات کارآزمایی تصادفی

نوع پاسخ مورد بررسی				
پیوسته	اسمی	نرخ	زمان بقاء	
توصیف داده‌ها	میانگین و انحراف معیار	جداول 2×2 یا $2 \times r$	تعداد وقایع، نرخ و شخص- زمان (به همراه فاصله اطمینان) در هر گروه	تعداد وقایع و شخص- زمان در هر گروه منحنی کاپلان- مایر
تعیین اندازه اثر	تفاضل بین میانگین‌ها آزمون t	تفاضل خطر، نسبت شانس، نسبت خطر، آزمون Z ، آزمون کای- دو	نسبت نرخ، آزمون Z	آزمون لگ-رتبه آزمون مانتل- کاکس
تعدیل متغیرهای مبنا	رگرسیون خطی چندگانه	روش مانتل- هانزل رگرسیون لجستیک	روش مانتل- هانزل رگرسیون پواسن	رگرسیون کاکس
تعیین اختلاف اثر درمان در زیرگروه‌های مورد مطالعه	اضافه کردن اثر متقابل در مدل‌های رگرسیون		به علاوه اثر متقابل، آزمون متناسب بودن مخاطره‌ها در گروه‌ها	
موارد خاص	مطالعه کارآزمایی تصادفی خوشه‌ای و طرح‌های متقاطع			

جدول ۸-۱۸: روش تحلیل داده‌ها در مطالعات متقاطع و طولی

جدول ۸-۱۸: روش تحلیل داده‌ها در مطالعات متقاطع و طولی

نوع پاسخ مورد بررسی				
نوع مواجهه	پیوسته	اسمی	نرخ	زمان بقاء
دو حالتی	تفاضل بین میانگین‌ها آزمون t	نسبت شانس، نسبت خطر آزمون کای-دو	نسبت نرخ آزمون Z	آزمون لگ-رتبه آزمون مانتل-کاکس
رسته‌ای	میانگین گروه‌ها آنالیز واریانس رگرسیون	نسبت‌های شانس بر اساس گروه مینا رگرسیون لجستیک	نسب‌های نرخ بر اساس گروه مینا رگرسیون پواسن	نسبت‌های مخاطره بر اساس گروه مینا رگرسیون کاکس
رتبه‌ای	افزایش میانگین در هر گروه رگرسیون خطی	افزایش لگاریتم شانس در هر گروه رگرسیون لجستیک آزمون کای-دو برای روند	افزایش لگاریتم نرخ در هر گروه رگرسیون پواسن	افزایش لگاریتم مخاطره مخاطره در هر گروه رگرسیون کاکس
پیوسته	رگرسیون خطی	رگرسیون لجستیک	رگرسیون پواسن	رگرسیون کاکس
تعدیل برای مخدوش کننده‌ها	رگرسیون خطی چندگانه	رگرسیون لجستیک چندگانه روش مانتل-هانزل	رگرسیون پواسن چندگانه روش مانتل-هانزل	رگرسیون کاکس چندگانه
موارد خاص	داده‌های خوشه‌ای (اندازه‌های مکرر در مطالعات طولی، اندازه‌گیری متعدد در طول زمان برای هر فرد، نمونه‌گیری خوشه‌ای)			

۸-۶-۳ مطالعات مورد-شاهدی

برای تحلیل مطالعات مورد-شاهدی غیر همسان شده روش‌های متداول تحلیل داده‌های اسمی قابل به‌کارگیری هستند. در چنین مواردی تحلیل بر مبنای جداول ساده 2×2 ، جدول $c \times 2 \times 2$ با روش مانتل-هانزل و احتمالاً مدل‌سازی با استفاده از روش رگرسیون لجستیک متداول است؛ اما برای مطالعات مورد-شاهدی همسان شده نیاز به روش‌های پیچیده‌تری وجود دارد. جدول ۸-۱۹ روش‌های مناسب برای تحلیل مطالعات مورد-شاهدی را نمایش می‌دهد.

جدول ۸-۱۹: مطالعه روش‌های مناسب برای تحلیل داده‌های مورد-شاهدی

طراحی	یک متغیر مواجهه	تعدیل برای متغیرهای مخدوش‌کننده
تصادفی ساده (مورد-شاهدی بدون همسان‌سازی)	جداول توافقی 2×2 نشان‌دهنده وضعیت هم‌زمان مواجهه و بیماری، نسبت شانس، آزمون کای-دو	رگرسیون لجستیک روش مانتل-هانزل
همسان‌سازی به گروهی (موردها و شاهد‌های متعدد)	جداول توافقی 2×2 در هر طبقه آزمون مانتل-هانزل نسبت شانس آزمون کای-دو	تحلیل به روش طبقه‌بندی رگرسیون لجستیک بر اساس کنترل هم‌زمان عوامل همسان شده و مخدوش‌کننده
همسان‌سازی فردی (زوج شده)	جداول 2×2 نشان‌دهنده توافق بین زوج‌های مورد-شاهد با توجه به عامل خطر نسبت شانس بر اساس زوج‌ها آزمون مک-نمار	تحلیل به روش طبقه‌بندی رگرسیون لجستیک بر اساس کنترل هم‌زمان عوامل همسان شده و مخدوش‌کننده
شاهد‌های زوج شده هم‌زمان برای هر مورد	نسبت شانس به روش مانتل-هانزل آزمون کای-دو طبقه‌بندی بر اساس گروه‌های همسان شده	تحلیل به روش طبقه‌بندی رگرسیون لجستیک بر اساس کنترل هم‌زمان عوامل همسان شده و مخدوش‌کننده

منابع

1. Pooling Project Research Group. Relation of blood pressure, serum cholesterol, smoking habit, relative weigh and ECG abnormality to incidence of major coronary events: final report of the pooling project. *J. Chronic. Dis* , 1978. 3, 201-306.
2. Krikwood BR, Sterne JAC. *Medical statistics* (2nd ed.). Blackwell Sciences, Massachusetts. 2003.
3. Cook RD, Weisberg S. *Applied regression Including computing and graphics*. Wiley, New York, 1999.
4. Mc Mahon et al. Risk factors of unchocerciasis in Sierra-Leone. *Trans Roy Soc Trop Med Hyg*, 1988. 82:595-600.
5. Christensen E, Neuberger J, Crowe J, Altman DG, Opper H, Portmann B, et al. Beneficial effect of azathioprine and prediction of prognosis in primary biniary cirrhosis. Final results of an international trial. *Gastroenterology* 1985, 89:1048-91.
6. Draper NR, Smith H. *Applied regression analysis* (3rd ed.). Wiley, New York 1998.
7. Atkinson AC. *Plots, transformations and regression*. Clarendon Press, Oxford, 1985.
8. Seber GAF, Wild CJ. *Nonlinear regression*. Wiley, Chichester. 1989.
9. Goldestein H. *Multilevel statistical models* (2nd ed.). Arnold, London 1995.
10. Agresti A. *Categorical data analysis* (2nd ed.). Wiley, New York, 2001.
11. Mc Cullagh P, Nelder JA. *Generalized linear models* (2nd ed.). Chapman & Hall/CRC. London 1989.
۱۲. محبی م ر. کاربرد مدل‌های در برگیرنده اثر تصادفی برای تعدیل ناهمگنی به وجود آمده بر اثر عدم مشاهده برخی متغیرهای کمکی در تحلیل بقا. پایان‌نامه کارشناسی ارشد رشته آمار زیستی دانشگاه تربیت مدرس ۱۳۸۲،
۱۳. ثناگو م. به‌کارگیری زمان بقا شتابنده به‌عنوان یک جایگاه مناسب وقتی فرضیه متناسب بودن برقرار نیست. پایان‌نامه کارشناسی ارشد رشته آمار زیستی دانشگاه تربیت مدرس، ۱۳۸۱
۱۴. یزدانی ج و محبی م ر. مدل‌بندی داده‌های بقا با مشاهدات سانسور شده فاصله‌ای. پایان‌نامه کارشناسی ارشد رشته آمار زیستی دانشگاه تربیت مدرس، ۱۳۸۲
۱۵. یاقوتی م، محبی م ر. مدل مرگ‌ومیر جمع‌پذیر و کاربرد آن در ارزیابی میزان بقا بیماران مبتلا به سارکوم

- بافت نرم اندام. پایان نامه کارشناسی ارشد رشته آمار زیستی دانشگاه تربیت مدرس، ۱۳۸۲.
۱۶. ثناگو م، جوادی م ع، کاظم‌نژاد ا، برادران رفیعی ع ر، محبی م ر. بررسی نتایج پیوند قرنیه اطفال بیمارستان لبافی‌نژاد طی سال‌های ۷۳-۱۳۶۵ مجله چشم‌پزشکی ایران، جلد ۱۶، شماره ۳، ۲۱-۱۰. ۱۳۸۲.
17. Kalbfleisch JD, Prentice RL. The statistical analysis of failure time data (2nd ed.). Wiley, New Jersey, 2002.
18. Klein JP, Mueschberger ML. Survival analysis: Techniques for censored and truncated data. Springer, New York, 1997.
۱۹. اکبری م ا، میرزایی ح، سوری ح. میزان بقاء ۵ ساله در سرطان پستان در بیماران مراجعه‌کننده به بیمارستان شهدا و جرجانی. مجله پژوهشی حکیم. تابستان ۱۳۸۵ دوره ۹، شماره ۲، ۴۴-۳۹.



شناسایی و کنترل تورش در مطالعات پزشکی

تورش یا سوگرایی را می‌توان نتیجه وجود خطاهای منظم در طراحی یا اجرای مطالعه تعریف کرد. خطاهای منظم می‌توانند از روش انتخاب افراد تحت مطالعه، نحوه جمع‌آوری اطلاعات درباره میزان مواجهه یا پیامد موردنظر، روش اندازه‌گیری پارامترها یا نحوه ارزیابی افراد تحت مطالعه نشئت گرفته باشند. همان‌طور که در ادامه این بخش به آن پرداخته می‌شود انواع بسیار زیادی تورش وجود دارد که می‌توانند نتایج مطالعه را تحت تأثیر قرار دهند. یکی از موضوع‌های مهم، تشخیص تورش از خطاهای تصادفی (خطاهای نمونه‌ای) است. خطاهای تصادفی به علت کم بودن حجم نمونه نسبت به کل جامعه برای برآورد پارامتر موردنظر اتفاق می‌افتد. نکته قابل‌توجه در مورد تورش‌ها این است که اتفاق افتادن آن‌ها به علت نوع مطالعه انتخاب‌شده نیست، بلکه به نحوه طراحی و اجرای مطالعه وابسته است. اگر طرح و اجرای مطالعه‌ای بدون تورش باشد، مطالعه را معتبر می‌خوانند زیرا نتایج (میانگین نتایج) درست هستند. مطالعه‌ای که دارای طراحی یا اجرای ناقص باشد، دارای تورش (نامعتبر) خوانده می‌شود. با این حال با توجه به وجود تغییرات نمونه‌ای، مطالعه‌ای همراه با تورش ممکن است به نتایجی که به حقیقت بسیار نزدیک است ختم شود.

هنگامی که تورش وجود دارد، اگر به‌طور فرضی تعداد نامحدودی مطالعه یکسان انجام شود، نتایج آن به‌طور متوسط دور از واقعیت هستند. برای مثال هنگامی که میانگین نسبت‌های خطر در تعداد زیادی (ازلحاظ نظری نامحدود) مطالعه مورد-شاهدی برابر ۲ باشد درحالی‌که واقعاً هیچ وابستگی وجود نداشته باشد، روش‌های به‌کاررفته دارای تورش هستند. این تعریف از تورش هیچ کاربردی در مطالعات عملی که محققان مایل به ارزیابی نتایج آن هستند ندارد. به همین علت باید تورش را به‌وسیله ارزیابی دقیق و نقادانه طرح مطالعه، روش‌ها و برنامه‌های به‌کاررفته در مطالعه تشخیص داد. پیشگیری و کنترل تورش شامل دو مرحله است.

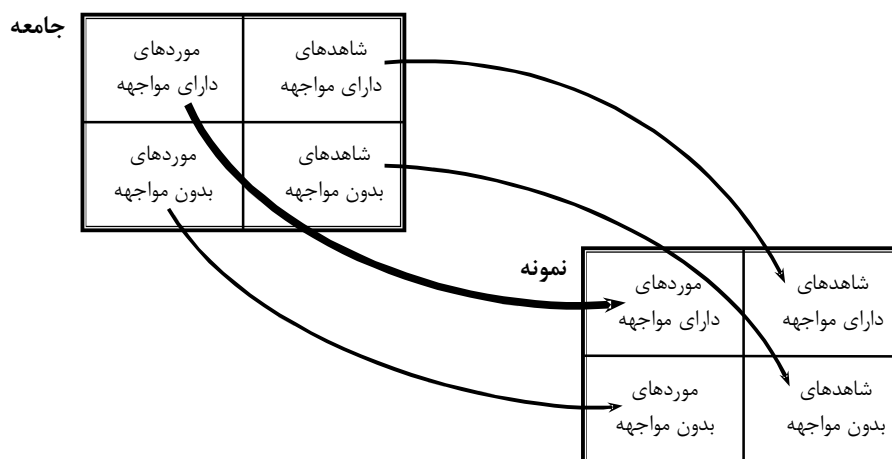
- ♦ اطمینان از مناسب بودن طرح مطالعه بر اساس اهداف و فرضیه‌های موردنظر
- ♦ نظارت و کنترل دقیق مراحل اجرایی شامل انجام اندازه‌گیری‌ها، جمع‌آوری و تحلیل داده‌ها

۹-۱ تورش‌های مرتبط با مطالعات مورد-شاهدی و هم‌گروهی

به‌طور کلی تقریباً تمام تورش‌هایی که در مطالعات با آن‌ها برخورد می‌کنیم در دو رده تورش انتخاب و تورش اطلاع طبقه‌بندی می‌شوند.

۹-۱-۱ تورش انتخاب

این تورش هنگامی رخ می‌دهد که افراد تحت مطالعه با توجه به پارامتر تحت بررسی، احتمال‌های انتخاب متفاوت داشته باشند. برای مثال همان‌طور که شکل ۹-۱ دیده می‌شود موردهای دارای مواجهه، احتمال بیشتری برای انتخاب شدن دارند. یکی از نمونه‌های بارز این نوع سوگرایی به تورش پایش معروف است. برای مثال هنگامی که یک مطالعه مورد-شاهدی برای ارزیابی رابطه استفاده از قرص‌های خوراکی ضدبارداری (OCP) و هر بیماری که دارای علائم مهم بالینی نظیر دیابت است، طراحی می‌شود، هرگونه علائم بالینی ناشی از بیماری در زنان مصرف‌کننده OCP سریع‌تر و در میانگین فراوان‌تر از بقیه تشخیص داده می‌شود، زیرا این اشخاص دفعات بیشتری تحت معاینه پزشک واقع می‌شوند. این موضوع سبب می‌شود که احتمالاً رابطه‌ای محکم بین مصرف OCP و دیابت حاصل شود.



شکل ۹-۱: تورش انتخاب: یکی از گروه‌های موردبررسی در مطالعه (در این شکل موردهای دارای مواجهه) احتمال بیشتری برای انتخاب و ورود به مطالعه دارا هستند.

تورش انتخاب هنگامی رخ می‌دهد که تشخیص وضعیت افراد تحت مطالعه به نحوی باشد که به‌طور منظم باعث تحریف (برآورد کمتر یا بیشتر از واقعیت) رابطه بین مواجهه و پیشامد موردبررسی شود. در مطالعات مورد-شاهدی نحوه تشخیص موردها و شاهد‌ها و در مطالعات هم‌گروهی وجود یا عدم وجود مواجهه از منابع وقوع این تورش هستند. تورش انتخاب مخصوصاً در مطالعات مورد-شاهدی مربوط به انتخاب بیماران بستری شده در بیمارستان به تورش برکسون معروف است.

جداول ۹-۱ تا ۹-۴ به بررسی نحوه تأثیر این تورش در یک مطالعه مورد-شاهدی فرضی اختصاص دارد.

جدول ۹-۱: مطالعه مورد- شاهدهی فرضی که تمام موردها و شاهدهای واجد شرایط بررسی شده‌اند (تورش اطلاع و عامل مخدوش کننده در این مطالعه وجود ندارد).

شاهدها	موردها	عامل خطر	کل جامعه
۱۸۰۰	۵۰۰	وجود دارد	
۷۲۰۰	۵۰۰	وجود ندارد	
۹۰۰۰	۱۰۰۰	مجموع	
$\frac{۱۸۰۰}{۷۲۰۰} = ۱/۴$	$\frac{۵۰۰}{۵۰۰} = ۱$	شانس مواجهه	
نسبت شانسی			
$\frac{(\frac{۵۰۰}{۵۰۰})}{(\frac{۱۸۰۰}{۷۲۰۰})} = ۴$			

جدول ۹-۲: مطالعه مورد- شاهدهی فرضی که بر مبنای نمونه‌ای بدون تورش از ۵۰٪ موردها و ۱۰٪ از شاهدهای جامعه معرفی شده در جدول ۹-۱ انتخاب شده است.

شاهدها ۱۰٪	موردها ۵۰٪	عامل خطر	نمونه
۱۸۰	۲۵۰	وجود دارد	
۷۲۰	۲۵۰	وجود ندارد	
$۹,۰۰۰ \times ۰/۱ = ۹۰۰$	$۱,۰۰۰ \times ۰/۵ = ۵۰۰$	مجموع	
$\frac{۱۸۰}{۷۲۰} = ۰/۲۵$	$\frac{۲۵۰}{۲۵۰} = ۱$	شانس مواجهه	
نسبت شانسی			
$\frac{(\frac{۲۵۰}{۲۵۰})}{(\frac{۱۸۰}{۷۲۰})} = ۴$			

استنباط: شانسی مواجهه و نسبت شانسی مواجهه، بدون تورش هستند.

جدول ۹-۳: مطالعه مورد- شاهدهی با وجود تورش انتخاب. در این مطالعه از ۵۰٪ موردها و ۱۰٪ از شاهدهای جامعه نمونه گرفته شده است.

شاهدها	موردها	عامل خطر	نمونه
۱۸۰	$۵۰۰ \times ۰/۶ = ۳۰۰$	وجود دارد	
۷۲۰	$۵۰۰ \times ۰/۴ = ۲۰۰$	وجود ندارد	
$۹,۰۰۰ \times ۰/۱ = ۹۰۰$	$۱,۰۰۰ \times ۰/۵ = ۵۰۰$	مجموع	
$\frac{۱۸۰}{۷۲۰} = ۰/۲۵$	$\frac{۳۰۰}{۲۰۰} = ۱/۵$	شانس مواجهه	
نسبت شانسی			
$\frac{(\frac{۳۰۰}{۲۰۰})}{(\frac{۱۸۰}{۷۲۰})} = ۶$			

استنباط: شانسی مواجهه در موردها دارای تورش و در شاهدها بدون تورش است. در مجموع شانسی مواجهه نسبت شانسی دارای تورش هستند.

جدول ۹-۴: مطالعه مورد-شاهدی با وجود تورش انتخاب نمونه‌ای از ۵۰٪ موردها و ۱۰٪ از شاهدها جامعه معرفی شده در جدول ۹-۱ گرفته شده است. در این مطالعه موردها و شاهدها دارای تورش انتخاب یکسان هستند.

شاهد‌ها	موردها	عامل خطر	نمونه
$1,800 \times 0/136 = 245$	$500 \times 0/6 = 300$	وجود دارد	
$7,200 \times 0/091 = 655$	$500 \times 0/4 = 200$	وجود ندارد	
$9,000 \times 0/1 = 900$	$1,000 \times 0/5 = 500$	مجموع	
$\frac{245}{655} = 0/374$	$\frac{300}{200} = 1/5$	شانس مواجهه	
$\frac{\left(\frac{300}{200}\right)}{\left(\frac{245}{655}\right)} = 4$		نسبت شانس	

استنباط: شانس‌های مواجهه‌ای دو گروه مورد و شاهد تورش دارد ولی نسبت شانس بدون تورش باقی می‌ماند.

در جدول ۹-۱، تمام موردها و شاهد‌های واقعی موجود در جمعیت مرجع ۱۰۰۰۰ نفری در مطالعه مورد-شاهدی بررسی شده‌اند. به همین علت جدول ۹-۱، نتایج واقعی که می‌تواند به‌عنوان مرجع استاندارد مورد استفاده واقع شود را ارائه می‌کند.

در جدول ۹-۲، نمونه‌ای بدون تورش انتخاب، از جدول مرجع گرفته شده است. نمونه گرفته شده ترکیبی از ۵۰٪ موردها و ۱۰٪ شاهد‌های کل جامعه است. نمونه‌گیری بدون تورش باعث می‌شود که شانس مواجهه و نسبت شانس محاسبه شده برای ارزیابی عامل خطر بر بیماری بدون تورش باشد. جدول‌های ۹-۳ و ۹-۴، نحوه اثر تورش انتخاب بر رابطه عامل خطر و بیماری را نمایش می‌دهد.

در جدول ۹-۳، به علت بی‌توجهی محقق در انتخاب موردها مطالعه دارای تورش است. در این مثال نیز مانند مثال جدول ۹-۲، محقق نمونه‌ای شامل ۵۰٪ از موردها انتخاب کرده است؛ اما نسبت موردهای دارای مواجهه بیشتر از موردهای بدون مواجهه است. در مثال جدول ۹-۴، تورش انتخاب به‌طور یکسان در موردها و شاهد‌ها وجود دارد. به همین علت شانس‌های مواجهه در موردها و شاهد‌ها دارای تورش است، اما نسبت شانس بدون تورش باقی می‌ماند.

یکی از روش‌های حذف اثر تورش ناشی از اختلاف بین گروه‌ها تلاش برای پخش شدن مساوی تورش بین گروه‌های تحت بررسی است. برای مثال در مطالعات مورد-شاهدی گذشته‌نگر باید شرایط انتخاب موردها و شاهد‌ها کاملاً مشابه باشد. در مطالعات مورد-شاهدی بهترین روش برای انتخاب گروه شاهد، استفاده از شاهد‌هایی است که در یک برنامه غربالگری شرکت کرده‌اند. برای مثال اگر خانم‌هایی که در برنامه غربالگری مربوط به سرطان پستان شرکت کرده‌اند و مبتلا تشخیص داده شده‌اند (موردها) با آن‌ها که سالم تشخیص داده شده‌اند (شاهد‌ها) مقایسه شوند، به علت آن‌که شاهد‌ها نیز به اندازه موردها در معرض عوامل خطری نظیر سابقه

فامیلی مثبت، سن بالا و سایر عوامل خطر شناخته شده قرار داشته‌اند، تورش کمتری نسبت به حالتی که گروه شاهد را از افراد عادی جامعه انتخاب کنیم وجود دارد. مثال فرضی که در جدول ۹-۴، نمایش داده شد، بیانگر همین موضوع است. در این مثال نیز تورش در دو گروه مورد و شاهد یکسان است. به همین دلیل با این که شانس‌های مواجهه دارای تورش است نسبت شانس بدون تورش باقی می‌ماند. در این مثال اندازه تورش انتخاب در موردها و شاهد‌ها یکسان است. این موضوع باعث جبران سوگرایی می‌شود [۱].

$$\text{تورش} = \frac{\text{شانس انتخاب موردها}}{\text{شانس واقعی انتخاب موردها}} = \frac{1/5}{1} = \frac{\text{شانس انتخاب شاهد‌ها}}{\text{شانس واقعی انتخاب شاهد‌ها}} = \frac{1}{2/67} = \frac{1}{4} = 1/5$$

برای جبران سوگرایی متعادل باید منابع تورش یکسانی در هر دو گروه مورد مقایسه وجود داشته باشد. در این صورت می‌توان به صورت زیر استدلال کرد.

$$\text{نسبت شانس واقعی (بدون تورش)} = \frac{\text{شانس مواجهه برای موردها}}{\text{شانس مواجهه برای شاهد‌ها}} = \frac{\text{[تورش]} \times \text{شانس مواجهه برای موردها}}{\text{[تورش]} \times \text{شانس مواجهه برای شاهد‌ها}}$$

در مثال جدول ۹-۴

$$\text{نسبت شانس} = \frac{\left(\frac{1}{5}\right) \times 1/5}{\left(\frac{1}{4}\right) \times 1/5} = \frac{\left(\frac{1/5}{1}\right)}{\left(\frac{1}{2/67}\right)} = 4$$

در عمل کسب اطمینان از جبران سوگرایی در گروه‌های مورد مقایسه بسیار مشکل است. برای مثال در مطالعه‌ای که برای ارزیابی رابطه بین مصرف قهوه و سرطان لوزالمعده انجام گرفت به منظور تعدیل تورش در گروه موردها (مبتلایان به سرطان لوزالمعده) و شاهد‌ها (سایر بیماران همان بیمارستان)، تیم پزشکی یکسانی میزان مواجهه یعنی مصرف قهوه را از آن‌ها سؤال می‌کردند. باین حال نکته بسیار مهم در این مطالعه، عدم توجه به این موضوع بود که بیماران مبتلا به سرطان لوزالمعده معمولاً دچار سوءهاضمه هستند و پزشکان معالج به آن‌ها توصیه می‌کنند کمتر قهوه مصرف کنند. به همین علت تلاش محققان برای جبران سوگرایی به گروه شاهدی منجر شد که بسیار کمتر از میانگین واقعی جامعه قهوه مصرف می‌کردند [۲].

این موضوع باعث مشاهده یک رابطه مثبت و قوی بین مصرف قهوه و سرطان لوزالمعده شد. رابطه مشاهده‌شده در مطالعات بعدی تأیید نشد [۳].

تمام مثال‌های ذکرشده در این قسمت مربوط به مطالعات مورد-شاهدی بودند، زیرا نتایج این نوع مطالعه و مطالعه مقطعی نسبت به تورش انتخاب بسیار حساس است. در مطالعات هم‌گروهی پیامد پس از انتخاب افراد تحت مطالعه یعنی گروه‌های دارای مواجهه و بدون مواجهه و در کارآزمایی‌های بالینی پیامد پس از تصادفی‌سازی

بیماران ایجاد می‌شود. به همین علت به نظر نمی‌رسد نحوه انتخاب اعضای نمونه تأثیر زیادی در روابط مشاهده شده داشته باشد.

در این مطالعات نکته مهم ایجاد تورش ناشی از موارد گم‌شده یا دارای پیگیری کوتاه‌مدت در برخی از گروه‌های تحت مطالعه است [۴].

۹-۱-۱-۱-۱-۱ تورش ارجاع به مرکز

شهرت یا خوش‌نامی یک مرکز درمانی ممکن است باعث مراجعه بیماران خاص و یا موارد وخیم‌تر آن بیماری به آن مرکز و در نتیجه ایجاد تورش شود. برای مثال بیشتر موارد وخیم بیماری در شهرستان‌ها به خاطر عدم وجود امکانات لازم و یا محوریت یک مرکز درمانی خاص در تهران به آن مرکز مراجعه می‌کنند. در صورتی که تنها بیماران مراجعه‌کننده به آن مرکز مورد بررسی قرار گیرند ممکن است برآورد شدت بیماری در مراجعین بزرگنمایی شود. چنین مشکلی در بیماران مراجعه‌کننده به مراکز ارجاعی نیز ممکن است روی دهد. دسترسی به مراکز تشخیص بیماری و یا اطلاع از برخی مراکز خاص تشخیص نیز ممکن است باعث تورش شود.

۹-۱-۱-۲-۱-۱ تورش داوطلب‌ها

داوطلب‌ها به‌طور معمول از افراد عادی جامعه سالم‌تر هستند، بیشتر ورزش می‌کنند و سیگار و قهوه کمتری نیز مصرف می‌کنند. هم‌چنین، احتمال دارد این افراد به دلایل خاصی (مثلاً نگرانی از بیماری سلامتی خود) دقیق‌تر و یا با ذکر جزئیات بیشتری اثر درمان یا پیامدهای درمان را ابراز دارند. آن‌ها علاقه‌مند به مشارکت در مطالعه هستند و اطلاعات به دست آمده از آن‌ها قابل‌تعمیم به جمعیت عمومی نیست.

۹-۱-۲-۱-۱-۲ تورش اطلاع

تورش اطلاع از تعریف ناقص متغیرهای تحقیق و یا نقص و کاستی در فرایند جمع‌آوری داده‌ها، ایجاد می‌شود. این خطا از طبقه‌بندی ناصحیح میزان مواجهه یا پیامد موردنظر در بخش‌های معینی از افراد تحت مطالعه ناشی می‌شود. جمع‌آوری داده درباره میزان مواجهه در افراد تحت مطالعه یکی از منابع ایجاد این تورش است. این مسئله عمدتاً در مطالعه مورد-شاهدی حادث می‌شود. به همین علت بیشتر مثال‌ها در این زمینه مربوط به مطالعات مورد-شاهدی است. دوره مهم از این نوع تورش، تورش خاطر (حافظه) و تورش مصاحبه‌گر است.

۹-۱-۲-۱-۲-۱-۲ تورش به خاطر آوردن

تورش به خاطر آوردن به علت یادآوری نادرست و یا ناقص مواجهه‌های گذشته رخ می‌دهد این تورش یکی از مهم‌ترین انواع تورش در اندازه‌گیری مواجهه است. به‌طور خاص در مطالعات مورد-شاهدی هنگامی که از افراد تحت مطالعه خواسته می‌شود مواجهه‌های قبلی خود را بازگو کنند، خطا در یادآوری وقایع قبلی باعث طبقه‌بندی نادرست سطوح مواجهه می‌شود. هنگامی که نحوه یا کیفیت تورش خاطره در مورد‌ها و شاهد‌ها، مانند هم باشد، تورش بدون جهت و هنگامی که نحوه یا کیفیت این تورش در مورد‌ها و شاهد‌ها متفاوت است به آن تورش

جهت‌دار گفته می‌شود. تأثیر تورش خاطره به‌طور تجربی در مطالعه هم‌گروهی «سلامت پرستاران» مورد بررسی قرار گرفت [۴]. در این مطالعه از اعضای شرکت‌کننده درخواست شد تا قابلیت برنزه شدن پوست خود را در ابتدای مطالعه و در طول پیگیری در گروه موردها، که مبتلا به سرطان پوست^{۲۱} شده بودند و افراد بدون عارضه (شاهدها) ثبت کنند. نتایج، نشان‌دهنده گزارش بیش از مقدار حقیقی توانایی کم برنزه شدن پوست در موردها نسبت به زمان قبل از بیماری بود. درحالی‌که اطلاعات مربوط به گروه شاهد در طول زمان ثابت باقی ماند.

۹-۱-۲-۲ تورش مصاحبه‌گر

هنگامی که وضعیت بیمار یا سالم بودن افراد تحت مطالعه در یک مطالعه مورد-شاهدی برای محققان مسئول جمع‌آوری اطلاعات معلوم باشد. تورش‌های مربوط به مشاهده نظیر تورش مصاحبه‌گر احتمال وقوع زیادی دارند. یکی از راه‌های ایجاد این نوع تورش هنگامی است که مصاحبه‌کننده در هنگام مصاحبه با موردها آن‌ها را ترغیب به یادآوری مواجهه‌های گذشته می‌کند و با این کار قواعد مصاحبه که در دستورالعمل تحقیق ذکر شده است را نقص می‌نماید.

نحوه ارزیابی و تشخیص این تورش بسیار مشکل است. بهترین روش برای پیش‌گیری از آن فراهم آوردن تدابیر ویژه در مرحله طراحی مطالعه است. اطمینان از بالا بودن کیفیت تحقیق، فراهم آوردن جزئیات کاملی از دستورالعمل‌های اجرایی، آموزش افراد تحت مطالعه، استاندارد کردن نحوه جمع‌آوری داده‌ها و نظارت بر جمع‌آوری داده‌ها روش‌های کلی جلوگیری از این تورش‌ها است. یکی از روش‌های عملی برای حذف این تورش، کور کردن وضعیت بیماری برای تیم مسئول جمع‌آوری اطلاعات است. بسیاری اوقات استفاده از این روش غیرممکن و یا بسیار مشکل است. در این مواقع استفاده از یک پرسشنامه وسیع الطیف که شامل سؤال‌هایی راجع به وضعیت سلامت موردها و شاهدها است می‌تواند تا حد زیادی سؤالات اصلی تحقیق را مخفی نگه دارد. این موضوع در مطالعات روانشناسی کاربرد بسیار دارد.

تورش‌های مرتبط با اندازه‌گیری پیامد در مطالعات مورد-شاهدی و هم‌گروهی احتمال وقوع دارد. عدم تعریف دقیق و کامل پیامد (بیماری) موردنظر و یا خطا در مرحله جمع‌آوری داده‌ها باعث بروز این تورش می‌شود.

۹-۱-۲-۳ تورش مشاهده‌گر

در مطالعات هم‌گروهی، امکان دارد تصمیم‌گیری راجع به وقوع پیامد مورد بررسی، تحت تأثیر آگاهی از وضعیت مواجهه افراد تحت مطالعه قرار بگیرد. این مسئله هنگامی که تشخیص پیامد موردنظر تا حدی مبهم باشد، نظیر حمله‌های میگرنی، کم شدن شنوایی یا ایجاد بیماری روانی بسیار جدی‌تر است. با این حال وقوع تورش مشاهده‌گر، در مراحل مختلف تشخیص پیامد که شامل تشخیص‌های پاتولوژیک و بالینی می‌شود نیز محتمل است. مثلاً برای تشخیص سیروز کبدی به علت مصرف الکل به کمک آزمایش بافت‌شناسی نمونه‌ها، اگر پاتولوژیست از اعتیاد یا عدم اعتیاد افراد تحت مطالعه به الکل مطلع باشد، این موضوع یکی از منابع بالقوه بروز تورش مشاهده‌گر تلقی می‌شود. یکی از مثال‌های مستند در این زمینه مطالعه‌ای است که اندازه‌گیری تورش مشاهده‌گر،

به علت اطلاع از نژاد بیماران در تشخیص بیماری کلیوی پیشرفته ناشی از فشارخون بالا (HSRD)^{۳۳} بررسی شده است. در این مطالعه سابقه پزشکی هفت بیمار به گروهی از متخصصان کلیه ارجاع شد. در این پرونده‌ها نژاد بیماران به تصادف سیاه یا سفید ذکر شده بود. در نهایت شانس تشخیص بیماری در افراد سیاه‌پوست دو برابر شانس تشخیص در بیماران سفیدپوست مشاهده شد [۶].

همان‌طور که در این قسمت اشاره شد هنگامی که تشخیص بیماری و اطلاع از وضعیت مواجهه مستقل نباشد تورش مشاهده‌گر رخ می‌دهد. به همین علت از لحاظ نظری اگر مشاهده‌گر از سابقه مواجهه بی‌اطلاع (کور) باشد، تورش مشاهده‌گر ایجاد نخواهد شد. در مواقعی که کور کردن امکان‌پذیر نباشد، استفاده از دو یا چند مشاهده‌گر برای تشخیص یکی از راهبردهای مناسب برای کاهش این تورش است.

۹-۱-۲-۴ تورش پاسخگویی

تورش خاطره و سایر تورش‌های اطلاع معمولاً مرتبط با اندازه‌گیری مواجهه در مطالعات مورد-شاهدی هستند؛ اما تورش در تشخیص و گزارش پیامد یا بیماری موردنظر معمولاً در طول پیگیری افراد تحت مطالعه (مثلاً یک مطالعه هم‌گروهی) رخ می‌دهد. علاوه بر احتمال بروز تورش مشاهده‌گر، امکان بروز تورش دیگری ناشی از نحوه پاسخگویی افراد تحت مطالعه است، وجود دارد. برای مثال هنگام جمع‌آوری داده‌ها مبتنی بر پاسخ‌های اظهار شده توسط افراد تحت مطالعه مانند دفعات وقوع حمله‌های صرع، تورش پاسخگویی بسیار بااهمیت است.

در صورت امکان، اطلاعات اظهار شده به وسیله افراد تحت مطالعه باید به وسیله ابزارهای مستند نظیر پرونده‌های بیمارستانی کنترل شوند. هرگاه که چنین امکانی فراهم نباشد، مانند مواقعی که جمع‌آوری اطلاعات درباره وقایعی که در زمان‌هایی به‌جز هنگام بستری بودن بیمار اتفاق می‌افتاد و یا امکان ارزیابی آزمایشگاهی آن غیرممکن است، یکی از روش‌های کاهش تورش پاسخگویی جمع‌آوری اطلاعات دقیق و جزئی درباره نشانه‌ها، علائم و عوارض احتمالی پیامد موردنظر است. این اطلاعات کمک مؤثری در ارزیابی صحت اطلاعات اظهار شده هستند. برای مثال در مطالعه‌ای که اندازه‌گیری وقوع سردردهای میگرنی بخشی از آن بود، به‌جز پرسش درباره وجود سردردهای بسیار شدید، پرسش‌هایی درباره وجود سرگیجه، کوفتگی و تهوع به همراه سر درد نیز از افراد تحت مطالعه پرسیده شد [۷]. ارزیابی وقوع برخی از عارضه‌ها نظیر آنژین صدری و مننژیت مزمن دارای پرسشنامه‌های استاندارد شده هستند [۸-۱۱].

۹-۱-۲-۵ تورش پایش (یا تشخیص) بیماران

این تورش هنگامی رخ می‌دهد که زیرگروه‌هایی از جامعه به علت وجود علامت یا پیش‌آگهی خاصی بیشتر مورد توجه تیم پزشکی قرار می‌گیرند. به همین علت آن‌ها بیشتر از سایر افراد جامعه معاینه شده و مورد آزمایش‌های پاراکلینیکی قرار می‌گیرند یا پیگیری آن‌ها طولانی و دقیق‌تر صورت می‌گیرد. این موضوع باعث تشخیص سریع‌تر و دقیق‌تر برخی عارضه‌ها در این زیرگروه‌ها می‌شود این تورش در مورد بیماری‌ها و یا مواجهه‌هایی نظیر دیابت، فشارخون و مصرف قرص‌های خوراکی ضدبارداری (OCP) که منجر به انجام مداوم و

۱. Hypertensive end-stage renal disease

دقیق برخی معاینات و آزمایش‌ها می‌شود، بسیار قابل توجه است. برای مثال به علت این که یکی از عوارض استفاده از OCP از کارافتادن کلیه است اگر خانم‌هایی که OCP استفاده می‌کنند به‌طور مکرر معاینه شوند و کراتینین آن‌ها اندازه‌گیری شود، خطر استفاده از OCP در ابتدا به بیماری‌های کلیوی بسیار اغراق‌آمیزتر از مقدار واقعی مشاهده می‌شود.

انجام آزمایش‌ها و معاینات مشابه در تمام زیرگروه‌های جامعه، بدون توجه به وضعیت مواجهه باعث به حداقل رسیدن این تورش در مطالعات می‌شود. به همین علت معاینات یکسان، مستمر و برنامه‌ریزی شده در مطالعات هم‌گروهی بزرگ برای تمام افراد تحت مطالعه انجام می‌شود.

یکی دیگر از روش‌های به حداقل رساندن این تورش در مطالعات هم‌گروهی، مخفی کردن وضعیت مواجهه از تیم مسئول تشخیص پیامد موردبررسی است. مواقعی که این روش قابل اجرا نباشد، مثلاً هنگامی که مطالعه گذشته‌نگر است (نظیر مطالعات مورد-شاهدی) با استفاده از موردهای قدیم باید در مورد زیرگروه‌های بدون مواجهه یا پیش‌آگهی نیز اطلاعات باکیفیت و مستمر درباره سوابق پزشکی، نتایج آزمایش‌ها و معاینات قبلی و پیگیری‌های پزشکی وجود داشته باشد. برای مثال هنگامی که محققان در نظر دارند به‌وسیله یک مطالعه مورد-شاهدی به ارزیابی اثر هورمون درمانی در یک بیماری با علائم کلینیکی مشخص نظیر دیابت نوع II بپردازد، توجه به مواردی نظیر اینکه آیا تمام بیماران دارای بیمه پزشکی هستند یا معاینات پزشکی آن‌ها در گذشته از نظر کیفیت و دفعات انجام مشابه بوده، بسیار مهم و ضروری است. نکته قابل توجه دیگر وضعیت آموزش، تحصیلات و طبقه اجتماعی و اقتصادی بیماران است. زیرا این عوامل معمولاً رابطه بالایی با در دسترس بودن و استفاده از مراقبت‌های پزشکی دارند.

در مجموع تورش پایش باعث می‌شود بیماری یا پیامد در برخی از زیرگروه‌های جامعه که دارای مواجهه یا پیش‌آگهی خاصی هستند، سریع‌تر و بیشتر تشخیص داده شود. به همین علت نسبت عارضه، در مطالعات هم‌گروهی و نسبت شانس در مطالعات مورد-شاهدی در این زیرگروه‌ها بالاتر از سایر گروه‌های موردبررسی می‌شود.

۹-۱-۳ تورش سوء طبقه‌بندی

تورش‌های اطلاع منجر به سوء طبقه‌بندی مواجهه و یا پیامد تحت بررسی می‌شوند. برای مثال هنگامی که تورش خاطره در یک مطالعه مورد-شاهدی ایجاد می‌شود، برخی افراد دارای مواجهه، در گروه بدون مواجهه طبقه‌بندی می‌شوند و برخی افراد بدون مواجهه در گروه دارای مواجهه طبقه‌بندی می‌شوند. این موضوع در مطالعات هم‌گروهی به‌صورت تشخیص نادرست پیامد موردبررسی احتمال وقوع دارد. در چنین مواقعی افراد دارای عارضه، سالم و افرادی که بدون عارضه موردبررسی هستند به‌عنوان بیمار تلقی می‌شوند. به‌طور کلی این تورش دارای دو رده سوء طبقه‌بندی جهت‌دار و سوء طبقه‌بندی بدون جهت است.

سوء طبقه‌بندی بدون جهت هنگامی رخ می‌دهد که میزان سوء طبقه‌بندی نادرست مواجهه مستقل از رده‌های پیامد موردبررسی باشد (یا برعکس).

جدول ۹-۵ مثالی فرضی از نحوه تأثیر سوء طبقه‌بندی بدون جهت را هنگامی که دو رده وجود داشته باشد،

نمایش می‌دهد [۱]. در این مثال اعضای دارای مواجهه به صورت بدون مواجهه طبقه‌بندی شده‌اند. در عمل احتمالاً این خطا در هر دو جهت رخ می‌دهد. این تورش در مطالعات مورد-شاهدی هنگامی رخ می‌دهد که حساسیت و ویژگی آزمون‌ها کمتر از ۱۰۰٪ باشد. به همین دلیل توجه به حساسیت و ویژگی در مطالعات (به ویژه مطالعات غربالگری) می‌تواند بسیار مفید باشد.

سوءطبقه‌بندی جهت‌دار هنگامی رخ می‌دهد که میزان سوءطبقه‌بندی در گروه‌های مورد مقایسه همانند نباشد. سوءطبقه‌بندی بدون جهت به طور کلی باعث کاهش مقدار وابستگی حقیقی بین مواجهه و پیامد می‌شود؛ اما سوءطبقه‌بندی جهت‌دار باعث ایجاد تورش در هر دو جهت می‌شود. نحوه اثر این نوع تورش در جدول ۹-۶ نمایش داده شده است [۱].

جدول ۹-۵: مثال فرضی از اثر سوءطبقه‌بندی بدون جهت برای دو رده دارای مواجهه و بدون مواجهه با میزان خطای ۳۰٪

بدون خطای طبقه‌بندی

مواجهه	موردها	شاهدها
دارد	۵۰	۲۰
ندارد	۵۰	۸۰
نسبت شانس		$\frac{(\frac{50}{50})}{(\frac{20}{80})} = 4$

خطای ۳۰٪ در طبقه‌بندی مواجهه

مواجهه	موردها	شاهدها
دارد	$50 - 15 = 35$	$20 - 6 = 14$
ندارد	$50 + 15 = 65$	$80 + 6 = 86$
نسبت شانس		$\frac{(\frac{35}{65})}{(\frac{14}{86})} = 3/3$

جدول ۹-۶ مثال فرضی از نحوه اثر سوءطریقهبندی جهت‌دار بر تشخیص مواجهه در شاهدها در یک مطالعه مورد-شاهدی هنگامی که نسبت شانس واقعی ۴ است.

نسبت شانس مشاهده شده	شیوع مواجهه در شاهدها	ویژگی	حساسیت
۲/۶	۰/۲	۰/۸۵	۰/۹
۱/۹	۰/۲	۰/۸۵	۰/۶
۳/۲	۰/۲	۰/۹۵	۰/۹
۱/۹	۰/۲	۰/۶	۰/۹
۳/۰	۰/۳۶۸	۰/۹	۰/۹
۲/۸	۰/۲	۰/۹	۰/۹
۲/۲	۰/۰۷۷	۰/۹	۰/۹

شناسایی و میزان اثر تورش سوءطریقهبندی در مطالعات پزشکی بسیار موردتوجه واقع شده است. برخی بررسی‌ها نشان داده است که این تورش حتی می‌تواند باعث تحریف کامل رابطه واقعی بین مواجهه و پیامد موردبررسی شود. در مواقعی که نسبت شانس یا خطر نسبی واقعی نزدیک به ۱ باشد، بروز تورش طبقه‌بندی بدون جهت می‌تواند به‌طور کامل رابطه حقیقی را بی‌اثر جلوه دهد. برای مثال هنگامی که مواجهه دارای شیوع ۱۶٪ و نسبت شانس واقعی ۱/۳ است بروز تورش سوءطریقهبندی بدون جهت می‌تواند نسبت شانس برآورد شده را برابر با ۱ نشان دهد (هنگامی که حساسیت و ویژگی هر دو برابر ۷۰٪ باشد). سوءطریقهبندی جهت‌دار باعث ایجاد تورش در هرکدام از جهت‌ها می‌شود. معمولاً داده‌هایی برای مقایسه و ارزیابی طبقه‌بندی‌ها در دسترس نیست؛ اما هنگامی که چنین موقعیتی وجود داشته باشد مقایسه‌ها بسیار سودمند خواهد بود. باین‌حال مقایسه نمونه‌ها از لحاظ درستی طبقه‌بندی تنها در مورد متغیر پاسخ کافی نیست و باید در مورد تمام متغیرهای مهم که دارای منابع بالقوه برای بروز این نوع تورش هستند اعمال شود. به‌عنوان مثال، ضبط مصاحبه‌ها با افراد مطالعه روشی سودمند است که بازبینی مجدد پاسخ‌ها را امکان‌پذیر می‌سازد. ارزیابی ابزار اندازه‌گیری، استفاده از ابزارهای عینی اندازه‌گیری مواجهه و پیامد مانند استفاده از دستگاه‌های سونوگرافی و سی‌تی اسکن باعث مطمئن شدن محققان از نتایج می‌شود. در مواقعی که حساسیت و ویژگی ابزارهای اندازه‌گیری از پیش معلوم هستند، امکان محاسبه و گزارش شاخص‌های تعدیل‌یافته برای اندازه‌خطر نسبی (برای مثال نسبت خطر اصلاح‌شده) وجود دارد. روش‌های آماری پیشرفته‌تری برای محاسبه شاخص‌های تعدیل‌یافته هنگامی که ارزیابی پیامد و اندازه‌گیری مواجهه دارای خطای طبقه‌بندی نادرست هستند موردنیاز است.

در این قسمت به بررسی تورش‌هایی می‌پردازیم که حاصل خطای توأم در انتخاب و اطلاع هستند این تورش‌ها در مطالعات مربوط به پایش بیماران، مطالعات مقطعی و مطالعات غربالگری شایع‌تر از سایر مطالعات هستند.

۹-۲ تورش‌های مرتبط با مطالعات مقطعی

تورش‌های مرتبط با مطالعات مقطعی به دو گروه تورش استفاده از بروز به‌جای شیوع و تورش تقدم پیامد بر مواجهه تقسیم می‌شوند. یکی از تورش‌های شایع در مطالعات مقطعی استفاده از برآوردهای شیوع حاصل از این مطالعات برای محاسبه خطر نسبی است. بین شیوع، بروز و طول دوره بیماری رابطه زیر برقرار است.

$$\frac{\text{شیوع}}{\text{شیوع}-۱} = \text{طول دوره بیماری} \times \text{بروز}$$

این رابطه را می‌توان به این صورت نیز نمایش داد.

$$(\text{شیوع} - ۱) \times (\text{طول دوره بیماری} \times \text{بروز} = \text{شیوع})$$

به این ترتیب شیوع به جز بروز و طول دوره بیماری با عبارت (شیوع - ۱) نیز مرتبط است. این موضوع در مطالعات مقطعی باعث می‌شود شیوع نسبی، تابعی از نسبت بروز در گروه مواجهه (که با علامت + نشان داده می‌شود) به بروز در گروه بدون مواجهه (که با علامت - نشان داده می‌شود) و همچنین نسبت طول دوره بیماری گروه مواجهه به گروه بدون مواجهه و نسبت عبارت (شیوع - ۱) در گروه‌های دارای مواجهه و بدون مواجهه باشد. یعنی:

$$\text{شیوع نسبی} = \frac{(\text{شیوع}^+) (\text{طول دوره بیماری}^+)}{(\text{شیوع}^-) (\text{طول دوره بیماری}^-)} \times \frac{\text{بروز}^+}{\text{بروز}^-}$$

نسبت‌های ۲ و ۳ باعث ایجاد دو نوع تورش با نام‌های تورش نسبت دوره بیماری و تورش نسبت مکمل شیوع می‌شود.

۹-۲-۱ تورش دوره بیماری

این تورش نوع خاصی از تورش انتخاب است. هنگامی که شیوع نسبی حاصل از یک مطالعه مقطعی به عنوان شاخصی برای نشان دادن میزان وابستگی بین مواجهه و بیماری به کار می‌رود، درحالی‌که دوره بیماری در دو گروه دارای مواجهه و بدون مواجهه همانند نیست، تورش طول دوره بیماری اتفاق می‌افتد. به علت این که طول دوره بیماری در بیماری‌های مزمن غالباً با مدت بقای بیماران مرتبط است. این تورش را گاهی اوقات تورش بقا می‌خوانند. هنگامی که بیماری موردبررسی نادر باشد و طول دوره بیماری در هر دو گروه یکسان باشد، شیوع نسبی تقریباً برآوردی بدون تورش از میزان بروز نسبی است؛ اما در مواقعی که مواجهه، طول دوره بیماری را تغییر می‌دهد نتایج دارای تورش است.

۹-۲-۲ تورش نسبت مکمل شیوع

هنگامی که طول دوره بیماری از مواجهه مستقل باشد، میزان شیوع نسبی باعث برآورد کمتر از مقدار واقعی رابطه میان مواجهه و پیامد موردبررسی می‌شود (باعث می‌شود برآورد میزان شیوع نسبی به سمت ۱ متمایل شود). اندازه این تورش به اندازه شیوع نسبی و شیوع در دو گروه دارای مواجهه و بدون مواجهه بستگی دارد. هنگامی که شیوع در گروه مواجهه بیشتر باشد، نسبت مکمل شیوع کمتر از ۱ است. این نسبت هنگامی که شیوع در هر دو گروه نادر است، نزدیک ۱ می‌شود. برای مثال اگر شیوع یک بیماری در گروه دارای مواجهه ۰/۰۴ و در گروه بدون مواجهه ۰/۰۱ باشد، شیوع نسبی برابر $\frac{۰/۰۴}{۰/۰۱} = ۴$ است و تورش ناشی از نسبت مکمل شیوع قابل صرف‌نظر

است ($۱ \cong ۰/۹۷ = \frac{۰/۹۶}{۰/۹۹}$). در مقابل هنگامی که شیوع در گروه دارای مواجهه نسبتاً بالا باشد، نسبت مکمل شیوع و

تورش ناشی از آن مهم خواهد بود. برای مثال اگر شیوع بیماری در گروه مواجهه ۰/۴ و در گروه بدون مواجهه برابر ۰/۱ باشد، شیوع نسبی مانند مثال قبل است، اما نسبت مکمل شیوع برابر $\frac{۰/۶}{۰/۹} = ۰/۶۷$ می‌شود. در این صورت برآورد بروز نسبی به‌وسیله شیوع نسبی باعث می‌شود بروز نسبی ۳۳٪ کمتر از مقدار واقعی برآورد شود.

۹-۲-۳ تورش تقدم پیامد بر مواجهه

این نوع تورش هنگامی رخ می‌دهد که توالی زمانی لازم برای برقرار شدن رابطه سببی در مطالعه در نظر گرفته نشود. این نوع تورش بیشتر در مطالعاتی که مواجهه و پیامد هم‌زمان جمع‌آوری شده‌اند رخ می‌دهد. برای مثال از مشاهده یک رابطه تصنعی بین پرفشاری خون و بالا بودن کراتینین امکان دارد استنباط شود فشارخون بالا باعث افزایش کراتینین می‌شود یا برعکس؛ اما انجام یک مطالعه آینده‌نگر بر روی افرادی با فشارخون بالا با سطح کراتینین عادی نشان داد فشارخون بالا یکی از عوامل پیش‌بینی‌کننده سطح کراتینین و در نهایت بیماری کلیوی است [۱].

۹-۳ تورش‌های متداول در کارآزمایی‌های بالینی

تورش یک کارآزمایی می‌تواند در مراحل برنامه‌ریزی، انتخاب شرکت‌کنندگان، اجرای مداخلات، سنجش پیامدها و تجزیه و تحلیل داده‌ها، تفسیر و گزارش نتایج و انتشار نتایج مطالعه اتفاق بیفتد. تورش ممکن است حتی زمانی که فردی گزارش یک کارآزمایی را می‌خواند رخ دهد.

دلیل اصلی برای بررسی، آشکار کردن، سنجیدن و کنترل تورش، ناشناخته بودن اثرات واقعی هرگونه مداخله است. در واقع، هدف کلی، فراهم آوردن نتایج به نحوی است که قابل تعمیم به کل جمعیت هدف باشد. اطمینان از وجود یا عدم تورش در نتایج یک تحقیق غیرممکن است، زیرا اثبات انحراف منظم نتایج تحقیق از واقعیت ناشناخته غیرممکن است. تعداد زیادی منابع احتمالی تورش طی سال‌ها شناخته شده‌اند. بحث تورش به‌طور سنتی روی تورش‌هایی متمرکز است که می‌تواند در هر مرحله از اجرای یک کارآزمایی، از انتخاب شرکت‌کنندگان در گروه‌های تحقیق تا اجرای مداخلات، اندازه‌گیری پیامد، تفسیر و گزارش نتایج روی دهد.

تصور کنید که یک داروی جدید جهت درمان ام-اس در مطالعات حیوانی و در کارآزمایی‌های مرحله اول به نتایج امیدبخشی دست یافته است. این نتایج نشان داده‌اند که دارو قادر به ایجاد تأخیر در شروع اختلال حرکتی شدید در بیمار است و طی سه ماه گذشته به‌طور وسیعی در رسانه‌ها تبلیغ شده است. در نتیجه این یافته‌ها گروه‌های هوادار بیماران، دولت را تحت فشار قرار داده‌اند که هر چه زودتر داروی جدید در دسترس بیماران قرار گیرد. با توجه به این که ام-اس یک بیماری ناتوان‌کننده است که میلیون‌ها نفر در جهان به آن مبتلا هستند و با توجه به اینکه درمان شناخته‌شده‌ای برای آن وجود ندارد، بررسی‌کنندگان (کلیه پزشکانی که سال‌ها با بیماران مبتلا به ام-اس در تماس بوده‌اند)، شرکت سازنده داروی جدید (که منابع زیاد صرف تکمیل دارو نموده است) و رسانه‌ها (که علاقه‌مند به تأیید دارویی که به‌طور وسیعی تبلیغ نموده‌اند هستند) همه علاقه‌مند به یافتن نتایج مؤثر در ترکیبات جدید هستند. پس از برگزاری جلسات بحث در مورد دوره درمان، دولت یک گروه کاری چندمرکزی

شامل نمایندگان مصرف‌کنندگان تشکیل داده و با انجام مرحله بعدی به صورت کارآزمایی بالینی تصادفی شده موافقت می‌شود. یک توافق‌نامه تحقیقی توسط گروه ویژه دیگری از بررسی‌کنندگان و مصرف‌کنندگان تهیه می‌شود و یک گروه تحقیق معروف در یک مرکز بزرگ مراقبت بهداشتی جهت اجرای مطالعه انتخاب می‌شوند. عناصر این سناریوی فرضی با توجه به تورش‌هایی که ممکن است طی کارآزمایی ایجاد شود، در بخش‌های زیر تشریح خواهند شد.

۹-۳-۱ تورش انتخاب

تورش انتخاب زمانی رخ می‌دهد که پیامدهای یک کارآزمایی، تحت تأثیر اختلافات منظم در روش انتخاب یا رد افراد برای یک کارآزمایی قرار می‌گیرد، یا در حالتی که مداخلات تنها به افرادی که قبول کرده‌اند در کارآزمایی شرکت کنند اختصاص داده شود. پیشگیری از این نوع تورش دلیل اصلی استفاده از کارآزمایی‌های بالینی تصادفی شده جهت مقایسه و ارزشیابی مداخلات مختلف است. با تصادفی کردن صحیح به کلیه شرکت‌کنندگان در مطالعه شانس یکسان جهت حضور هر یک از گروه‌های مطالعه داده می‌شود. به عبارت دیگر چنانچه کارآزمایی به درستی تصادفی شود، بررسی‌کنندگان و شرکت‌کنندگان مطالعه نمی‌توانند تخصیص گروه‌ها را تحت تأثیر قرار دهند.

به هر حال یک روش کاملاً تصادفی شده برای تقسیم شرکت‌کنندگان به گروه‌های مختلف، کارآزمایی‌های بالینی تصادفی شده را از تورش انتخاب حفظ نمی‌کند. تورش انتخاب می‌تواند زمانی اعمال شود که برخی از افراد که احتمالاً واجد شرایط انتخاب بوده‌اند به دلیل آگاهی از گروهی که در آن قرار گرفته‌اند از مطالعه خارج شوند. مثلاً تیم مسئول انتخاب بیماران برای شرکت در کارآزمایی ام-اس معتقد است که بیماران افسرده احتمالاً کمتر به داروی جدید پاسخ می‌دهند. در کارآزمایی، پرسشنامه‌ای برای تعیین افسردگی شرکت‌کنندگان طراحی نشده و این تیم به دنباله اعداد تصادفی، برای اختصاص بیماران به گروه‌های تحت مطالعه، دسترسی دارند. تیم مسئول انتخاب، می‌تواند آگاهانه یا ناخودآگاه با ایجاد مانع در دریافت داروی جدید توسط بیماران افسرده موجب اعمال تورش شود. اعمال تورش حداقل از دو راه صورت می‌گیرد: اول بیماران افسرده را که در گروه دریافت‌کننده داروی جدید قرار گرفته‌اند با معیار خروج از مطالعه هماهنگ می‌شوند و یا در اغلب موارد آن‌ها در گروه دارونما قرار می‌گیرند. دوم اطلاعات مربوط به کارآزمایی را در اختیار بیماران افسرده‌ای که در گروه دریافت داروی جدید قرار گرفته‌اند، قرار می‌دهند. به طوری که آنان از رضایت دادن برای شرکت در کارآزمایی منصرف می‌شوند. در پایان کارآزمایی چنانچه حدس بررسی‌کننده درست باشد و بیماران افسرده پاسخ کمتری به داروی جدید داشته باشند کارآزمایی اثر اغراق‌آمیزی از تأثیر داروی جدید در درمان ام-اس نشان می‌دهد. مدارک مستند تأیید می‌کنند که اگر توالی تصادفی کردن شرکت‌کنندگان در زمان تهیه رضایت‌نامه شرکت در کارآزمایی، از بررسی‌کنندگان مخفی نماند، اثر مداخلات جدید می‌تواند بزرگ‌نمایی شود.

۹-۳-۲ تورش پی بردن (اطلاع)

تورش پی بردن زمانی اتفاق می‌افتد که نتایج به دست آمده از یک کارآزمایی به‌طور سیستماتیک به دلیل اطلاع

از مداخله‌ای که هر یک از شرکت‌کنندگان دریافت می‌کنند، مختل شود. تورش پی بردن ممکن است توسط فرد تجویزکننده مداخلات، دریافت‌کننده مداخلات (شرکت‌کنندگان)، بررسی‌کننده‌ای که پیامدها را مورد ارزیابی و تجزیه تحلیل قرار می‌دهد و حتی کسی که نتایج را تحلیل می‌کند یا گزارش کارآزمایی را می‌نویسد اعمال شود.

بهترین راه حفظ کارآزمایی از تورش پی بردن، ناآگاه نگه‌داشتن افراد درگیر در کارآزمایی از نوع مداخلات، است. به این عمل کورکردن یا پوشاندن نیز می‌گویند. راهبردهای کاهش تورش پی بردن را در دو مرحله از کارآزمایی می‌توان پیاده کرد. اولین مرحله، جمع‌آوری اطلاعات تجویز مداخلات و به دست آوردن اطلاعات پیامد است. دومین مرحله بعد از جمع‌آوری اطلاعات، هنگام تجزیه و تحلیل داده‌ها و گزارش نتایج است.

تورش پی بردن طی جمع‌آوری داده‌ها و از طرق مختلفی ایجاد می‌شود. مثلاً افراد اجراکننده مداخلات می‌توانند با تغییر هدفمند (سوگرایانه) در مداخلات ارائه‌شده به شرکت‌کنندگان، در کارآزمایی تورش ایجاد نمایند. در مثال کارآزمایی ام-اس، ممکن است بیماران در گروه داروی جدید زودتر و بیشتر از گروه دارونما فیزیوتراپی شوند. اگر شرکت‌کنندگان بدانند که در گروه دارونما قرار دارند، این احتمال وجود دارد که مأیوس شده و از نظر روانی تمایل کمتری به گزارش بهبودی داشته باشند. به‌علاوه اگر افراد مسئول ارزیابی و گزارش نتایج بدانند که بیماران در کدام گروه تحقیق قرار دارند، ممکن است آگاهانه یا ناخودآگاه پیامدهای بیماران دریافت‌کننده داروی جدید را مطلوب‌تر از بیماران گروه دارونما گزارش کنند.

در شرایط آرمانی، تورش پی بردن را باید با کور کردن افراد اجراکننده مداخلات، دریافت‌کنندگان مداخلات و افراد مسئول ارزیابی و گزارش پیامدها کاهش داد. در اکثر موارد، اگر افرادی که مسئول تجزیه و تحلیل یا گزارش نتایج کارآزمایی هستند از گروه‌های درمانی مطلع باشند، اثرات مداخله جدید افزایش می‌یابد. مثلاً مسئول تجزیه و تحلیل داده‌های کارآزمایی می‌تواند پیامدها یا نقاط زمانی را که حداکثر مفید بودن مداخله جدید را نشان می‌دهند انتخاب کند و پیامدها یا نقاط زمانی که نشان‌دهنده بی‌اثر یا مضر بودن مداخله جدید است، نادیده بگیرد. همین‌طور فرد مسئول گزارش نتایج کارآزمایی می‌تواند پیامدها و نقاط زمانی خاصی را مورد تأیید قرار دهد که نشان‌دهنده حداکثر تأثیر مداخله جدید باشند و یافته‌هایی را که نشان‌دهنده تأثیر مساوی یا کمتر مداخله جدید باشند را کم‌اهمیت یا نادیده بگیرد.

این منبع تورش را می‌توان با بی‌اطلاع نگه‌داشتن تجزیه و تحلیل‌کنندگان داده‌ها و مسئولین گزارش نتایج کارآزمایی از هویت گروه‌های مطالعه، کنترل کرد. برای مثال در یک مداخله دارای دو گروه، می‌توان داده‌ها را به‌صورت رمزهای الف و ب مورد تجزیه و تحلیل قرار داد. با تکمیل تجزیه و تحلیل، نتایج را می‌توان با همان رمزها به افراد مسئول نوشتن گزارش ارائه کرد. رمزها قبل از تکمیل مراحل تجزیه و تحلیل داده‌ها و گزارش نباید افشا شوند. این راهبرد به‌ندرت مورد استفاده قرار می‌گیرد. به‌طور کلی فراوانی و اهمیت تورش پی بردن ناشی از این مرحله تاکنون مورد مطالعه قرار نگرفته است.

۹-۳-۳ سایر تورش‌ها در کارآزمایی‌های بالینی

♦ تورش ناشی از بررسی‌های ناکافی انصراف‌ها، کناره‌گیری‌ها و تخطی از پروپوزال

به‌طور آرمانی، کلیه شرکت‌کنندگان در یک کارآزمایی باید مطالعه را به اتمام برسانند، از پروتکل تحقیق پیروی کرده و اطلاعات مربوط به پیامدهای موردنظر را در کلیه نقاط زمانی فراهم کنند؛ اما در عمل کارآزمایی‌ها دارای داده‌های گم شده هستند. فقدان داده‌ها ممکن است به دلیل کناره‌گیری شرکت‌کنندگان قبل از پایان کارآزمایی یا عدم رعایت پروتکل تحقیق به‌طور آگاهانه یا اتفاقی و سنجیده شدن نادرست بعضی پیامدها در یک یا چند نقطه زمانی باشد. صرف‌نظر از علت، بررسی نامناسب اطلاعات گم‌شده می‌تواند منجر به تورش شود. برای مثال اگر در کارآزمایی ام-اس بیمارانی که داروی جدید در موردشان مفید نبوده به دلیل اثرات جانبی مضر، اغلب از مطالعه انصراف دهند، به علت عدم حضور آن‌ها اثرات مفید داروی جدید بیشتر می‌شود و اثر جانبی آن کمتر از مقدار واقعی نشان داده می‌شود.

♦ تورش نشر

بعضی شواهد نشان‌دهنده گرایش در بررسی‌کنندگان، مسئولین تدوین و ارائه مقاله، گروه منتقدان و سردبیران جهت قبولی دست‌نویس‌هایی برای چاپ است که کارآزمایی‌های با نتایج مثبت در آن گزارش شده است. اربیبی انتشار در متاآنالیز چیست؟ بسیاری از مطالعات در نقاط مختلف دنیا انجام می‌شود ولی نتایج آن‌ها به دلایل مختلف منتشر نمی‌شود یا حداقل ممکن است در مجلات ایندکس نشده منتشر گردند. وجود تفاوت‌های معنی‌دار، احتمال چاپ مقاله را افزایش می‌دهد. چاپ نشدن مقالات بدون نتایج معنی‌دار را *file-drawer* می‌گویند. نتایج فراتحلیل حاصل از این مقالات دارای خطای منظمی به نام اربیبی انتشار است. پس منظور از اربیبی انتشار این است که یک فرا تحلیل شامل تمام مطالعات انجام شده نیست. در صورت وجود اربیبی انتشار، نتایج فراتحلیل دارای بیش برآوردی (*overestimate*) خواهد بود. روش‌های بررسی اربیبی انتشار شامل: نمودار جنگلی، نمودار فونل، نمودار چندک نرمال، روش همبستگی رتبه‌ای بگ، روش رگرسیون ایگر، روش رگرسیون نمودار فونل و روش برش و پر کردن است که اطلاعات مبسوط آن در منابع شماره ۱۷، ۱۸ و ۱۹ این فصل قابل‌دسترس است.

♦ تورش زبان

نوعی از تورش نشر، تورش زبان است و نشانگر آن است که دست‌نوشته‌ها بسته به سمت‌وسوی نتایج به دست آمده به مجلات با زبان‌های مختلف ارائه و به چاپ می‌رسند. بیشتر مطالعات با نتایج مثبت به زبان انگلیسی چاپ شده است. مشخص شده که در بعضی کشورها محققان ممکن است تنها نتایج مثبت را گزارش کنند. نظیر کارآزمایی‌های بالینی تصادفی شده ارزیابی‌کننده کارایی طب سوزنی، که در کشورهای چین، ژاپن، هنگ‌کنگ و تایوان اجرا شده است.

♦ تورش ناشی از تأخیر زمانی

این تورش زمانی روی می‌دهد که سرعت نشر بستگی به سمت‌وسو و قدرت نتایج کارآزمایی دارد. به‌طورکلی به نظر می‌رسد که نشر کارآزمایی‌های دارای نتایج منفی دو برابر بیشتر از نتایج مثبت طول می‌کشند. همچنین مطالعات انجام شده روی این تورش نشان می‌دهد که کارآزمایی بعد از تکمیل ممکن است سال‌ها طول بکشد تا

ارائه و منتشر شود. به طوری که زمان صرف شده در تکمیل کارآزمایی مساوی زمان صرف شده برای نشر است.

♦ تورش توان خارق‌العاده

این تورش توسط روزنامه‌نگاران (و به طور فزاینده‌ای توسط ناشرین اینترنتی) که کارآزمایی‌ها را بر اساس سمت و سوی یافته‌های به طور هدفمند (سوگرایانه) انتخاب، بزرگنمایی و انتشار می‌دهند، روی می‌دهد. این مسئله می‌تواند زمانی اتفاق بیفتد که مثلاً روزنامه‌نگاران اقدام به جلب توجه عموم به نتایج یک کارآزمایی که مدعی اثرات مفید یک مداخله جدید جهت یک بیماری لاعلاج است، می‌کنند.

♦ تورش اخلاقی

بزرگ‌نمایی یا دست کم گرفتن مطالعه بستگی به میزان موافقت یا مخالفت آن با اخلاقیات خواننده دارد. برای مثال خواننده‌ای که سقط جنین را عمل غیر اخلاقی می‌داند، نتایج مطالعه‌ای را که نشان دهنده ارتباط بین سقط جنین با سرطان پستان است، بزرگ‌نمایی می‌کند.

♦ تورش مجله خوشنام

این تورش زمانی روی می‌دهد که نتایج مطالعاتی که در مجله خوشنام منتشر می‌شود، بزرگ‌نمایی شود.

♦ تورش مجله غیر خوشنام

عکس تورش مجله خوشنام، این تورش زمانی روی می‌دهد که نتایج مطالعاتی که در مجله غیر خوشنام منتشر می‌شود کم ارزش نمایانده شود.

♦ تورش نویسندگان سرشناس

این تورش زمانی روی می‌دهد که نتایج مطالعات منتشر شده توسط نویسندگان سرشناس بزرگ‌نمایی می‌شوند.

♦ تورش نویسندگان ناشناس یا غیر سرشناس

این تورش زمانی روی می‌دهد که نتایج مطالعات منتشر شده توسط نویسندگان غیر سرشناس دست کم گرفته می‌شوند.

♦ تورش مؤسسه مشهور

زمانی ایجاد می‌شود که نتایج مطالعات منتشر شده توسط مؤسسات مشهور بزرگ‌نمایی می‌شوند.

♦ تورش مؤسسه ناشناخته

این تورش مربوط به تورش قبلی است. زمانی که نتایج مطالعات منتشر شده از مؤسسات ناشناخته یا غیر خوشنام به طور هدفمند (سوگرایانه) دست کم گرفته می‌شوند روی می‌دهد.

♦ تورش کارآزمایی بزرگ

در این تورش کارآزمایی‌های بزرگ، بزرگ‌نمایی می‌شوند.

♦ تورش کارآزمایی چندمرکزی

زمانی روی می‌دهد که کارآزمایی‌های چند مرکزی بزرگ‌نمایی شود. این کارآزمایی‌ها لزوماً دارای نمونه‌های بزرگ‌تر نیستند.

♦ تورش کارآزمایی کوچک

زمانی روی می‌دهد که نتایج کارآزمایی‌های دارای نمونه‌های کوچک دست کم گرفته شوند، بخصوص وقتی ناقص نظر خواننده باشد. مثلاً احتمال هر گونه تأثیر معنی‌دار یا بالینی یافت شده در یک کارآزمایی کوچک را به شانس نسبت دادن و یا فقدان نتایج معنی‌دار بودن را ناشی از کاهش قدرت مطالعه دانستن. یکی از دلایل ایجاد این تورش هستند.

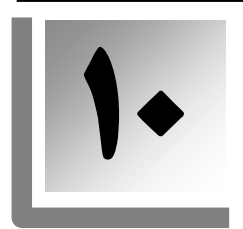
♦ تورش جغرافی

تورش جغرافی زمانی روی می‌دهد که تحقیقات بر حسب کشور یا ناحیه‌ای که در آن اجرا شده‌اند بزرگ‌نمایی یا دست کم گرفته شوند.

منابع

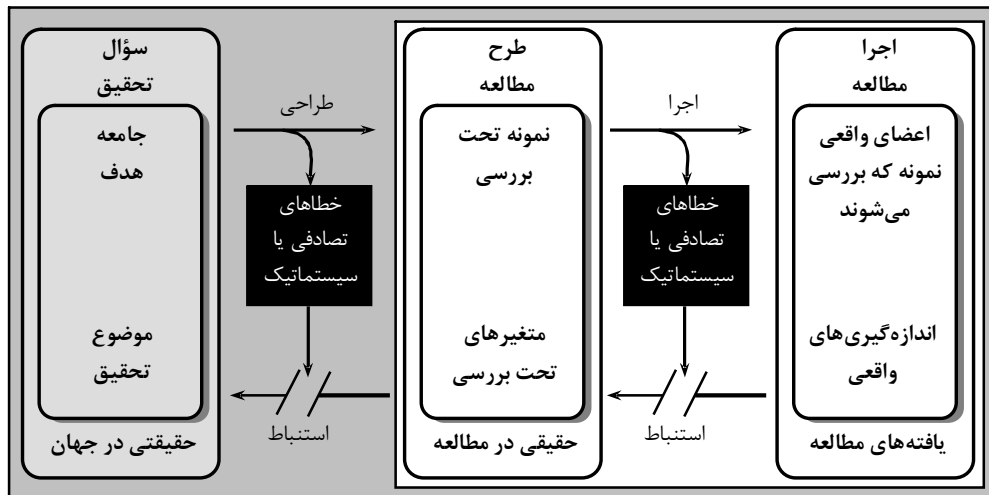
1. Szklo M, Nieto FJ. *Epidemiology: Beyond the basics*. Aspen, Maryland, 2000.
2. Mac Mahon B, Yens, Trichopoulos D, et al. Coffee and cancer of the pancreas. *N Engl J Med*. 1981. 304: 630-3.
3. Hsieh CC, Mac Mahon B, Yen S, et al. Coffee and pancreatic cancer. *N Engl J Med*. 315: 587-8, 1986 Letter.
4. Gustafson P. *Measurement error and misclassification in statistics and Epidemiology*. Chapman & Hall, New York 2004.
5. Weinstock MA, Colditz GA, Willett WC, et al. Recall (report) bias and reliability in the retrospective assessment of melanoma risk. *Am J Epidemiol*. 133:240-5, 1991.
6. Perneger TV, Whelton PK, Klag MJ, Rossiter KA. Diagnosis of hypertensive end stage renal disease: effect of patient's race. *Am J Epidemiol*. 141:10-15, 1995
7. Stewart WF, Linet MS, Celentano DD, et al. Age and sex: Specific incidence rates of migraine with and without visual aura. *Am J Epidemiol*. 134:1111-20, 1991
8. Rose GA. Chest pain g questionnaire. *Milbank Mem Fund Q*, 43:32-9, 1965.
9. Bass E13, Follansbee WP, Orchard TJ. Comparison of a supplemented Rose questionnaire to exercise thulium testing in men and women. *J Clin Epidemiol*. 1989, 42:385-93.
10. Garber CE, Corleton RA, Heller Gv. Comparison of Rose questionnaire angina to exercise thallium scintigraphy: Different findings in males and females. *J Clin Epidemiol*, 1992. 45:715-20.
11. Sorlie PD, Cooperl, Schreiner PJ, et al. Repeatability and validity of the Rose questionnaire for angina pectoris in the arteriosclerosis risk in communities study. *J Clin Epidemiol*. 49:719-25, 1996.
۱۲. سوری ح، رحیمی م، محسنی ح. بررسی ارتباط بین استرس‌های شغلی و حوادث ناشی از کار: یک مطالعه مورد-شاهد. *مجله تخصصی اپیدمیولوژی ایران*. زمستان ۱۳۸۴، دوره اول، شماره ۲، ۵۸-۵۳.
13. Boccia S, La Torre G, Persiani R, D'Ugo D, van Duijn CM, Ricciardi G. A critical

- appraisal of epidemiological studies comes from basic knowledge: a reader's guide to assess potential for biases. *World J Emerg Surg.* 2007 Mar, 15; 2: 7.
14. Haslam N. Bias in psychopathology research. *Curr Opin Psychiatry.* 2006.
15. Sica GT. Bias in research studies. *Radiology.* 2006 Mar;238(3):780-9. Review.
16. Bell CM, Urbach DR, Ray JG, Bayoumi A, Rosen AB, Greenberg D, Neumann PJ. Bias in published cost effectiveness studies: systematic review. *BMJ.* 2006 Mar 25;332(7543):699-703.
17. Whitehead A. *Meta-Analysis of Controlled Clinical Trials*, John Wiley & Sons Ltd, (2002).
18. Rothstein HR, Sutton AJ, Borenstein M. *Publication Bias in Meta-Analysis*, John Wiley & Sons Ltd, (2005).
۱۹. سوری ح. بررسی و نقد مقالات اپیدمیولوژی، مجله دانشکده پزشکی دانشگاه علوم پزشکی گیلان، ۱۳۸۱؛ دوره ۱۱، بهار، شماره ۴۱: صفحات ۶۹-۶۴.



هدایت مطالعه: مطالعه راهنما، پیش‌آزمون، کنترل کیفیت و اصلاحات پروتکل

مباحث قبلی بیشتر مرتبط با طراحی مطالعه بود. این قسمت مطالب مرتبط با سمت چپ شکل ۱۰-۱ (یعنی قسمت خاکستری‌رنگ) را شامل می‌شود. در این فصل به مسائل مربوط به اجرای مطالعه یعنی مطالعه راهنما، پیش‌آزمون، کنترل کیفیت، تخطی از پروتکل و اصلاح پروتکل بررسی می‌شوند.



شکل ۱۰-۱: نحوه اجرای مطالعه

ابتدا مطالعه راهنما را بررسی می‌کنیم. حتی بهترین طراحی‌های متفکرانه در عمل ممکن است تغییر کند. میزان پاسخ‌های به‌دست‌آمده یا میزان پیگیری ممکن است خیلی کمتر از مقدار پیش‌بینی شده باشد و اندازه‌گیری‌ها ممکن است به نتایج غیرعملی منجر شود. بهترین راه برای آشکار شدن این مشکلات انجام مطالعه راهنما است. با انجام مطالعه راهنما قبل از آغاز مطالعه اصلی می‌توان در پروتکل، تغییرات مناسبی اعمال کرد.

مرحله بعدی، کنترل کیفیت است. بسیاری از محققان توجه زیادی به اجرای مطالعه ندارند. زیرا این موضوع خسته‌کننده است و کمتر از مرحله تحلیل و طراحی به آن توجه می‌شود. نتایج مطالعه‌ای که به‌خوبی طراحی شود

در نتیجه دقت پایین در اجرا و بروز خطاها در این مرحله می‌تواند گمراه‌کننده باشد. راه‌حل این مشکل بهبود سیستم و افزایش کیفیت داده‌ها در زمان اجرای مطالعه است.

۱-۱۰ اهمیت مطالعات راهنما (Pilot Studies)

واژه "مطالعات راهنما یا پیلوت" به نسخه کوچکی از یک مطالعه با مقیاس کامل و هم‌چنین پیش‌آزمودن اختصاصی یک ابزار پژوهشی خاص، مانند پرسشنامه یا برنامه مصاحبه گفته می‌شود. مطالعات راهنما یک جزء مهم از یک طرح مطالعاتی خوب هستند. انجام مطالعه راهنما تضمینی بر موفقیت مطالعه اصلی نیست، اما احتمال موفقیت را افزایش می‌دهد.

واژه "مطالعه راهنما" به دو صورت در پژوهش‌های کاربردی استفاده می‌شود. از یک‌سو، این واژه به قابلیت اجرای مطالعه اشاره می‌کند که عبارت‌اند از مطالعات با مقیاس کوچک یا کارآزمایی‌هایی که به‌منظور آماده‌سازی مطالعه اصلی انجام می‌شوند. از سوی دیگر امکان دارد در مطالعه راهنما به پیش‌آزمون (Pre-testing) یا ارزیابی یک ابزار پژوهشی خاص پرداخت. یکی از مزایای انجام مطالعه راهنما این است که می‌توان پیش از انجام مطالعه اصلی، آگاهی یافت که در کدام قسمت ممکن است طرح پژوهشی با شکست روبرو شود، متوقف شود یا این‌که روش‌ها یا ابزار پیشنهادی، نامناسب یا بیش‌ازحد پیچیده هستند. دلایل دیگری از قبیل متقاعد کردن تأمین‌کنندگان بودجه در این مورد که پروپوزال پژوهشی ارزش سرمایه‌گذاری دارد، از علت‌های دیگر انجام مطالعه راهنما است. بنابراین دلایل مختلفی برای انجام مطالعات راهنما وجود دارد. جدول ۱-۱۰ به بررسی دلایل انجام مطالعه راهنما اختصاص دارد. مطالعات راهنما بر مبنای روش‌های کمی و یا کیفی انجام می‌شوند. مطالعات با مقیاس وسیع نیاز به انجام مطالعات راهنما متعدد قبل از انجام بررسی اصلی دارند.

جدول ۱-۱۰: دلایل انجام مطالعات راهنما

دلایل انجام مطالعه راهنما
تهیه ابزار پژوهشی و آزمودن کفایت این ابزار
ارزیابی قابلیت اجرای یک مطالعه یا بررسی امکان اجرای آن در مقیاس وسیع
طراحی پروتکل پژوهشی
ارزیابی واقع‌گرایانه از عملی بودن پروتکل پژوهشی
اطمینان از مؤثر بودن چارچوب و تکنیک نمونه‌گیری
ارزیابی میزان موفقیت روش‌های پیشنهادی جهت ورود افراد به مطالعه
شناسایی مشکلات قانونی یا اخلاقی که ممکن است در حین به‌کارگیری روش‌های پیشنهادی رخ دهد.
برآورد تغییرپذیری (variability) پیامدها جهت تعیین حجم نمونه
جمع‌آوری داده‌های اولیه
تعیین منابع موردنیاز از قبیل کارکنان یا منابع مالی برای مطالعه طراحی‌شده
تنظیم سؤال و طرح پژوهشی
آموزش فرآیند پژوهش به پژوهشگران
متقاعد ساختن تأمین‌کنندگان بودجه در این مورد که گروه پژوهشی دارای صلاحیت و دانش لازم است.
متقاعد ساختن تأمین‌کنندگان بودجه در این مورد که مطالعه اصلی قابل اجرا و دارای ارزش سرمایه‌گذاری است.
متقاعد ساختن سایر گروه‌های ذینفع در این مورد که مطالعه اصلی ارزش حمایت دارد.

برای مثال برای انجام سرشماری نفوس و مسکن ۱۳۸۵ که بزرگ‌ترین بررسی ده‌ساله در کشور است، تغییرات متدولوژیک و سایر تغییرات نسبت به پرسشنامه سرشماری سال ۱۳۷۵ آزمون شد. این سرشماری‌های امتحانی اطلاعات اساسی در زمینه واکنش مردم نسبت به سؤالات و سبک نگارش جدید پرسشنامه، نحوه انجام عملیات میدانی و روش کاهش بی‌پاسخ‌ها را فراهم کرد و میزان موفقیت روش‌های مدیریت، اجرا، جمع‌آوری و پردازش ارزیابی شد.

در مطالعات راهنما، سعی بر این است تا مشکلات عملی بالقوه برای انجام پژوهش مشخص شود. سایر مشکلات از قبیل ناقص بودن مدارک، پایین بودن میزان پاسخ‌دهی موارد گم‌شده و تخطی از پروپوزال را نیز می‌توان از طریق مطالعه راهنما شناسایی و تدابیر احتیاطی و اطمینان‌بخش را اتخاذ نمود. مراحل پیش‌آزمون برای یک پرسشنامه بر روی یک گروه کوچک از داوطلبانی که تا حد امکان به جمعیت هدف مطالعه شباهت دارند، در جدول ۱۰-۲ بررسی شده است. مطالعات راهنما قادر به آشکار ساختن سیاست‌ها یا مشکلات و مسائل محلی هستند که ممکن است بر روی فرایند پژوهش تأثیر بگذارد.

۱۰-۱-۲ مشکلات مطالعات راهنما

مطالعات راهنما محدودیت‌هایی نیز دارند که عبارت‌اند از: احتمال پیشگویی‌ها یا برداشت‌های نادرست بر اساس داده‌های مطالعه راهنما، مشکلات ناشی از آلودگی (Contamination) و مسائل مرتبط با بودجه. در این قسمت هر یک از این موضوعات مورد بحث قرار می‌گیرد.

جدول ۱۰-۲: شیوه‌های انجام مطالعه راهنما به‌منظور بهبود اعتبار داخلی یک پرسشنامه

نحوه انجام مطالعه راهنما برای پرسشنامه‌ها
پرسشنامه را دقیقاً به همان صورتی که در مطالعه اصلی، ارائه می‌شود، به افراد مطالعه راهنما ارائه دهید. از افراد تحت مطالعه تقاضا کنید موارد مبهم و سؤالات مشکل را شناسایی کنند. زمان بکار رفته برای تکمیل پرسشنامه را ثبت کنید و در مورد منطقی بودن آن تصمیم بگیرید. کلیه سؤالات غیرضروری، مشکل یا مبهم را حذف کنید. کافی بودن گزینه‌های موجود برای هر سؤال را ارزیابی کنید. مطمئن شوید که پاسخ‌ها را می‌توان برحسب اطلاعات موردنیاز، تفسیر نمود. بررسی کنید که افراد تحت بررسی به همه سؤالات، پاسخ داده باشند. هر سؤالی را که مطابق انتظار به آن پاسخ داده نشده است، مجدداً جمله‌بندی کنید و یا محل آن را در پرسشنامه تعویض نمایید. پرسشنامه را خلاصه و تجدیدنظر کنید و در صورت امکان مطالعه راهنمای دیگری انجام دهید.

به اتمام رساندن موفقیت‌آمیز مطالعه راهنما تضمینی برای موفقیت مطالعه اصلی در ابعاد وسیع نیست. اگرچه نتایج مطالعه راهنما نشان‌دهنده میزان احتمالی پاسخ‌دهی در بررسی اصلی است؛ اما تضمین‌کننده آن نیست، زیرا پایه آماری ندارد و تقریباً در همه موارد بر اساس حجم نمونه ناکافی به‌دست آمده است. به‌علاوه، ممکن است سایر مشکلات جز در زمان اجرای مطالعه با مقیاس بزرگ‌تر، مشخص نشود.

مسئله مورد ملاحظه بعدی، آلودگی است. آلودگی از دو طریق ممکن است ایجاد شود:

- زمانی که داده‌های مطالعه راهنما در نتایج مطالعه اصلی وارد شوند.
 - زمانی که شرکت‌کنندگان در مطالعه راهنما در مطالعه اصلی نیز شرکت کنند اما از آن‌ها داده‌های جدید جمع‌آوری شود.

یکی از ویژگی‌های اساسی مطالعه راهنما در مطالعات کمی این است که در هنگام گزارش نتایج، داده‌های آن برای آزمون فرضیه به کار نمی‌رود و در مطالعه واقعی تحلیل نمی‌شود. واضح است که اگر نتایج مطالعه راهنما نشان دهد که مشکلاتی در رابطه با ابزار پژوهشی وجود دارد و لازم است تغییراتی در آن صورت بگیرد، این داده‌ها ممکن است ناقص یا نادرست باشند؛ اما زمانی که از ابزار پژوهشی استاندارد و معتبر (Validated) استفاده می‌شود و مطالعه راهنما سایر جنبه‌های متدولوژیک مانند میزان ورود افراد به مطالعه را تعیین می‌کند، داده‌های حاصل با ارزش هستند.

یک مشکل شایع‌تر، تصمیم‌گیری در این مورد است که آیا شرکت‌کنندگان یا مکان‌های مورد مطالعه در مطالعه راهنما، وارد مطالعه اصلی شوند یا خیر. مسئله مهم این است که، این افراد قبلاً در مواجهه با مداخله قرار گرفته‌اند و در نتیجه ممکن است پاسخ آن‌ها نسبت به افرادی که قبلاً با این مداخله مواجهه نداشته‌اند، متفاوت باشد. برای مثال شرکت‌کنندگان با نحوه به‌کارگیری ابزار یا شیوه اندازه‌گیری آشنا می‌شوند، اما از سوی دیگر ممکن است تکرار مجدد مراحل پژوهش بر روی آن‌ها تأثیر منفی داشته باشد و با موارد زیاد تخطی از پروتکل توسط افراد تحت مطالعه پژوهشی مواجه شویم. زیرا این پروتکل دیگر برای آن‌ها جذابیت ندارد.

بنابراین حضور شرکت‌کنندگان در مطالعه راهنما، در مطالعه اصلی به علت این که تنها افراد بررسی شده در مطالعه راهنما تجربه بیشتری کسب می‌کنند (که این موضوع در مورد تمام افراد تحت مطالعه، صدق نمی‌کند) باعث ایجاد نوعی تورش در مطالعه می‌شود؛ اما در بعضی موارد به‌راحتی نمی‌توان شرکت‌کنندگان مطالعه راهنما را از مطالعه اصلی حذف نمود، زیرا در این صورت حجم نمونه مطالعه اصلی بسیار کوچک خواهد شد. این مسئله به‌خصوص در مواردی که طرح نمونه‌گیری خوشه‌ای است بیشتر نمود پیدا می‌کند. در این موارد می‌توان برای ارزیابی این که فرآیند پیش‌آزمون یا مطالعه راهنما تا چه حد بر روی اندازه اثر مداخله، تأثیر می‌گذارد تحلیل حساسیت (Sensitivity analysis) یا تحلیل در زیرگروه‌ها (Sub-group Analysis) انجام داد.

در مطالعات کیفی، مسئله آلودگی اهمیت کمتری دارد و پژوهشگران اغلب قسمتی از داده‌های مطالعه راهنما یا کل آن را در مطالعه اصلی بکار می‌گیرند. جمع‌آوری و تحلیل داده‌های ناشی از پژوهش‌های کیفی اغلب تدریجی و مستمر است و مصاحبه دوم یا مصاحبات بعدی، معمولاً مؤثرتر از مصاحبه قبلی است زیرا مصاحبه‌گر، با استفاده از بینش و معلوماتی که از مصاحبه‌های قبلی به‌دست آورده است، برنامه‌های مصاحبه و سؤالات خاصی از آن را اصلاح می‌کند. به همین دلیل بعضی از افراد استدلال می‌کنند که در رویکردهای کیفی، انجام مطالعات راهنما به‌صورت مجزا، ضروری نیست.

مشکلات دیگر ممکن است زمانی ایجاد شوند که مطالعه راهنما نیاز به یک سرمایه‌گذاری وسیع داشته باشد، در نتیجه برای گروه پژوهشی مشکل باشد که در صورت ناموفق بودن مطالعه راهنما پژوهش را متوقف کنند. پژوهشگران ممکن است به‌جای این که نتیجه بگیرند که مطالعه پیشنهادی آن‌ها با توجه به منابع، زمان، جمعیت و سایر عوامل موجود و در دسترس، غیرقابل انجام است، ترغیب شوند تغییرات زیادی در مطالعه اصلی اعمال

کنند. از سوی دیگر، در صورتی که مطالعه راهنما یک مطالعه بزرگ و جامع باشد، تأمین‌کنندگان بودجه تمایل خود را به سرمایه‌گذاری بیشتر در این مطالعه از دست می‌دهند، زیرا دیگر به آن به‌عنوان یک پژوهش جدید (Original) نمی‌نگرند، بخصوص اگر نتایج مطالعه راهنما چاپ و منتشر شده باشد.

۱۰-۱-۳ گزارش مطالعات راهنما

تورش نشر ممکن است باعث شود مطالعه راهنما هیچ‌گاه گزارش نشود. زیرا مجلات تنها تمایل به پذیرش مقاله‌هایی دارند که نتایج آن‌ها از لحاظ آماری معنی‌دار است و اثراتی را که از نظر آماری معنی‌دار نیستند، گزارش نمی‌کنند. مطالعه‌ای در زمینه پژوهش‌های مربوط به سیگار کشیدن غیرفعال نشان داده که بین میانه (Median) نشر نتایج مطالعات معنی‌دار و مطالعاتی که معنی‌دار نیستند، دو سال تفاوت وجود دارد. هم‌چنین مقالاتی که در آن به موضوعاتی که هنگام اجرای یک مطالعه راهنما شناسایی می‌شوند، جذابیت کمتری برای ناشران دارند.

۱۰-۱-۴ نتیجه‌گیری

مطالعات راهنما کمتر مورد بحث و استفاده قرار می‌گیرند و کمتر گزارش می‌شوند. به‌ندرت می‌توان گزارش کامل مطالعات راهنما را در متون پژوهشی یافت. در صورتی هم که این مطالعات گزارش شوند، اغلب تنها روش‌های پژوهشی یا ابزار پژوهشی خاصی را تأیید می‌کنند. در بسیاری از موارد، در مقالات پژوهشی تنها به یک جزء مطالعه راهنما مانند "پیش‌آزمون" یا بررسی یک پرسشنامه اشاره می‌شود. این مقالات به‌سادگی بیان می‌کنند: "اعتبار و پایایی پرسشنامه، آزمون شد."

زمانی که جزئیات مطالعات راهنما در مقاله و گزارش‌ها بیان می‌شود، اغلب پژوهشگران اذعان می‌دارند که از مطالعه راهنما درس‌هایی آموختند و تغییرات لازم را اعمال کردند، بدون این‌که جزئیات آنچه را که دقیقاً آموخته‌اند، به خواننده منتقل کنند. برخی از این فرآیندها و پیامدهای حاصل شده از مطالعات راهنمای موفق و غیر موفق می‌توانند برای کسانی که پروژه‌های مشابهی را آغاز می‌کنند، مفید باشد. این امر بسیار مهم است زیرا مطالعات راهنما "زمان‌بر، ناامیدکننده و همراه با مشکلات غیرقابل‌پیش‌بینی هستند؛ اما به‌رحال بهتر است قبل از آن که زمان، هزینه و تلاش‌های زیادی صرف مطالعه اصلی شود، مطالعه راهنما انجام شود."

مطالعات راهنمایی که به‌خوبی طراحی و اجرا شده‌اند، می‌توانند محققان را از بهترین شیوه انجام پژوهش و گاهی پیامدهای احتمالی آن آگاه سازند. بنابراین محققان باید تشویق شوند تا مطالعات راهنما را کامل‌تر گزارش کنند و به‌خصوص اصلاحات واقعی را که در طراحی مطالعه و فرآیند پژوهش صورت گرفته است به‌طور دقیق و با جزئیات بیان کنند.

۱۰-۱-۵ نهایی کردن پروتکل

هدف نهایی مطالعه راهنما، تصمیم‌گیری در مورد چگونگی طراحی روش‌های اعمال مداخله، انجام اندازه‌گیری و مدیریت نیروی انسانی است. مطالعات راهنما برای بررسی روش‌های مدیریت نیروی انسانی، برآوردی از تعداد

افراد در دسترس که حاضر به همکاری هستند و کارایی روش‌های مختلف مدیریتی در مطالعه به کار می‌رود. مطالعات راهنما دورنمایی از ساختار جامعه‌ای که از آن نمونه‌گیری شده است در اختیار محقق قرار می‌دهد و توزیع سن، جنس، نژاد و شاخص‌های دیگری که در مطالعه مهم هستند برآورد می‌شود. مطالعه راهنما در بهبود روش‌های اندازه‌گیری متغیرها مهم هستند. پیش‌آزمون‌های مقایسه‌ای کوچک، برای آزمون روش‌های مختلف در مراحل ابتدایی به کار می‌روند. در مرحله پیش‌آزمون می‌توان واکنش‌های فرعی، قابل فهم بودن سؤالات پرسشنامه و روش‌های بهبود مطالعه را بررسی کرد. امکان دارد در مرحله برنامه‌ریزی، پیش‌آزمون‌های وسیع روی افرادی که عضو جامعه اصلی هستند انجام شود. این کار برای بررسی اعتبار و اطمینان از درستی روش‌های اجرایی صورت می‌گیرد. اگر نمونه‌ها در مطالعه شامل افرادی با سن، جنس، نژاد و فرهنگ‌های گوناگون باشد و اندازه‌گیری‌ها به این گوناگونی حساس باشند در پیش‌آزمون باید از هر زیرگروه اصلی نمایندگان مناسبی انتخاب شوند. اگر در مطالعات، یک مداخله جدید به کار رفته باشد، پیش‌آزمون‌ها بسیار مفید هستند. مداخلات طراحی شده به‌منظور تغییر رفتار، مانند برنامه‌های پیش‌گیری از ایدز در مدارس، تردیدآمیز هستند.

۱۰-۲ کنترل کیفیت

یکی از موضوعات بسیار مهم در تحقیقات بالینی که معمولاً مورد بی‌توجهی قرار می‌گیرد، اطمینان از کیفیت و سازگاری داده‌ها است. زیرا مراحل جمع‌آوری داده‌ها معمولاً تکراری و خسته‌کننده است و اندازه‌گیری‌ها اغلب درست‌تر از آنچه هستند به نظر می‌رسند. اگر روش‌های کنترل کیفیت مناسب برای تهیه داده‌های معتبر وجود نداشته باشد، وجود داده‌های نادرست و گم‌شده باعث به وجود آمدن مشکلاتی می‌شود.

۱۰-۲-۱ داده‌های گم‌شده

اگر نسبت زیادی از اندازه‌ها، داده‌های گم‌شده باشند و یا حتی اگر چند اندازه گم‌شده در مطالعه وجود داشته باشد نتایج گمراه‌کننده و دارای تورش است. مثلاً مطالعه‌ای با پیگیری طولانی مدت درباره یک عمل جراحی که دارای میزان مرگ ۵٪ است در نظر بگیرید، اگر ۱۰٪ بیماران در پیگیری‌ها از مطالعه خارج شوند (گم شوند) و مرگ در بین این افراد شایع باشد، این عارضه به‌طور خیلی جدی کمتر از اندازه واقعی برآورد می‌شود. نتایج گمراه‌کننده مربوط به داده‌های گم‌شده گاهی اوقات در مرحله تجزیه و تحلیل قابل تصحیح است.

روش‌های آماری متنوعی به‌ویژه در تحلیل چند متغیره که وجود داده‌های گم‌شده در مقابل تعدادی متغیر مستقل باعث می‌شود اطلاعات زیادی در دسترس باشد، برای جایگزین کردن پاسخ گم‌شده بر اساس اطلاعات موجود وجود دارد. با این حال تنها راه‌حل مناسب، طراحی و انجام مطالعه به صورتی است که داده‌های گم‌شده به حداقل برسد. اگرچه این روش‌ها مفید هستند، اما اگر تعداد قابل توجهی مشاهدات گم‌شده وجود داشته باشد هیچ تضمینی برای بدون تورش بودن پاسخ‌ها وجود ندارد.

۱۰-۲-۲ داده‌های نادرست و با دقت پایین

داده‌های نادرست با دقت پایین باعث به وجود آمدن مشکلاتی می‌شوند که اغلب ناشناخته باقی می‌مانند. این مشکل به‌خصوص وقتی پیش می‌آید که بیشتر از یک نفر به جمع‌آوری داده‌ها می‌پردازد (وقتی محقق مطالعه‌ای را طراحی می‌کند و جمع‌آوری داده‌ها را به همکارانش واگذار کند). در مرحله تحلیل داده‌ها بعضی از این اندازه‌گیری‌ها به همراه استفاده مداوم از یک روش تحلیل نامناسب به‌طور جدی باعث تورش نتایج می‌شوند. وقتی خطاهای مربوط به واردکردن داده‌ها بعد از انجام مطالعه آشکار نشوند، مشکل جدی‌تری به وجود می‌آید.

۱۰-۲-۳ داده‌های گمراه‌کننده

محققان بالینی که یک سیستم تحقیق را رهبری می‌کنند باید در نظر داشته باشند که ممکن است همکاران تیم تحقیق به‌دقت همکاری نکنند. این افراد نادرست واردکردن اطلاعات را آسان‌ترین راه انجام وظایف خود یافته‌اند. راه‌های اطمینان از عدم وقوع چنین خطاهای بزرگی عبارت‌اند از: دقت در انتخاب همکاران و اعضای طرح، برقراری ارتباط با تیم تحقیق و فهماندن اهمیت موضوع تحقیق به آن‌ها، آگاه بودن از احتمال وقوع خطاهای عمدی و ارزیابی اتفاقی داده‌ها هنگام انجام مطالعه [۱].

۱۰-۲-۴ کنترل کیفیت مراحل بالینی تحقیق

کنترل کیفیت مراحل بالینی تحقیق از مرحله اجرا شروع می‌شود و در تمام طول اجرای مطالعه ادامه می‌یابد. (جدول ۱۰-۳)

- دستورالعمل عملیاتی تحقیق

دستورالعمل عملیاتی متمم بخش پروتکل مطالعه است. دستورالعمل عملیاتی به‌طور دقیق افراد مطالعه، نحوه اندازه‌گیری متغیرها و روش انجام مطالعه را تعریف می‌کند. دستورالعمل عملیاتی هنگامی که تحقیق به‌وسیله افراد متعدد انجام می‌شود، به‌ویژه وقتی که چند محقق در چند مرکز مختلف همکاری می‌کنند، بسیار ضروری است. حتی وقتی یک محقق همه کارها را به‌تنهایی انجام می‌دهد. دستورالعمل عملیاتی در کاهش تغییرات تصادفی و تغییرات در تکنیک اندازه‌گیری در طول زمان به او کمک می‌کند.

موضوعات مهم یک دستورالعمل عملیاتی، در جدول ۱۰-۳ بیان شده است. این جداول شامل همه ابزارها و فرم‌های مورد استفاده در مطالعه است که شامل چگونگی تماس با افراد مورد مطالعه، انجام مصاحبات، پر کردن پرسشنامه، کدگذاری و اصلاح فرم‌ها، جمع‌آوری نمونه‌ها، تحلیل داده‌ها و روش‌های کنترل کیفیت است.

جدول ۱۰-۳: مراحل کنترل کیفیت در پژوهش‌های بالینی

کنترل کیفیت
مرحله طراحی
مشخص کردن روش‌های مدیریت و تحلیل داده‌ها تدوین سیستم کنترل کیفیت لزوم و امکان کور کردن آموزش تیم تحقیق تأیید مهارت‌های تیم تحقیق
مرحله اجرا
هدایت و رهبری مستمر جلسات کاری دوره‌ای تأیید مجدد مهارت‌های تیم تحقیق گزارش‌های دوره‌ای

- آموزش و تأیید مهارت

در مطالعاتی که به صورت گروهی انجام می‌شوند افرادی که اندازه‌گیری‌های مهم را انجام می‌دهند آموزش می‌بینند و صلاحیت آن‌ها آزمون و تأیید می‌شود.

- رهبری و هدایت

در مطالعاتی که به صورت گروهی انجام می‌شود وظیفه کنترل کیفیت به عهده مجری اصلی است. او نمی‌تواند همه اندازه‌گیری‌های همکارانش را مشاهده کند، اما اگر در مورد درستی داده‌ها تردید داشته باشد می‌تواند صحت انجام اندازه‌گیری‌ها و یا مشاهده و ثبت اطلاعات را آزمون کند و یا عملکرد اعضای تیم تحقیق را کنترل نماید.

- جلسات کاری

از آغاز مرحله طراحی، محقق باید با همه اعضای تیم تحقیق جلسات منظم داشته باشد. در جلسات باید دستورالعمل مراحل بعدی توزیع شود و افرادی که مسئولیتی در هر مرحله مطالعه داشته‌اند، گزارش کار ارائه دهند. این جلسات فرصتی برای کشف و حل مشکلات فراهم می‌کند و باعث می‌شود هرکس با توجه به پیشرفت برنامه و زمان‌بندی مطالعه مسئولیت‌های خود را انجام دهد. جلسات منظم مرجع خوبی برای ارزیابی تحقق اهداف مطالعه است.

- بازیابی

همه اعضای تیم باید در سیستم بازیابی، که شیوه‌های مورد استفاده توسط هر عضو تیم تحقیق در آن‌ها بررسی

می‌شود، همکاری داشته باشند. بعد از کسب اجازه از افراد مورد مطالعه، مجری اصلی طرح می‌تواند حداقل یک مثال کامل از مصاحبه‌ها را بیان کند و یا فرآیند تکنیکی هر عضو تیم تحقیق را نمایش دهد. استفاده از یک چک‌لیست استاندارد در طول مشاهده (که از قبل تهیه شده باشد) مفید است. ارتباط بین مجری اصلی طرح و هر عضو تیم تحقیق به وسیله مرور چک‌لیست آسان می‌شود. پیش‌نویس‌ها برای ثبت نتایج واقعی معاینه و تجزیه و تحلیل نتایج واحد کنترل کیفیت به کار می‌روند.

نگاه‌های کاوشگرانه به تیم تحقیق به‌عنوان بازبین، روحیه دادن، جلب همکاری و مطمئن شدن از به کار بردن یکسان روش‌های استاندارد شده در بین اعضای تیم پژوهش بسیار مفید است. یک مزیت دیگر نگاه‌های کاوشگرانه به سیستم این است که همه اعضای تیم تحقیق احساس می‌کنند در روند کنترل کیفیت سهیم هستند.

– گزارش‌های دوره‌ای

جدول‌بندی داده‌ها برای بررسی کیفیت تکنیکی مراحل بالینی تحقیق و اندازه‌گیری‌ها در فواصل منظم مهم است. با این روش می‌توان اندازه‌های گم‌شده، نادرست و نادقیق را مشخص کرد. مثلاً اگر در یک تیم غربالگری فشارخون، میانگین سطح فشارخون اندازه‌گیری شده در دو ماه متوالی تفاوت داشته باشد می‌توان نتیجه گرفت روش‌های اندازه‌گیری متفاوتی به کار رفته است و به‌طور مشابه تغییرات گسترده در انحراف معیار اندازه‌ها در ماه‌های متوالی بیانگر نامناسب بودن تکنیک اندازه‌گیری است.

کنترل کیفیت اغلب شامل آزمون‌هایی برای سنجش اعتبار و تکرارپذیری اندازه‌گیری‌های مطالعه است. برای مثال جمع‌آوری دو مجموعه داده از یک سری افراد برای آزمون وجود توافق بین اندازه‌ها، چک کردن سازگاری درونی در میان داده‌های مربوط به هم و استفاده کردن از معیار مرجع برای معیار ساختن اندازه‌گیری‌ها در نمونه‌ای از افراد مورد مطالعه، مراحل هستند که برای اطمینان از کیفیت تحقیق انجام می‌شوند (به فصل چهار مراجعه شود).

۱۰-۲-۵ روش‌های ویژه برای مداخله‌های دارویی

لازم است در کارآزمایی‌های بالینی مربوط به استفاده از داروها به‌خصوص آن‌ها که کور هستند، نظارت دقیق بر رعایت دستور مصرف دارو وجود داشته باشد. این موضوع به‌وسیله برنامه‌ریزی دقیق با شرکت سازنده دارو، کنترل نحوه توزیع دارو، نظارت بر صحت دارو و دوز مصرفی حاصل می‌شود.

۱۰-۲-۶ هماهنگی برای انجام کنترل کیفیت

در مطالعات بزرگ بهتر است که یکی از اعضای تیم تحقیق مسئول اجرای مناسب روش‌های کنترل کیفیت باشد. هدف از این کار مشخص کردن مشکلات احتمالی قبل از رخداد آن‌ها و جلوگیری از بروز آن‌ها است.

۱۰-۲-۷ کنترل کیفیت فرآیندهای آزمایشگاهی

با استفاده از روش‌های توضیح داده‌شده برای کنترل کیفیت مراحل بالینی، می‌توان کیفیت فرآیندهای

آزمایشگاهی را نیز کنترل کرد (جدول ۱۰-۴).
جابجایی نمونه‌های آزمایشگاهی افراد و ماهیت خطاپذیر روش‌های آزمایشگاهی منجر به اتخاذ راهبردهای خاص در این زمینه می‌شود [۲] و [۳].

جدول ۱۰-۴: کنترل کیفیت مراحل آزمایشگاهی

کنترل کیفیت
مرحله طراحی
مرحله اجرا
راهکارهای موجود در جدول ۱۰-۱ کنترل دوره‌ای تجهیزات امکان کور کردن

- توجه به برچسب‌گذاری

وقتی نمونه آزمایشی خون فردی اشتباهاً با نام فرد دیگری برچسب خورده است، تصحیح و یا کشف خطاهای صورت گرفته تقریباً غیرممکن است. تنها راه‌حل، جلوگیری و اجتناب از خطاهای جابجایی، چک کردن دقیق نام و کد شناسایی افراد تحت مطالعه موقع برچسب‌گذاری نمونه‌های آزمایشگاهی است. با استفاده از یک کامپیوتر می‌توان مجموعه‌ای از برچسب‌های لوله‌های آزمایشگاهی محتوی خون را پرینت گرفت و آنها را ثبت کرد. این روش به روند برچسب‌گذاری سرعت می‌دهد و از جابجایی ارقام جلوگیری می‌کند. وقتی سرم جدا شده از یک لوله آزمایشگاهی به لوله دیگر منتقل می‌شود، بهتر است لوله آزمایشگاهی جدید را از قبل برچسب زده و دو لوله را نزدیک هم نگهداری کرد. در این روش با خواندن یکی دیگری نیز چک می‌شود هم‌چنین می‌توان از بارکدهای اتوماتیک استفاده کرد.

- کور کردن

اگر محقق هنگام اندازه‌گیری نمونه‌هایی که از بیماران گرفته می‌شود اطلاعی از گروه مطالعه یا مقادیر متغیرهای کلیدی دیگر نداشته باشد، کور کردن انجام شده است. حتی در فرآیندهای شبیه تعیین گلوکز خون به‌طور اتوماتیک نیز، این روش باعث کاهش تورش می‌شود و از به کار بردن قضاوت‌های شخصی هنگام گزارش نتایج جلوگیری می‌کند. وقتی در یک مطالعه چندمرکزی نمونه‌ها، نتایج آزمایشگاهی یا تصاویر پزشکی به آزمایشگاه مرکزی فرستاده می‌شوند بهتر است برچسب‌ها به‌صورت دوسوکور باشد. این روش معیاری مناسبی برای ارزیابی دقت روش‌های آزمایشگاهی است و می‌تواند به صورتی طراحی شود که هم سازگاری مشاهده تکی و هم سازگاری بین مشاهدات را آزمون کرد.

۱۰-۲-۸ کنترل کیفیت داده‌ها

محقق باید قبل از آغاز مطالعه، سیستم مدیریت داده‌ها را پیش‌آزمون کند. این فرایند شامل طراحی فرم‌ها برای ثبت اندازه‌ها، انتخاب بانک اطلاعاتی برای ویرایش و مدیریت داده‌ها و طراحی جداول خالی برای اطمینان از اینکه همه متغیرهای تحقیق جمع‌آوری می‌شوند، است. جدول ۱۰-۵ مراحل کنترل کیفیت داده‌ها را فهرست کرده است.

- طراحی فرم‌ها

طراحی فرم‌های جمع‌آوری داده‌ها، تأثیر زیادی روی کیفیت داده‌ها دارد. طراحی سیستم ورود داده‌ها قبل از شروع مرحله ورود اطلاعات، به محقق اجازه می‌دهد پاسخ‌ها را در طول مرحله طراحی استاندارد کند و در مرحله ورود داده‌ها خطاها را کاهش دهد. همچنین باعث می‌شود که محقق از به کار بردن سوالاتی که کذبندی یا تفسیر آن‌ها مشکل است اجتناب کند. اندازه‌گیری‌های ذهنی (نظیر تشخیص‌های کلینیکی) نیاز به تعریف عملیاتی واضح دارند. این تعاریف باید به صورت دقیق و واضح در دستورالعمل عملیاتی و فرم‌های مربوط به ثبت اندازه‌گیری آورده شوند. گزینه‌ها باید دارای ارتباط و توالی منطقی باشند. فرم‌های خوانا و ترغیب‌کننده باعث می‌شود افراد تحت مطالعه توجه بیشتری به تکمیل کردن فرم‌ها داشته باشد. پیش‌آزمون باعث می‌شود از وضوح سوالات و آسانی دستورالعمل تکمیل فرم اطمینان حاصل شود. برچسب زدن به هر صفحه با تاریخ، نام و شماره شناسه پرونده افراد باعث می‌شود اطلاعات گم نشوند.

جدول ۱۰-۵: کنترل کیفیت داده‌ها در مرحله طراحی

کنترل کیفیت داده‌ها (مرحله طراحی مطالعه)
راهکارهای موجود در جدول ۱۰-۱
تنها متغیرهای موردنیاز باید جمع‌آوری شوند
تهیه جدول‌های توخالی برای طراحی مرحله تجزیه و تحلیل داده‌ها
تهیه فرم‌های اطلاعاتی که دارای ویژگی‌های زیر باشند:
قابل فهم
سازگار
واضح و منظم
چاپ مناسب
معتبر شده
برچسب‌گذاری شده در هر صفحه
تهیه بانک اطلاعاتی که به مشاهدات خارج از محدوده، گم‌شده و ناسازگار حساس باشند.

- جمع‌آوری و مدیریت داده‌ها

مدیریت منسجم و ساختارمند، قسمتی مهم در کنترل کیفیت بالا بردن کارایی و صحت داده‌ها است. در اندازه‌گیری‌های بالینی، کنترل کیفیت و سازگاری داده‌ها باید هنگامی انجام شود که فرد هنوز در دسترس است [۹]. یکی از اعضای تیم تحقیق باید فرم‌ها را ویرایش کند و فرد دیگری کامل و درست بودن اطلاعات را بازبینی کند (جدول ۱۰-۶). اگر ویرایش کننده در جمع‌آوری داده‌ها شرکت نداشته باشد احتمال کشف مقادیر گم‌شده و خطاهای دیگر بیشتر می‌شود.

یک روش دیگر، ویرایش کامپیوتری است. در این روش کامپیوتر به طریقی برنامه‌ریزی می‌شود که داده‌های گم‌شده، اطلاعات ناسازگار و مقادیر خارج از دامنه به‌طور اتوماتیک شناسایی شود. این روش برای مطالعات کوچک مناسب نیست. روش جایگزین انتخاب یک نمونه تصادفی از فرم‌های اصلی و تهیه مجدد جدول داده‌های اولیه و مقایسه این دو با هم است.

جدول ۱۰-۶: کنترل کیفیت داده‌ها در مرحله اجرا مطالعه

کنترل کیفیت داده‌ها (مرحله اجرای مطالعه)
اطمینان از صحیح بودن نام، شماره شناسایی، روز انجام معاینه در هر صفحه
اطمینان از تکمیل شدن فرم‌ها برای هر قسمت از معاینات
اطمینان از جا نیفتادن قسمت‌های مختلف پرسشنامه
کنترل صحت داده‌های واردشده به‌وسیله بازبینی تصادفی
تهیه جداول فراوانی دوره‌ای برای کشف خطاهای احتمالی

جدول‌بندی‌های دوره‌ای

بررسی توزیع فراوانی متغیرهای مهم در فواصل زمانی منظم به محقق امکان می‌دهد کامل بودن و کیفیت داده‌ها را بررسی کند. این امر باعث می‌شود تصحیح خطاهای قبلی ممکن شود و خطاهای احتمالی در ادامه مطالعه قابل پیشگیری باشد.

۱۰-۲-۹ کنترل کیفیت در مطالعات چندمرکزی

در بسیاری از تحقیقات حجم نمونه موردبررسی بسیار بالا است و بیشتر از تعداد نمونه‌ای است که در یک مرکز قابل‌دسترسی است. این موضوع به انجام مطالعات چندمرکزی منجر می‌شود. این مطالعات به‌وسیله تیم‌های تحقیق در مراکز مختلف انجام می‌شود. گاهی اوقات این مراکز در یک شهر یا یک منطقه هستند و یک محقق می‌تواند همه تیم‌های تحقیق را سرپرستی کند [۱۰]. با این حال اغلب مطالعات چندمرکزی به‌وسیله چند محقق در شهرهای مختلف انجام می‌شود. در این نوع مطالعات مراحل خاصی لازم است تا همه مراکز به یک روش، مطالعه را انجام دهند و داده‌های قابل‌مقایسه و ادغام پذیر برای تحلیل و انتشار نتایج فراهم شود. یک مرکز

هماهنگ‌کننده باید ارتباط بین مراکز مختلف را میسر کند، دستورالعمل‌ها را هماهنگ کند و جنبه‌های دیگر کنترل کیفیت، آموزش کارکنان هر مرکز، بازمینی، مدیریت داده‌ها، تحلیل و انتشار را به عهده بگیرد.

یک خطای متداول، گرایش به جمع‌آوری داده‌های اضافی است. توجه این موضوع که مرحله معاینه بالینی اولیه در هنگام ورود بیماران به مطالعه، تنها زمانی است که محققان شانس اندازه‌گیری متغیرهای زمینه‌ای را دارند، باعث می‌شود محققان هر متغیر جالب‌توجه (مثلاً ثبت داده‌های معمولی حاصل از معاینه بدنی که ارتباط ناچیزی با سؤال تحقیق دارد) را ثبت کنند. معمولاً محققان داده‌های بسیار زیادی نسبت به آنچه تحلیل و یا منتشر می‌شود، جمع‌آوری می‌کنند. مشکل این روش، صرف بیهوده زمان جهت اندازه‌گیری پارامترهای کم‌اهمیت است. افراد تحت مطالعه خسته و ناراحت می‌شوند و داده‌های مهم‌تر کیفیت خود را از دست می‌دهند. مشکل دیگر حجم زیاد و پیچیدگی داده‌ها است که کنترل کیفیت و تحلیل داده‌ها را مشکل‌تر می‌سازد. بهتر است نیاز به جمع‌آوری متغیرها بررسی شود و بسیاری از متغیرهایی که به دلخواه در نظر گرفته شده‌اند حذف شود. کاهش متغیرهای اضافه اعتبار و دقت متغیرهای مهم را افزایش می‌دهد.

۳-۱۰ بهبود پروتکل هنگام اجرای مطالعه

صرف‌نظر از نتایج پیش‌آزمون، مشکلات طراحی، اجرا و اندازه‌گیری پارامترها وقتی مطالعه آغاز شده باشد ظاهر می‌شوند. قانون کلی، اعمال کمتر تغییرات بعد از آغاز مطالعه است. به‌هرحال گاهی اوقات بهبود پروتکل می‌تواند باعث اعتبار بخشیدن به مطالعه شود.

۱۰-۳-۱ تغییرات جزئی

تصمیم برای آنکه آیا تغییرات جزئی باعث بهبود مطالعه می‌شود یا نه، اغلب تعادلی بین مزایای روش‌شناسی پروتکل بهبودیافته و عوض کردن یکنواختی مطالعه، صرف زمان و هزینه برای تغییر سیستم است. تصمیماتی که یک تعریف عملیاتی را دقیق‌تر می‌سازد نسبتاً ساده هستند. مثلاً در "مطالعه‌ای که سیگار کشیدن یکی از معیارهای خروج است برای افرادی که سیگار را ترک کرده‌اند چه تصمیمی باید گرفته شود؟" این تصمیم باید به‌طور یکنواخت برای تمام افراد و در تمام مراکز انجام شود.

۱۰-۳-۲ بهبود اساسی

تغییرات بزرگ‌تر پروتکل مطالعه از قبیل تغییر در معیارهای ورود و خروج بیماران یا تغییر مداخله مشکلات جدی‌تری را به دنبال دارد. اگرچه ممکن است دلایل خوبی برای اعمال این تغییرات وجود داشته باشد اما اگر این تغییرات منجر به تفسیر متفاوت یافته‌ها شود باید تحلیل و گزارش داده‌ها به‌طور جداگانه موردبررسی قرار گیرد.

منابع

1. Townsend MC, Morgan J, Durkin D, et al. Quality control aspects of pulmonary function testing in the MRFIT. *Control Clin Trials*. 1986, 7:1795-1925.
2. Canner PL, Krol WF, Forman SA. External quality control programs. *Control Clin Trials*. 1983, 4:441-66.
3. Widdowson BM, Kuehneman M, Duchene AG, et al. Quality control of biochemical data in the MRFIT. *Control Clin Trials*. 1986, 7:175-335.
4. The MRFIT Research Group. MRIFT: Risk factor changes and mortality results. *JAMA*. 248:1465-77 1982.
5. The MRFIT Research Group. CHD death, nonfatal acute MI and other clinical outcomes in the MRFIT. *Am J Cardiol*. 1986, 58:1-13.
6. Brownson RC, Petitti DB. *Applied epidemiology*. Oxford University Press, London. 2000.
7. Szklo M, Nieto FJ. *Epidemiology, beyond the basics*. An Aspen Publication, Gaithersburg, Maryland. 2000.
8. Gordis L. *Epidemiology*. Elsevier Inc. Pennsylvania, 2004.
9. Mohammadi MR, Soori H, et al. Selegiline in comparison with Methylphenidate in attention deficit hyperactivity disorder children and adolescents in the double-blind, randomized clinical trial. *J of Child And Adolescent Psychopharmacology*. 2004;14 (3): 426-33.
10. Mohammadi MR, Soori H, et al. An Epidemiological Survey of Psychiatric Disorders in Iran. *Clin Pract Epidemiol Ment Health*. 2005 Sep 26;1(1):1-16.

ضمیمه الف

تنظیم پروپوزال و تامین منابع تحقیق

پروپوزال یا پیشنهادیه تحقیقاتی، قسمت مکتوب یک مطالعه است که نوشتن آن به محقق توان تنظیم، وضوح و بهبود همه عناصر مطالعه را می‌دهد. به این ترتیب برخی مشکلات عملی پروژه بهبود می‌یابد. به علاوه حتی اگر برای مطالعه نیازی به منابع مالی نباشد، پروپوزال جهت راهنمایی در اجرا لازم است. پروپوزال سند مکتوبی برای جذب منابع مالی از مراکز تحقیقاتی است. پروپوزال شامل قسمت‌های فنی مطالعه، پروتکل مطالعه، بودجه، جداول زمان‌بندی اجرای طرح و اطلاعات تکمیلی دیگر است. در این بخش ساختار یک پروپوزال موفق مورد بحث قرار می‌گیرد.

الف-۱ تنظیم پروپوزال‌ها

بخش مشکل آماده کردن یک پروپوزال، تنظیم، نوشتن و بازبینی آن است که نیاز به صرف چند ماه وقت دارد. گام‌های زیر می‌تواند برای رسیدن به یک آغاز خوب مفید باشد.

۱- در مورد این که پروپوزال به کجا فرستاده می‌شود تصمیم بگیرید. هر مرکز تحقیقاتی روش‌های منحصر به فردی جهت جذب پروپوزال دارد. بنابراین محقق باید با تصمیم‌گیری در مورد جایی که پروپوزال فرستاده می‌شود، تعیین محدودیت‌های مالی و تشریح اهداف پروپوزال برای مراکز تحقیقاتی مورد نظر، شروع کند.

۲- تشکیل تیم تحقیق و انتخاب یک سرپرست. بیشتر پروپوزال‌ها به وسیله تیمی شامل چند نفر که تجربیات زیادی در طراحی و اجرای تحقیقات علمی دارند، نوشته می‌شود. این تیم ممکن است کوچک (فقط محقق و یک مشاور) یا بزرگ باشد (شامل همکاران، یک مشاور آمار حیاتی یک مسئول امور مالی و مسئول پشتیبانی). این تیم باید توانایی و مهارت مورد نیاز برای طراحی و آغاز مطالعه را داشته باشد. یکی از اعضای تیم باید مسئول اجرای طرح باشد. معمولاً این فرد محقق اصلی است. محقق اصلی طرح‌های تحقیقی، باید فردی باتجربه که اطلاعات و تدبیر او برای تصمیم‌گیری‌ها کافی و مفید است باشد. احتمالاً سابقه و مطالعات قبلی اعضای تیم تحقیقاتی شانس تأمین بودجه را افزایش می‌دهد. در بعضی مطالعات یک مدیر پروژه در نظر گرفته می‌شود. این فرد معمولاً محقق تازه‌کاری است که وظیفه‌اش مدیریت روزانه و

هماهنگ کننده برنامه‌های تنظیم پروپوزال است.

گاهی اوقات مدیر پروژه همان محقق اصلی مطالعه است. وظایف محقق اصلی یا مدیر پروژه عبارت‌اند از: هدایت منظم مطالعه، تهیه و تنظیم پروپوزال، تنظیم جلسات منظم تیم تحقیق و اطمینان حاصل کردن از این که همه کارها به‌طور کامل و به‌موقع انجام می‌شود.

۳- پیگیری دستورالعمل برای دریافت کمک‌های مالی. تمام مراکز تحقیقاتی دستورالعمل مکتوبی

دارند که محقق باید قبل از نوشتن پروپوزال آن را به‌دقت مطالعه کند. این اطلاعات شامل طرح کلی تنظیم پروپوزال، محدودیت صفحات، اطلاعاتی در مورد مقدار بودجه‌ای که می‌تواند درخواست شود و اصولی که باید در تهیه پروپوزال رعایت شود، است؛ اما این دستورالعمل شامل همه اطلاعات مهمی که محقق برای شناختن عملکردها و ملزومات دریافت کمک‌های مالی لازم دارد نیست. مراکز تحقیقاتی یک تیم علمی دارند که وظیفه ارزیابی پروپوزال‌ها را بر عهده دارند. هنگام تهیه پروپوزال بهتر است در مورد موضوع طرح با فردی که مسئول هماهنگی درخواست‌ها است، بحث‌هایی صورت گیرد. به این ترتیب نکات موردنظر مؤسسه (از قبیل محدودیت‌های بودجه، هدف و توضیحات موردنیاز در پروپوزال) به‌روشنی مطرح می‌شود و از این که طرح تحقیق در محدوده درخواست‌های مؤسسه است، اطمینان حاصل می‌شود. رد یک پروپوزال به دلیل عدم توجه به جزئیات، تجربه‌ای ناکام و اجتناب‌ناپذیر است.

۴- تهیه جدول زمانی و برقرار کردن جلسات منظم. تهیه جدول زمانی برای تکمیل کردن کارها،

باعث می‌شود اعضای تیم تحقیق تحت فشار باشند تا سر وقت وظایف خود را ارائه کنند. بهتر است قبل از این که پروپوزال به یک مؤسسه پژوهشی فرستاده شود، بودجه آن دوباره با قسمت مالی مؤسسه بازبینی شود. عدم توجه به این جزئیات باعث به وجود آمدن بحرانی می‌شود که به پروپوزال‌های خوب انجام شده دیگر آسیب می‌رساند.

۵- پیدا کردن یک پروپوزال نمونه. بهتر است یک نسخه از پروپوزال‌های موفق مؤسسه‌ای که قرار است

طرح را مورد حمایت مالی قرار دهد، تهیه شود. پروپوزال‌های موفق روش مطمئنی برای آگاه شدن از نحوه نگارش و محتوی یک پروپوزال خوب است. محقق می‌تواند بهترین نظرات را به دست آورد و پروپوزالی طراحی کند که واضح‌تر، منطقی‌تر و متقاعدکننده‌تر باشد. همچنین بررسی گزارش‌های مکتوبی که مؤسسه برای همه پروپوزال‌های موفق و ناموفق تهیه کرده است مفید است.

برای جستجوی فرصت‌های مالی می‌توان پروپوزال‌های جالب را با استفاده از اینترنت پیدا کرد.

۶- تنظیم ساختار پروپوزال. مراحل تنظیم پروپوزال در جدول الف-۱ ارائه شده است. فرم پروپوزال

ساختار بندی شده در تقسیم مسئولیت‌های قسمت‌های مختلف پروپوزال مفید است.

۷- بازبینی، پیش‌آزمون و اصلاح مداوم. نوشتن پروپوزال یک فرایند تدریجی است معمولاً نسخه‌های

زیادی تهیه می‌شود که هر کدام منعکس کننده ایده‌ها، توصیه‌ها و تجربیاتی است. قبل از اینکه نسخه نهایی نوشته شود پروپوزال باید به‌وسیله همکارانی که با موضوع آشنا هستند به‌طور جدی بازبینی شود. باید به کیفیت سؤال تحقیق، اعتبار طرح، روش‌های به‌کاررفته و نکات ویرایشی توجه خاص شود.

بہتر است قبل از ارائه، پروپوزال به طور جدی مورد بررسی نقادانه قرار بگیرد. در غیر این صورت پروژه به دلیل عجله و مشکلات نوشتاری رد می‌شود. وقتی پروپوزال تقریباً آماده شد گام نهایی، بازبینی دقیق جهت سازگاری داخلی پروپوزال، انطباق با دستورالعمل مرکزی که پروپوزال به آنجا ارائه می‌شود و اصلاح خطاهای ویرایشی (نوشتاری) است.

الف-۲ عناصر تشکیل دهنده یک پروپوزال

عناصر تشکیل دهنده یک پروپوزال در جدول الف-۱ ارائه شده است. ممکن است بعضی مؤسسات اطلاعات کمتر و یا ساختار متفاوتی نیاز داشته باشند. محقق باید پروپوزال را بر اساس مؤسسه‌ای که حمایت مالی را به عهده دارد تنظیم کند. این موارد باید برای همه مطالعات نوشته شود حتی اگر مؤسسه مالی به آن نیاز نداشته باشد.

الف-۲-۱ مرحله شروع

عنوان باید مختصر، مفید و واضح باشد. عنوان مناسب باعث به وجود آمدن اولین تأثیر مناسب می‌شود و یادآوری مداومی از محتوی و طرح مطالعه است. یک عنوان خوب باعث خلاصه کردن محتوی مطالعه می‌شود. چکیده، خلاصه مفیدی از مطالب پروپوزال است. بهتر است با سؤال تحقیق شروع شود. سپس طرح و روش‌ها بیان شوند و با جمله مهمی از یافته‌های اساسی مطالعه به پایان برسد. بیشتر مؤسسه‌ها برای چکیده محدودیت کلمات قائل می‌شوند. بنابراین بهتر است از عبارات مؤثر و روشن استفاده شود. به طور کلی بعد از اینکه قسمت‌های دیگر پروپوزال آماده شد چکیده نوشته می‌شود و سپس مورد بازبینی قرار می‌گیرد. چکیده تنها صفحه‌ای است که توسط اغلب داورها خوانده می‌شود و باید به راحتی نماینده طرح اصلی پروپوزال باشد تا برای هر فرد دیگری غیر از داور هم قابل استفاده شود. فهرست مطالب کمک اساسی برای خواننده است. پروپوزال‌ها ممکن است شامل بخش‌های بزرگ‌تر، پیچیده‌تر یا غیرمعمول باشند. در این موارد باید فهرست مطالب ذکر شود و با فصل‌ها یا زیر فصل‌های موجود در متن مطابقت داشته باشد.

الف-۲-۲ بخش‌های اجرایی

تقریباً همه مؤسسه‌ها خواهان یک بخش اجرایی هستند که شامل بودجه، ارائه تخصص‌های کارکنان، مؤسسه ارائه دهنده پروپوزال میزان دسترسی به تجهیزات، فضای مورد نیاز و مهارت‌ها است. بخش بودجه معمولاً مطابق با راهنمایی‌های مؤسسه مالی تنظیم می‌شود. فرم بودجه نباید در آخرین قسمت طراحی پروپوزال تهیه شود. بسیاری از موارد نیازمند گذشت زمان هستند (مثلاً برای به دست آوردن برآورد مناسبی از هزینه‌ها، فضا مورد نیاز مهارت و کارکنان گذشت زمان مهم است).

جدول الف-۱: عناصر اصلی نوشتن پروپوزال

عنوان
چکیده
فهرست مطالب
فرم بودجه بندی
معرفی محققان مطالعه
منابع مورد نیاز، تجهیزات فنی و فضای فیزیکی مورد نیاز
اهداف ویژه
اهمیت انجام مطالعه
مرور منابع
روش‌ها
بازبینی طرح مطالعه چار چوب زمانی انجام مطالعه افراد تحت مطالعه معیارهای انتخاب طرح نمونه‌گیری نحوه دستیابی به اعضای نمونه
اندازه‌گیری‌ها
متغیرهای مستقل اصلی (مداخله در صورت انجام مطالعه تجربی) متغیرهای پاسخ مخدوش‌کننده‌های بالقوه
طرح مطالعه راهنما
تجزیه و تحلیل آماری
روش تحلیل داده‌ها آزمون فرض‌ها و برآورد حجم نمونه
نکات اخلاقی مورد توجه
هماهنگی و مشاوره با مؤسسات شرکت کننده در مطالعه
فهرست منابع
ضمایم

پس از مشخص شدن طرح کلی پروپوزال سرپرست می‌تواند کار خود را شروع کند و مقدار بودجه را پیشنهاد کند. مقررات و فرصت زمانی ارائه پروپوزال به هر مؤسسه باید رعایت شود. سرپرست با تجربه می‌تواند به محقق کمک کند تا همیشه بر قوانین اداری، اشکالات غیرمنتظره و دیر کردهای احتمالی فائق آید. هم‌چنین سرپرست می‌تواند در انجام پیش‌نویس متن بخش‌های مربوط به بودجه و منابع جمع‌آوری ضمیمه‌ها و سایر موارد کمک مؤثری باشد.

باید در فرم بودجه همه هزینه‌ها کاملاً واضح باشد. معمولاً حقوق شامل بیشترین هزینه طرح تحقیقاتی بالینی است. بنابراین باید برای همه افراد حقوق و ساعت انجام وظیفه در فهرست بودجه مشخص شود. دقت شود که شرح وظایف بیان شده برای محقق و اعضای دیگر تحقیق نباید در ذهن داوران تردیدی به‌جای بگذارد و باید یقین حاصل شود که کار برآورد شده برای هر فرد فقط در جهت موفقیت طرح است.

اغلب داوران به درصد هزینه اختصاص داده شده به اعضای اصلی پروژه توجه می‌کنند. به‌ندرت اعضای که بر اساس فرم بودجه مسئولیت خیلی کمی دارند مورد ارزیابی قرار می‌گیرند. (چون آن‌ها وظایف خیلی زیاد دیگری دارند و نمی‌توانند برای مطالعه پیشنهادی وقت زیادی صرف کنند). اگر نیازهای مطالعه تغییر کنند یا هزینه‌های غیرمنتظره وجود داشته باشد، ممکن است بهترین طرح بودجه نیز تغییر کند. اگر محقق بخواهد هزینه بعضی موارد را تغییر دهد یا وظایف محققان اصلی را کم یا زیاد کند باید از مرکز تحقیقاتی تأمین‌کننده بودجه پژوهش کسب اجازه نماید. معمولاً مؤسسه‌ها تا زمانی که محقق خواهان افزایش حمایت مالی نباشد با درخواست‌های معقول برای تغییر بودجه‌بندی موافقت می‌کنند. مشخصات محققان معمولاً در صفحه دوم پروپوزال ذکر می‌شود که شامل تحصیلات دانشگاهی، شغل فعلی و قبلی و افتخارات است و باید به ترتیب زیر تهیه شود: عنوان همه تألیفات، همه مقالات و کتاب‌های منتشر شده در سه سال اخیر.

تجهیزات و مواد لازم برای پروژه شامل رایانه و وسایل فنی، مواد و تجهیزات آزمایشگاهی، فضای اداری و کتابخانه باید به‌طور کامل توضیح داده شود. این موارد معمولاً در پروپوزال‌هایی که قبلاً نوشته شده در دسترس است.

الف - ۲-۳ بخش‌های اهداف و مرور منابع

اهداف ویژه از نظر محتوی همان سؤالات تحقیق هستند اما آن طوری که از نتایج تحقیق انتظار می‌رود، بیان می‌شوند. هر هدف باید با یک یا دو جمله توضیح داده شود. بیشتر پروپوزال‌ها اهداف متعددی دارند که باید به ترتیب و توالی زمانی مطرح شوند. منطقی‌ترین روش این است که ابتدا اهداف توصیفی سپس اهداف تحلیلی را بیان کنیم.

از بخش اهداف ویژه برای تنظیم بخش‌های دیگر می‌توان استفاده کرد. بهتر است بخش‌های مرور منابع و روش‌ها پشت سر هم آورده شود.

وقتی در یک مطالعه جنبه‌های زیادی مورد بررسی قرار می‌گیرد بهتر است فهرست کاملی از اهداف ویژه در اختیار خواننده قرار گیرد. اگر اهداف ویژه زیاد باشند بهتر است فقط مهم‌ترین و جالب‌ترین آن‌ها بیان شود. در هر صورت تعداد صفحات نباید زیاد باشد. بخش مرور منابع، پروپوزال را از نظر محتوی پر بار می‌کند. جزئیات این

بخش باید تا آنجا که ممکن است مکتوب باشد. به این ترتیب برای شخصی که در آن زمینه خبره نیست قابل درک می‌شود. باید در مورد این که این مطالعه خاص چگونه انجام می‌شود و چرا انجام آن اهمیت دارد، اطلاعات کافی ارائه شود. هدف این بخش آگاه کردن محقق از چگونگی انجام تحقیق، مشکلات و نیازها است. بازبینی همه‌جانبه یا جزئی، به تمرکز اهداف ویژه، پیچیدگی موضوعات و غیره بستگی دارد. معمولاً بازبینی ۲۰ یا ۳۰ منبع از جدیدترین و بهترین منابع کافی است. در بسیاری از پروپوزال‌ها انجام یک مطالعه راهنما که موجه بودن سؤال تحقیق و عملی بودن مطالعه را تأیید کند بسیار مهم است.

الف-۲-۴ بخش روش‌های علمی

داورها، بخش روش‌ها را موشکافانه بررسی می‌کنند. یک دلیل متداول رد شدن پروپوزال ضعف در روش‌های تکنیکی است.

اولین مسئله چگونگی تنظیم این بخش است. گاهی اوقات مراکز تحقیقاتی در مورد چگونگی تنظیم روش‌ها راهنمایی‌هایی تهیه می‌کنند در غیر این صورت، فهرست موجود در جدول الف-۱ را پیشنهاد می‌کنیم. در این مرحله یک جدول خلاصه در مورد محتوای بخش روش‌ها بسیار مفید است و بازبینی طرح همراه با ارائه نمودارها و جداول جهت در جریان قرار دادن خواننده بسیار مفید است.

بهتر است بخش آماری با روش‌های تحلیل داده‌ها آغاز شود. روند منطقی آن می‌تواند به صورت زیر باشد. ابتدا جداول توصیفی، سپس روش تحلیل روابط بین متغیرها و بعداً این مرحله مسئله حجم نمونه مطرح می‌شود که بهتر است با جمله فرض صفر و انتخاب آزمون آماری مناسب، قبل از تعیین حجم نمونه شروع شود و با استفاده از α و β حجم مؤثر نمونه برآورد شود. بخش اجرایی پروپوزال باید شامل جدول زمان‌بندی باشد یعنی مشخص شود که فازهای اصلی مطالعه چه وقت شروع و چه وقت پایان می‌یابند.

الف-۳ شاخص‌های یک پروپوزال خوب

ویژگی‌های یک پروپوزال خوب عبارت‌اند از:

- ۱- **کیفیت علمی طرح تحقیق:** طرح تحقیق باید بر اساس یک سؤال تحقیق مناسب تهیه شود. طرح و روش‌های انجام آن دقیق و عملی باشد. تیم محقق تجربه مهارت کافی داشته و متعهد به انجام تحقیق باشد.
 - ۲- **کیفیت فنی پروپوزال:** ارائه مناسب، یکی از مهم‌ترین عوامل تعیین‌کننده‌های سرنوشت پروپوزال است. حتی اگر سؤال تحقیق مهم و طرح مطالعه عالی باشد، ارائه ضعیف باعث می‌شود داوران، پروپوزال را رد کنند. پروپوزال باید مختصر و مفید و جذاب باشد و با نوشته‌های مبهم در مورد موضوعات فرعی تمرکز داوران را از بین نبرد. پروپوزالی که به خوبی طراحی شده، به دقت نوشته شده، به طور جذابی ارائه شده و عاری از اشتباه باشد خواننده را از محتوای تحقیق مطمئن می‌سازد.
- بهتر است هنگام نوشتن پروپوزال به مشخصات فردی بازبین کننده توجه داشته باشیم. اغلب داورها درگیر

انبوه زیادی از پروپوزال‌های طولانی هستند، بنابراین باید نکات برجسته پروژه مشخص شود تا با خواندن سریع و بدون دقت از قلم نیافتد می‌توان با ارائه نتایج واضح، بخش‌های کوتاه با زیر فصل‌های با معنی، خلاصه‌های دقیق و کوتاه، جداول مختصر و مفید و نمودارهای ساده توجه داور را به نکات بسیار مهم پروپوزال جلب کرد. بهتر است در صفحات فضا‌های سفید و خالی وجود داشته باشد.

بیشتر بازیکنان بسیار باتجربه هستند و به جملات اضافی و فرم‌های متعدد درخواست بودجه توجهی نمی‌کنند. پروپوزالی که در آن اهمیت پروژه اغراق‌آمیز باشد، باعث به وجود آمدن شک و تردید می‌شود. با شوق و اشتیاق نوشتن ایده خوبی است، اما بهتر است محقق در مورد محدودیت‌های پروژه واقع‌بین باشد. بیشتر داورها در امر تعیین مشکلات جزئی طرح یا عملی بودن طرح پروژه تحقیق خیره هستند.

محقق به جای حذف نواقص اساسی می‌تواند آن‌ها را صراحتاً بیان کند و معایب و مزایای طرح انتخابی را توضیح دهد. با این حال توجه زیاد به این مسائل کاری اشتباه است، زیرا ممکن است جنبه‌های ضعیف پروپوزال مورد توجه واقع شود و نکات قوت پروپوزال نادیده گرفته شود. هدف این است که داوران مطمئن شوند محقق از بروز مشکلات اساسی جلوگیری می‌کند و روش متفکرانه و واقع‌بینانه‌ای در برخورد با آن‌ها دارد.

ضمیمه ب

طراحی پرسشنامه و هدایت مصاحبه

ابزار سنجش و اندازه‌گیری وسایلی هستند که محقق به کمک آن‌ها می‌تواند اطلاعات موردنیاز را برای تجزیه و تحلیل و بررسی پدیده موردنظر و نهایتاً کشف حقیقت گردآوری نماید.

سه روش اصلی برای جمع‌آوری داده‌هایی در مورد رفتارها، اطلاعات سلامتی و تاریخچه شخصی وجود دارد: مشاهده، پرسشنامه و مصاحبه. مشاهده مشهورترین و طبیعی‌ترین ابزار جمع‌آوری اطلاعات است. پرسشنامه ابزاری است که پاسخ‌دهنده آن را تکمیل می‌کند ولی مصاحبه به‌وسیله مصاحبه‌کننده اجرا می‌شود.

استفاده از روش مشاهده در طی سالیان متوالی و با پیشرفت علم و اختراع دستگاه‌های دقیق روز به روز دقیق و کارآمدتر می‌شود. یکی از مهم‌ترین دلایل استفاده از روش مشاهده این است که با این روش، پدیده و یا رفتار تحت مطالعه به‌طور مستقیم و زنده بررسی می‌شود.

هنگامی که پدیده موردبررسی، باید در ابعاد وسیع بررسی شود، نتوان روش‌های دیگر را به کار برد (نظیر مطالعه رفتار کودکان)، حضور محقق یا پرسشگر بر پاسخ افراد تحت مطالعه تأثیر بگذارد، این روش بسیار مناسب است.

از مشکلات این روش وابستگی بسیار زیاد به برداشت و استنباط شخصی مشاهده‌گر است. این موضوع به‌ویژه در مورد رفتارهای انسانی که انگیزه‌های متفاوت می‌توانند موجب رفتارهای مشابه شود، نمود پیدا می‌کند.

زمان آغاز و پایان مشاهده می‌تواند تأثیر عمقی بر نتایج به‌دست‌آمده داشته باشد.

در پرسشنامه‌ها از روشی یکسان جهت پاسخگویی به سؤالات ساده از قبیل سن و عادت سیگار کشیدن استفاده می‌شود. پرسشنامه‌ها ارزان‌تر از مصاحبه هستند و به زمان زیادی نیاز ندارند.

مصاحبه‌ها روش بهتری برای جمع‌آوری پاسخ‌های سؤالات پیچیده‌ای که نیاز به توضیح بیشتری دارند، هستند و مصاحبه‌کننده از کامل بودن پاسخ‌ها مطمئن می‌شود. وقتی اعضای شرکت‌کننده توانایی‌های مختلفی برای خواندن و فهم سؤالات دارند، مصاحبه بهتر است. گران بودن، وقت‌گیر بودن و تحت تأثیر قرار گرفتن پاسخ‌ها از روابط بین مصاحبه‌کننده و پاسخ‌دهنده از معایب مصاحبه است.

مصاحبه و پرسشنامه قابل استاندارد شدن هستند اما اجرای مصاحبه هر دفعه کمی تغییر می‌کند. هر دو روش به خطای مربوط به حافظه حساس هستند و با درجات مختلف تحت تأثیر پاسخ‌های موجه اجتماعی هستند.

مصاحبه ممکن است گران یا غیرممکن باشد. وقتی هر دو روش قابل‌اجرا باشد انتخاب بستگی به

مقرون به صرفه بودن و پیچیدگی سؤالات مورد بررسی دارد. مصاحبه و پرسشنامه به دو دسته کلی تقسیم می‌شوند: استاندارد و غیراستاندارد که توسط محقق ابداع و معرفی می‌شود.

ابزارهای استاندارد به دلیل کاربردهای فراوانی که داشته‌اند در طول زمان در تحقیقات مختلف اصلاح و بهبود یافته‌اند و از قابلیت اعتماد و تکرارپذیری مناسبی برخوردار هستند. معمولاً در بیشتر تحقیقات می‌توان یکی از این ابزارهای استاندارد شده برای اندازه‌گیری را پیدا کرد. محققان باید قبل از شروع تحقیق و در هنگام طراحی پروپوزال با جستجو در منابع مرتبط با موضوع، راجع به این موضوع تصمیم‌گیری نماید. هنگامی که ابزارهای استاندارد شده‌ای برای اندازه‌گیری و سنجش وجود داشته باشند استفاده از ابزارهای ساخته‌ی محقق منطقی نیست و در واقع کاربرد ابزار ساخته‌ی محقق منحصر به زمانی است که ابزار استاندارد برای تحقیق وجود نداشته باشد. باین حال موارد زیادی وجود دارد که ابزارهای استاندارد (نظیر پرسشنامه) به اصلاح و تعدیل بر اساس پارامترهایی نظیر فرهنگ، زبان، شرایط ویژه افراد تحت مطالعه و غیره نیاز دارند.

ب-۱ طراحی پرسشنامه و ابزار جمع‌آوری داده‌ها

بسیاری از داده‌های تحقیقات بالینی با استفاده از پرسشنامه یا مصاحبه جمع‌آوری می‌شوند. بر پرسشنامه مجموعه‌ای از پرسش‌های از قبل تعیین شده برای جمع‌آوری اطلاعات (داده‌های بالینی، وضع اجتماعی، گروه‌های شغلی و غیره) است. در بسیاری از مطالعات، اعتبار نتایج به کیفیت این ابزار وابسته است. در این قسمت نحوه طراحی مناسب پرسشنامه را توضیح می‌دهیم.

ب-۱-۱ طراحی سؤالات باز و بسته

دو نوع اصلی سؤالات، سؤالات باز و بسته هستند که برای اهداف مختلفی به کار می‌روند. وقتی شنیدن پاسخ‌های افراد تحت مطالعه برای محققان مهم باشد سؤالات باز به کار می‌روند. مثلاً یک نمونه سؤال باز عبارت است از: "به نظر شما چه عاداتی باعث افزایش خطر حمله قلبی می‌شود؟"

سؤالات باز، باعث می‌شوند محقق با محدودیت‌های کمتری به سؤال پاسخ دهد. این سؤالات در مرحله اولیه بررسی و هنگام طراحی مفید هستند، زیرا محققان را با نحوه نگرش و تفکر افراد تحت مطالعه آشنا می‌کند. عبارات و لغات مورد استفاده در پاسخ‌ها مبنای ساختن گزینه‌های بهتر در مراحل بعدی است. روش‌های کدگذاری و تحلیل پاسخ‌های سؤالات باز نسبت به کدگذاری سؤالات بسته زمان بیشتری را طلب می‌کند که یکی از معایب این نوع سؤالات به شمار می‌آید. سؤالات بسته متداول تر هستند و باعث ساخته شدن معیارهای استاندارد بیشتری می‌شوند. پاسخ این سؤالات از میان چند پاسخ از قبل تعیین شده انتخاب می‌شود. مثلاً یک نمونه سؤال بسته عبارت است از:

به نظر شما کدام یک از موارد زیر بیشتر باعث افزایش خطر حمله قلبی می‌شود. (فقط یک پاسخ صحیح است

□ سیگار کشیدن

□ چاقی

□ نگرانی

از آنجایی که سؤالات بسته فهرستی از پاسخ‌های ممکن را تهیه می‌کنند سریع‌تر و راحت‌تر پاسخ داده و به راحتی تحلیل می‌شوند. در مجموع فهرست پاسخ‌های ممکن، معنی سؤال را واضح‌تر می‌کند. سؤالات بسته در طراحی مقیاس‌های چندگزینه‌ای برای تهیه رتبه‌بندی مفید هستند. معایب سؤالات بسته زیاد است. این سؤالات به پاسخ‌های جهت‌دار منجر می‌شوند. فهرست پاسخ‌ها ممکن است شامل مناسب‌ترین پاسخ نباشد. مشکل اخیر با انجام یک پیش‌آزمون با استفاده از پاسخ‌هایی که به یک سؤال باز داده می‌شود قابل حل است. به هر حال ممکن است مجموعه پاسخ‌ها شامل همه گزینه‌های ممکن نباشد. قرار دادن گزینه‌هایی از قبیل “سایر موارد” یا “هیچ کدام از موارد بالا” در پرسشنامه بسیار مهم است. وقتی یک پاسخ خاص مورد نظر است پاسخ‌های ممکن نباید همپوشانی داشته باشند.

وقتی سؤالی بیشتر از یک جواب دارد داشتن گزینه “همه موارد” مناسب نیست. برای سؤالات چند جوابی بهتر است برای هر گزینه از پاسخ‌های “بله، خیر، نمی‌دانم” استفاده شود. مثال: کدام یک از موارد زیر باعث افزایش خطر حمله قلبی می‌شود؟

نمی‌دانم	خیر	بله	
			چاقی
			نگرانی

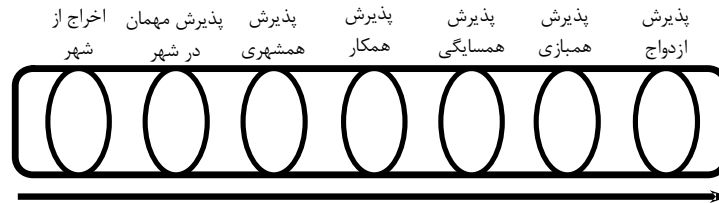
ب-۱-۲ طیف مقیاس پیوسته (VAS)^{۲۳}

طیف مقیاس پیوسته، روش دیگری برای کدگذاری پاسخ‌های سؤالات بسته با استفاده از خطوط یا اشکال دیگر است. در این روش از مقیاس پیوسته‌ای که شامل بدترین حالت تا بهترین حالت است استفاده می‌شود. از افراد تحت مطالعه خواسته می‌شود پاسخ خود را روی این مقیاس علامت‌گذاری کنند. کلمات پایانی باید بیانگر مقادیر اغراق‌آمیز انتهایی باشد. برای راحتی اندازه‌گیری اغلب از خطوط ۱۰ سانتی‌متری استفاده می‌شود و مقیاس فاصله با واحد سانتی‌متر است.

در مقیاس طیف پیوسته به علت اینکه شاخص‌ها روی یک مقیاس پیوسته برآورد می‌شوند، بسیار مورد توجه هستند. این مقیاس نسبت به تغییرات حساس‌تر است. یک روش دیگر تهیه چند دایره به جای یک خط است که ممکن است برای رتبه‌بندی آسان‌تر باشد؛ اما بعضی پاسخ‌ها برای فهمیدن مشکل هستند. بنابراین توضیح و ارائه مثال‌هایی از چگونگی پاسخ به سؤال مهم است. برای مثال در اندازه‌گیری درجه تمایل یا عدم تمایل فرد یا افرادی از گروه «الف» نسبت به گروه‌های اجتماعی دیگر می‌توان طیفی نظیر نمودار ب-۱ در اختیار افراد تحت

۱. The visual analog scale

مطالعه قرار دارد.



جهت افزایش تمایل و علاقه

نمودار ب-۱: سنجش تمایل و علاقه به وسیله طیف

ب-۱-۳ ساختار پرسشنامه

مرسوم است که روی جلد پرسشنامه، با جملات کوتاه هدف مطالعه و چگونگی جمع‌آوری داده‌ها توضیح داده شود. در آغاز مصاحبه نیز اطلاعات مشابه بیان می‌شود. برای اطمینان از درستی و استاندارد بودن پاسخ‌ها همه ابزارهای سنجش باید دستورالعمل خاصی برای تکمیل شدن داشته باشند. این موضوع نه‌تنها برای پرسشنامه‌هایی که توسط افراد تحت مطالعه تکمیل می‌شود باید رعایت شود بلکه برای فرم‌هایی که توسط مصاحبه‌گر به‌صورت شفاهی پرسیده می‌شود نیز باید رعایت شوند. گاهی اوقات یک مثال تکمیل‌شده از یک سؤال ساده مفید است.

برای بهبود پاسخ‌دهی، بهتر است سؤالات برحسب موضوع گروه‌بندی شوند و به‌وسیله جملات کوتاه توضیح داده شوند. بهتر است سؤالات بی‌طرف از قبیل وضعیت سلامتی در ابتدای پرسشنامه باشند. سؤالات حساس‌تر در اواسط پرسشنامه قرار گیرند و سؤالاتی در مورد مسائل شخصی، از قبیل درآمد در اواخر پرسشنامه مطرح شوند. روش پاسخگویی برای هر سؤال یا مجموعه‌ای از سؤالات که با سؤالات دیگر پرسشنامه تفاوت داشته باشد باید به‌طور واضح توضیح داده شود. اگر پرسشنامه شامل چارچوب‌های زمانی مختلف باشد، آوردن چارچوب زمانی در بالای هر مجموعه سؤال جدید مفید است. مثلاً سؤالاتی از قبیل:

- ♦ در طول سال گذشته چند دفعه دکتر شما را ویزیت کرده است؟
 - ♦ در طول سال گذشته چند دفعه در اورژانس یک مرکز پزشکی به‌عنوان بیمار معاینه شده‌اید؟
 - ♦ در طول سال گذشته چند دفعه در یک بیمارستان بستری شده‌اید؟
- می‌تواند به‌صورت زیر خلاصه شوند:
- ♦ در طول سال گذشته چند دفعه الف) دکتر شما را ویزیت کرده است؟
 - ب) در اورژانس یک مرکز پزشکی به‌عنوان بیمار معاینه شده‌اید؟
 - ج) در یک بیمارستان بستری شده‌اید؟

طراحی ظاهری پرسشنامه باید طوری باشد که تا آنجا که ممکن است به‌آسانی به همه سؤالات پاسخ داده

شود. اگر پرسشنامه خیلی پیچیده باشد ممکن است پاسخ‌دهندگان یا مصاحبه‌کنندگان سؤالاتی را از قلم بیندازند یا اشتباه پاسخ دهند که باعث به وجود آمدن داده‌های گم‌شده یا غلط می‌شود و یا حتی ممکن است همه پرسشنامه نامعتبر شود.

یک پرسشنامه منظم با فضای کافی در مقایسه با پرسشنامه‌ای نامنظم و شلوغ بیشتر مورد توجه قرار می‌گیرد و برای استفاده آسان‌تر هستند.

اگرچه محققان اغلب به پرسشنامه‌هایی با تعداد صفحات کمتر علاقه‌مند هستند، اما وقتی سؤالات زیادی در یک صفحه وجود داشته باشد مشکل‌ساز می‌شود. مقیاس‌های پاسخ باید فضای کافی جهت نوشتن عدد صحیح یا کشیدن علامت داشته باشند تا جواب‌های بالا یا پایین به تصادف علامت نخورند. یک سؤال باز باید فضای کافی داشته باشد تا پاسخگو به راحتی در آن فضا پاسخ را بنویسد. پرسشنامه‌هایی که برای افرادی با مشکل بینایی طراحی می‌شوند باید درشت و پررنگ باشد.

پاسخ‌های ممکن برای سؤالات بسته باید زیر هم فهرست شوند، همراه با مربع‌هایی جهت علامت زدن یا عددهایی که دور آن‌ها دایره کشیده می‌شود.

♦ **مثال:** هر روز شما چند داروی مختلف مصرف می‌کنید؟ (فقط یک جواب)

هیچ

۱-۲

۳-۴

۵-۶

۷ یا بیشتر

گاهی اوقات محققان مایل هستند جزئیات دقیق‌تری راجع به افرادی که پاسخ‌های خاصی در بعضی سؤالات دارند کسب کنند. یکی از روش‌های مناسب برای انجام این هدف استفاده از سؤالات دنباله‌دار است. در این نوع سؤالات محققان با توجه به پاسخ‌های اولیه، افراد تحت مطالعه را غربال می‌کنند.

♦ **مثال:** آیا تا به حال به شما گفته شده که فشارخون بالا دارید؟

بله

خیر

اگر بله اولین بار در چه سنی به شما گفته شد که فشارخون بالا دارید؟

..... سالگی

اگر خیر به سؤال ۱۱ بروید.

سؤالات دنباله‌ای در زمان صرفه‌جویی می‌کنند و باعث می‌شوند پاسخ‌دهنده سؤالات نامربوط و تکراری را نادیده بگیرد. راهنمایی پاسخ‌دهنده به سؤال مناسب بعدی، با استفاده از جملاتی از قبیل "به سؤال ۱۱ بروید" انجام می‌شود.

♦ جمله‌بندی

هر لغت در یک سؤال می‌تواند اعتبار پاسخ‌ها را تحت تأثیر قرار دهد. هدف ساختن سؤالات ساده و بدون ابهام است تا افراد تحت مطالعه بتوانند صادقانه و به‌طور صحیح به سؤالات جواب دهند و واهمه‌ای از عواقب احتمالی پاسخ‌ها نداشته باشند.

♦ وضوح

تا جایی که ممکن است سؤالات باید واضح و اختصاصی باشند. به‌طور کلی تعریف لغات باید در فهرست لغات توضیح داده شود. مثلاً برای اندازه‌گیری میزان فعالیت‌های ورزشی در سؤال "در طول یک هفته خاص چند ساعت را به ورزش اختصاص می‌دهید؟" بهتر است سؤال زیر را جایگزین کنید: "معمولاً چقدر ورزش می‌کنید؟"

♦ سادگی

سؤالات باید ساده باشند و در آنها لغات متداول استفاده شود و از به کار بردن اصطلاحات تخصصی و پیچیده اجتناب شود. مثلاً مردم به راحتی متوجه جمله "کدام داروها را بدون نسخه پزشک می‌توانید بخرید؟" می‌شوند ولی با اصطلاح "مصرف خودسرانه دارو" آشنایی کمتری دارند. جملات باید کوتاه و ساختار دستوری ساده داشته باشند. از به کارگیری لغات القایی و کلمات کلیشه‌ای که باعث می‌شود جواب دلخواه القا گردد باید اجتناب شود. در سؤال "در ماه گذشته چند دفعه به مقدار زیاد نوشابه نوشیده‌اید؟" عبارت "به مقدار زیاد" القاکننده اثر منفی نوشابه قندی بر سلامتی است و بهتر است سؤال به صورت زیر مطرح شود: "طول ماه گذشته چند دفعه در طول روز بیشتر از پنج لیوان نوشابه نوشیده‌اید؟" [۶].

گاهی اوقات بهتر است جملات را با لحنی آغاز کرد تا پاسخگو تصور کند عادت یا عملی که خلاف عرف یا نامناسب است توسط افراد زیادی انجام می‌شود و انجام آن غیرمنتظره نیست مثلاً "بسیاری از بیماران گاهی اوقات دستور مصرف دارو را فراموش می‌کنند. آیا تا به حال شما دستور مصرف دارویتان را فراموش کرده‌اید؟" جمع‌آوری اطلاعات در مواردی نظیر عادت‌های بد و یا درآمد مشکل است. بعضی از مردم ترجیح می‌دهند به این سؤالات را در پرسشنامه جواب دهند نه در مصاحبه؛ اما یک مصاحبه‌کننده بسیار ماهر گاهی اوقات جواب‌های خوبی دریافت می‌کند. در مصاحبه‌های شخصی بهتر است پاسخ‌های موردنظر روی کارت‌هایی نوشته شود و پاسخ‌دهنده پاسخ‌ها را علامت بزند.

♦ زمان‌بندی

سؤالات اصلی پرسشنامه معمولاً برای اندازه‌گیری فراوانی رفتارهای خاص شبیه عادات غذایی یا مصرف دارو طراحی شده‌اند. برای اندازه‌گیری فراوانی رفتار پاسخ‌دهنده باید جواب را در مقیاس زمان نظیر تعداد بطری نوشابه نوشیده شده در طول یک روز ارائه دهد. اگر این رفتار روزانه تکرار شود مثلاً هرروز صبح یک قرص مصرف شود، سؤال خیلی ساده مطرح می‌شود: شما در یک روز چند قرص مصرف می‌کنید؟ بسیاری از رفتارهای روزانه، فصل‌به‌فصل یا سالانه تغییر می‌کنند. برای اندازه‌گیری این رفتارها ابتدا محقق

باید تصمیم بگیرد که جنبه‌های رفتاری مهم در مطالعه چه شاخص‌هایی هستند، میانگین یا حالت‌های فرین^{۳۴}. مثلاً در مطالعه اثر مزمن الکل در افزایش خطر سکتة قلبی احتمالاً نیاز به اندازه‌گیری میانگین مصرف الکل در یک زمان خاص است در صورتی که مطالعه‌ای که برای بررسی نقش الکل در وقوع صدمات ناشی از زمین خوردن طراحی می‌شود احتمالاً نیاز به اندازه‌گیری فراوانی دفعاتی دارد که نوشیدن بیش از اندازه الکل سبب بر هم خوردن تعادل افراد مورد مطالعه شده است.

سؤالاتی در مورد میانگین رفتار به دو روش پرسیده می‌شود: پرسیدن درباره رفتار معمول یا شمردن در طول دوره‌ای از زمان.

مثلاً محققان ممکن است میانگین مصرف نوشابه را با پرسش زیر اندازه‌گیری کنند:

- ♦ تقریباً در طول هفته چند شیشه نوشابه مصرف می‌کنید؟ (هر شیشه نوشابه معادل ۲۰۰ CC است)
- نوشابه در هفته

این ساختار ساده و خلاصه است. پاسخ‌دهنده با یک برآورد ساده میانگین رفتار خود را بیان می‌کند. چون نوشیدن در فاصله‌های زمانی کوتاه تغییر می‌کند ممکن است پاسخ‌دهنده برای اینکه تصمیم بگیرد در مورد کدام هفته صحبت کند مشکل داشته باشد. در مواجهه شدن با سؤالاتی که در مورد رفتارهای روزانه، افراد تحت مطالعه اغلب گزارشی از فعالیت‌های متداول خود را بیان می‌کنند و حالت‌های فرین را در نظر نمی‌گیرند. پرسش در مورد مصرف نوشابه در طول هفته معمولاً برآورد کمتری از مصرف نوشابه در پایان هفته ارائه می‌دهد. روش جایگزین دیگر مشخص کردن زمان دقیق مواجهه است.

- ♦ مثال: در طول ۷ روز گذشته چقدر نوشابه مصرف کرده‌اید؟

..... نوشابه در ۷ روز گذشته

تمرکز روی یک دوره زمانی ویژه و کوتاه (مانند هفته گذشته) توانایی به خاطر آوردن و تخمین دقیق‌تر را برای پاسخگو فراهم می‌کند. این امکان وجود دارد دوره زمانی تعریف شده در سؤال مشابه روزهای معمول سال نباشد. اگر فاصله زمانی شامل تعطیلات آخر هفته یا مناسبت‌های خاص باشد مثلاً ممکن است باعث شود میانگین مصرف نوشابه بیشتر برآورد شود. وقتی در مورد رفتار خاصی سؤال می‌شود برای پاسخ‌دهنده یادآوری عملکرد در فاصله زمانی طولانی‌تر مشکل است و پاسخ آن‌ها به سمت رفتار معمول‌تر گرایش دارد.

اگر رفتار تحت بررسی منظم و متداول نباشد استفاده از یک تقویم بسیار مفید است. در این روش محقق می‌تواند میانگین رتبه روزانه را محاسبه کند.

♦ سؤالات دوپهلو

هر سؤال فقط یک مفهوم را باید توضیح دهد. گاهی اوقات پاسخ سؤالاتی که شامل لغاتی از قبیل “و” و “یا” هستند نامفهوم به نظر می‌رسند. فرض کنید سؤالی برای ارزیابی مصرف کافئین طراحی شده باشد. “در طول یک روز شما چند فنجان قهوه یا چای می‌نوشید؟” کافئین قهوه بیشتر از کافئین چای است. به همین علت اندازه‌گیری

میزان کافین مصرفی دقیق نیست.

وقتی سؤالی برای ارزیابی دو مقوله در یک زمان مطرح می‌شود بهتر است در دو سؤال تبدیل شود. "در طول یک روز چند فنجان قهوه می‌نوشید؟" و "در طول یک روز چند فنجان چای می‌نوشید؟"

♦ مفروضات قراردادی

گاهی اوقات سؤالات شامل فرض‌هایی است که به همه اعضای مورد مطالعه مربوط نمی‌شود. مثلاً یک سؤال استاندارد در مورد "اضطراب" از پاسخ‌دهنده می‌خواهد دفاعاتی که دچار چنین احساسی شده است را بازگو کند. سؤالی نظیر "من احساس می‌کنم حتی با کمک خانواده‌ام هم نمی‌توانم نگرانی‌ام را برطرف کنم". برای کسانی که دارای خانواده نیستند و یا خانواده از آن‌ها حمایت نمی‌کند مشکل است.

♦ سؤالات و گزینه‌های نامرتب

سؤال و گزینه‌های پاسخ باید به هم مربوط باشند. این نکته مطلب مهمی است که ساده به نظر می‌رسد ولی اغلب به درستی رعایت نمی‌شود. مثلاً برای سؤال "آیا در هفته گذشته درد داشته‌اید؟" گاهی اوقات گزینه‌های جواب عبارت‌اند از "هرگز، به ندرت، خیلی اوقات، غالباً" که از نظر دستوری نادرست است و ممکن است پاسخ‌دهنده را گمراه کند. سؤال باید به صورت زیر تغییر کند: "در طول هفته گذشته چند دفعه احساس درد داشته‌اید؟" یا اینکه جواب‌های سؤال اول به صورت "بله خیر" باشد. مشکل دیگر وقتی رخ می‌دهد که گزینه‌های پاسخ توافق یا عدم توافق باشد. مثلاً اگر سؤالی به صورت زیر مطرح شود "من گاهی اوقات افسرده هستم" و گزینه‌های پاسخ "موافق هستم، موافق نیستم" باشد برای آن‌هایی که اغلب افسرده هستند پاسخ گمراه‌کننده است. موافق نبودن با این جمله یعنی اینکه شخص اغلب افسرده است یا هرگز افسرده نیست. در این حالت بهتر است از سؤالات واضح‌تر استفاده شود. مثلاً "چقدر شما احساس افسردگی می‌کنید؟" با گزینه‌های پاسخ "هرگز، گاهی اوقات، اغلب".

♦ امتیازبندی برای اندازه‌گیری متغیرهای کیفی

ارزیابی کمی مفاهیم ذهنی از قبیل کیفیت زندگی به وسیله یک سؤال مشکل است و معمولاً شاخص‌های خلاصه به وسیله امتیازبندی یک سری سؤال اندازه‌گیری می‌شود.

استفاده از چند گزینه برای ارزیابی یک مفهوم در مقایسه با یک سؤال تکی یا چند سؤال توأم، مزایایی دارد. در مقایسه با روش‌های دیگر مقیاس‌های چندگزینه‌ای دامنه پاسخ‌های ممکن را افزایش می‌دهد. (مثلاً دامنه امتیاز در یک مقیاس چندگزینه‌ای برای ارزیابی کیفیت زندگی از ۱ تا ۱۰۰ است. در صورتی که یک سؤال تکی ممکن است چهار یا پنج پاسخ از بسیار نامناسب تا کاملاً مناسب را شامل شود.)

از معایب مقیاس‌های چندگزینه‌ای این است که نتایج (مثلاً کیفیت زندگی = $46/2$) به سختی قابل فهم هستند.

معمولاً برای کمیت‌هایی از قبیل طرز برخورد، رفتارها و کیفیت زندگی از طیف لیکرت استفاده می‌کنند. این مقیاس‌ها فهرستی از سؤالات را در اختیار پاسخ‌دهنده قرار می‌دهد. سپس از پاسخ‌دهنده خواسته می‌شود تا

نزدیک‌ترین گزینه به عقیده خود را انتخاب کند. درنهایت به پاسخ‌ها امتیاز داده می‌شود.
 ♦ در هر سؤال دور عددی که از نظر شما بهترین پاسخ است، خط بکشید.

بسیار مخالف	مخالف	بی‌نظر	موافق	بسیار موافق	
۵	۴	۳	۲	۱	
					الف سیگار کشیدن در امکان عمومی باید غیرقانونی شود
					ب تبلیغات برای سیگار باید ممنوع شود
					ج درآمد ملی باید صرف مبارزه با سیگار شود

یک محقق به‌سادگی می‌تواند امتیاز کل پاسخ‌دهنده را با جمع کردن امتیاز هر سؤال یا میانگین به ازای سؤال‌های پاسخ‌داده‌شده، محاسبه کند. مثلاً امتیاز کل کسی که با سؤال اول و دوم بسیار موافق است (۱ امتیاز) اما با سؤال سوم مخالف است (۴ امتیاز)، ۶ است.

سازگاری درونی هر طیف با معیارهای آماری از قبیل آزمون آلفای کرونباخ قابل‌اندازه‌گیری است [۲]. این آزمون، سازگاری درونی مقیاس را ارزیابی می‌کند. آلفای کرونباخ با استفاده از همبستگی بین امتیازات هر سؤال محاسبه می‌شود. مقادیر بالای ۰/۷ برای این معیار معمولاً قابل‌قبول است و مقادیر ۰/۸ یا بیشتر بسیار عالی است. مقادیر کمتر نشان می‌دهد که هر گروه از سؤالات شاخص‌های مختلفی را اندازه‌گیری می‌کند (سازگاری درونی وجود ندارد).

ب-۱-۴ ساختن مقیاس‌ها و پرسشنامه‌های جدید

گاهی اوقات محقق شاخصی را اندازه‌گیری می‌کند که برای آن پرسشنامه و یا روش مصاحبه استاندارد وجود ندارد. در این حالت باید سؤالات یا مقیاس‌های جدید طراحی شود. محدوده این فعالیت می‌تواند از ساختن یک سؤال جدید راجع به یک متغیر فرعی در مطالعه (مثلاً روزی چند بار مسواک می‌زنید) تا توسعه و آزمون یک طیف چندگزینه‌ای برای اندازه‌گیری یک پیامد (رضایت‌مندی شغلی) مطالعه، گسترده باشد. ممکن است برای توسعه ابزار جدید برای اندازه‌گیری یک مفهوم مهم مانند اندازه‌گیری ضریب هوشی، روش‌های سیستماتیک و طولانی‌مدت برای تبدیل نسخه اولیه به نتیجه نهایی، نیاز باشد. زمانی که پرسشنامه یا مصاحبه آماده شد گام بعدی بررسی نقادانه به‌وسیله اشخاص متخصص و ماهر است.

ب-۱-۵ مراحل تدوین پرسشنامه در مطالعه

برای جمع‌آوری یک مجموعه پرسشنامه برای هر مطالعه چند مرحله وجود دارد. این مراحل شامل تهیه فهرستی از متغیرها که ممکن است برای مطالعه مناسب باشند، جمع‌آوری پرسشنامه‌های موجود و طراحی یک پرسشنامه جدید که واضح، جذاب و آسان باشد، انجام پیش‌آزمون، برای برآورد زمان موردنیاز برای اجرای پرسشنامه و آزمون تکرارپذیری و اعتبار پاسخ‌ها است.

♦ تهیه فهرستی از متغیرها

قبل از طراحی یک مصاحبه یا پرسشنامه محقق باید فهرستی از اطلاعات جمع‌آوری شده و مفاهیم اندازه‌گیری شده در مطالعه را تهیه کند (مثلاً پیامدها و مخدوش‌کننده‌ها).

♦ جمع‌آوری پرسشنامه‌های موجود

برای اندازه‌گیری هر متغیر، یک چک‌لیست را کامل کنید. یک روش مفید تهیه یک پوشه برای هر متغیر است. سپس سؤالات انتخابی برای موضوع تهیه می‌شود. باید بهترین پرسشنامه یا شیوه مصاحبه برای اندازه‌گیری پیامدها و مواجهه‌ها استفاده شود.

معمولاً منابع مختلف برای پرسشنامه‌ها و مصاحبه‌ها وجود دارد. ابتدا بهتر است پرسشنامه‌های محققان دیگر که معیارهای موردنظر را مطالعه کرده‌اند جمع‌آوری شود. بسیاری از پرسشنامه‌های استاندارد در کتاب‌ها، مقاله‌های مروری و منابع الکترونیکی از قبیل cdها و سایت‌های اینترنتی موجود است. با استفاده از عباراتی نظیر "پرسشنامه‌های مربوط بررسی الگوی تغذیه" می‌توان سایت موردنظر را پیدا کرد و پرسشنامه‌ها را جمع‌آوری نمود. همچنین پرسشنامه‌های منتشر شده سایر مطالعات با موضوعات مشابه، یا تماس گرفتن با محققان دیگر نیز می‌تواند برای گردآوری پرسشنامه مفید باشد.

استفاده از پرسشنامه‌های مطالعات دیگر در زمان صرفه‌جویی می‌کند و می‌توان یافته‌ها را با نتایج مطالعات دیگر مقایسه کرد. ممکن است پرسشنامه‌های موجود کاملاً برای سؤال یا جامعه مورد مطالعه مناسب نباشد و یا اینکه خیلی طولانی باشد؛ اما اگر بعضی موضوعات مناسب نبود می‌توان چند سؤال را حذف یا اضافه کرد و یا تغییر داد (همه این مشکلات هنگامی رخ می‌دهند که پرسشنامه که برای یک فرهنگ یا موضوع خاص تهیه و تنظیم شده در موقعیت‌های دیگر به کار گرفته شود).

♦ تهیه پیش‌نویس

پیش‌نویس اولیه در مقایسه با پرسشنامه اصلی باید شامل سؤالات بیشتری در مورد موضوع باشد ولی از نظر ساختار باید کاملاً شبیه پرسشنامه نهایی باشد.

♦ اصلاح کردن

محقق باید پیش‌نویس اولیه را به‌دقت مطالعه کرده و خود را یک پاسخ‌دهنده تصور کند. به همه سؤالات پاسخ دهد و همه تفسیرهای ممکن برای هر سؤال را بررسی کند. هدف تغییر لغات یا عباراتی است که ممکن است حتی تعداد کمی از پاسخ‌دهندگان را گمراه کند یا درست فهمیده نشود. پیدا کردن لغات کوتاه یا نامفهوم و ترجمه آن‌ها به عبارات ساده‌تر و قابل فهم‌تر نیز یکی از اهداف است. سؤالات پیچیده که شامل عباراتی از قبیل "و"، "یا" هستند قابل تبدیل به دو یا چند سؤال ساده‌تر هستند. بهتر است از افراد متخصص و ماهر برای مرور و طراحی پرسشنامه استفاده کرد. هدف هر سؤال باید خیلی واضح بیان شود.

♦ کوتاه کردن پرسشنامه مطالعه

معمولاً در مطالعات داده‌های بیشتری در مقایسه با آنچه باید تحلیل شود جمع‌آوری می‌شود. مصاحبه‌ها و پرسشنامه‌های طولانی ممکن است پاسخ‌دهنده را خسته کنند و باعث کاهش اعتبار پاسخ‌ها شوند. وقتی پرسشنامه به‌وسیله پست در اختیار مردم قرار می‌گیرد، توجه کمتری به پرسشنامه‌های طولانی می‌شود. سؤالاتی که پاسخ آن‌ها به هدف اصلی مطالعه مربوط نمی‌شود باعث صرف وقت اضافی برای ورود، تصحیح و تحلیل داده‌ها می‌شود و دقت مطالعه را کاهش می‌دهد. برای اینکه تصمیم بگیرید سؤالی در پرسشنامه وجود داشته باشد یا نه ابتدا در مورد تحلیل و ارائه گزارش نتایج آن در مطالعه فکر کنید. با خلاصه کردن جداول پایانی مطمئن می‌شوید که همه متغیرهای موردنیاز در مطالعه وجود دارد. اگر شک دارید موضوع یا معیاری در تحلیل‌های بعدی استفاده می‌شود یا نه بهتر است آن را حذف کنید.

♦ پیش‌آزمون

برای وضوح، تصحیح و برآورد زمان تکمیل پرسشنامه، باید پیش‌آزمون انجام شود. نقش پیش‌آزمون هنگام بررسی اعتبار و تکرارپذیری پرسشنامه‌ها بسیار بارزتر از مورد اخیر است.

♦ معتبر سازی

مانند هر معیار دیگری، باید اعتبار و تکرارپذیری پرسشنامه‌ها و مصاحبه‌ها بررسی شود. اولین مرحله بررسی اعتبار ظاهری سؤالات است. این مرحله سلیقه‌ای بوده و در آن موضوعاتی که شاخص‌های موردنظر محقق را در بردارند، بررسی می‌شوند. درستی پرسشنامه‌های جدید در مقایسه با معیارهای مرجع استاندارد، بررسی می‌شود. وقتی معیار استاندارد مرجع در دسترس نیست پاسخ‌های دو گروه از بیماران که در سطوح مختلفی از شرایط موردنظر قرار دارند، مقایسه می‌شوند. وقتی تفاوت زیادی بین دو گروه وجود داشته باشد نتیجه می‌گیریم سؤالات معتبرتر هستند. اعتبار پیش‌گویی یک پرسشنامه به‌وسیله ضریب همبستگی با وقایع آینده بررسی می‌شود. مثلاً در سؤال "آیا شما می‌توانید بدون عینک یا لنزهای تماسی روزنامه بخوانید؟" جواب‌ها با پاسخ‌های پرسشنامه معتبر موجود که برای بیماران مبتلا به آب‌مروارید و افرادی با چشمانی سالم تکمیل شده‌اند مقایسه می‌شود. با مقایسه پاسخ‌های بیماران مبتلا به آب‌مروارید قبل و بعد از عمل جراحی اعتبار پرسشنامه آزمون می‌شود.

♦ توالی سؤالات

در اکثر مواقع بهترین روش برای آغاز پرسش‌ها، سؤالات روزمره و عادی است. اگر سؤالات آغازین بسیار شخصی باشند امکان دارد به شرکت‌کنندگان احساس بیگانه بودن با پرسشنامه القا شود. این موضوع که تمام اطلاعات مربوط به افراد مطالعه حتی اطلاعاتی نظیر سن و نژاد و غیره محرمانه باقی می‌ماند حتماً باید به اطلاع تمام شرکت‌کنندگان مطالعه رسانده شود. ایجاد انگیزه راجع به موضوع پژوهش و تهیه سؤالات جالب به‌طوری که افراد تحت مطالعه باحوصله به تمام سؤالات پاسخ دهند بسیار مهم است. احتمالاً بهترین روش در طراحی توالی سؤالات قرار دادن سؤالات مهم و اساسی در اواسط پرسشنامه است.

این موضوع که پاسخ‌هایی که افراد تحت مطالعه برای سؤالاتی باز می‌نویسند، سؤالات بعدی را تحت تأثیر قرار ندهند باید موردتوجه قرار گیرد.

ب-۱-۶ میزان پاسخگویی

میزان پاسخگویی^{۲۵} به پرسشنامه‌ها بسته به اقشار جامعه که مورد بررسی قرار گرفته‌اند متغیر است؛ اما عموماً میزان میزان پاسخگویی را پایین گزارش کرده‌اند. بر اساس تحقیقات به عمل آمده میزان پاسخگویی حدود ۴۰٪ است و اگر نمونه‌گیری به صورت تصادفی ساده از کل جامعه باشد حتی پایین‌تر از این مقدار خواهد بود [۳].

رعایت موارد زیر ممکن است میزان پاسخگویی را بالاتر ببرد:

- ♦ استفاده از کاغذ مناسب و باکیفیت بخصوص اگر پرسشنامه پست می‌شود.
- ♦ استفاده از خط خوانا و واضح.
- ♦ درشت بودن خط بخصوص هنگامی که افراد تحت مطالعه دچار مشکلات بینایی هستند.
- ♦ فضای خالی برای پاسخ به سؤال‌ها در نظر گرفته شود.
- ♦ دستورالعمل پاسخگویی به سؤال‌ها با یک فونت متفاوت و مشخص برای هر قسمت تهیه شود.
- ♦ از تنظیم صحیح جداول مربوط به علامت زدن جواب‌ها اطمینان حاصل کنید.
- ♦ از نسخه تایپ‌شده پرسشنامه استفاده کنید.
- ♦ صفحات پرسشنامه را شماره‌گذاری کنید.
- ♦ در پرسشنامه‌های پستی، پاکت نامه تمبرخورده به همراه آدرس برگشت بروی آن در اختیار افراد تحت مطالعه قرار داده شود.
- ♦ یک نسخه چکیده پژوهش را در اختیار افراد تحت مطالعه قرار داده شود.
- ♦ تاریخ برگشت پاسخ‌ها را مشخص کنید.
- ♦ در یک نامه ضمیمه، اهداف پژوهش حاضر را شرح دهید و همچنین محققان را معرفی نمایید.
- ♦ از این نکته که افراد تحت مطالعه، پرسشنامه را در محیط و زمان مناسب کامل می‌کنند مطمئن شوید.
- ♦ پرسشنامه‌ها را وقتی اعضای نمونه بسیار گرفتار و مشغول هستند در اختیار آن‌ها قرار ندهید. مثلاً تعطیلات آخر سال، امتحان‌های پایان ترم و غیره زمان‌های مناسبی برای تکمیل پرسشنامه نیست.

ب-۱-۷: مزایا و معایب پرسشنامه

مزایا

- ♦ باعث صرفه‌جویی در هزینه و زمان می‌شود.
- ♦ دسترسی به گروه زیادی از افراد جامعه به طریق نسبتاً آسان و ارزان امکان‌پذیر می‌شود.
- ♦ تماس چهره به چهره محققان و افراد تحت مطالعه، که باعث تأثیرات اجتماعی و روانی بر روی افراد تحت مطالعه می‌شود، حذف می‌شود.
- ♦ پژوهش‌گران زمان زیادی برای طراحی سؤال‌ها در اختیار دارند و می‌توانند سؤال‌های سنجیده‌تری نسبت به مصاحبه طرح کنند.

معایب

- ♦ پاسخ‌های داده شده می‌توانند ظاهری و غیرواقعی باشند و افراد تحت مطالعه برخلاف واقعیت اطلاعاتی به محققان ارائه کنند.
- ♦ فقدان تماس چهره به چهره محققان و افراد تحت مطالعه ممکن است باعث گمراهی و بدفهمی در پاسخ به سؤال‌ها شود.
- ♦ پرسشنامه‌ها برای اعضای خاصی از جامعه نظیر کودکان یا افراد بی‌سواد مناسب نیستند.
- ♦ میزان پاسخ‌دهی معمولاً پایین است و درصد پرسشنامه‌های ناکامل یا بی‌دقت کامل شده بسیار بالا است.

ب-۲ مصاحبه

مصاحبه یک روش تحقیق نیست بلکه مجموعه‌ای از روش‌های تحقیق است که تنها در یک موضوع مشترک هستند. مصاحبه شامل گفتگو بین دو نفر است که در آن یک فرد نقش محقق را بازی می‌کند [۴]. مصاحبه می‌تواند بسیار حساب‌شده و دارای ساختار در حد یک پرسشنامه شفاهی و یا به‌طور کامل بدون ساختار و در حد یک گفتگوی متداول باشد. اکثر مصاحبه‌ها جایی بین این دو حد انتهایی قرار می‌گیرند [۵].

مصاحبه نامنظم	مصاحبه نیمه منظم	مصاحبه منظم
Unstructured interview	Semi-Structured interview	Structured interview

ب-۲-۱ مصاحبه منظم

در این نوع مصاحبه، مصاحبه‌گر درباره سؤال‌هایی که باید پرسیده شوند، ترتیب و توالی آن‌ها و کلمات دقیقی که در هر پرسش باید گنجانده شود، الگوی دقیقی در دست دارد و از روی دستورالعمل صادر شده عمل می‌کند. در مصاحبه منظم به‌طور معمول مصاحبه‌گر سؤال‌ها را برای افراد مطالعه از روی فرم‌های مخصوصی که دارای دستورالعمل و کدهای ویژه متناسب با هر جواب است خوانده و جواب‌ها را در فرم علامت می‌گذارد. تنها تفاوت مصاحبه منظم و پرسشنامه در این است که افراد مطالعه در مصاحبه منظم فرم سؤال‌ها را در اختیار ندارند و به سؤال‌ها به‌طور شفاهی پاسخ می‌دهند. در این صورت عواملی نظیر جذابیت سؤال‌ها، واضح بودن دستورالعمل پر کردن فرم‌ها و عدم توانایی در خواندن دقیق آن‌ها چندان مؤثر نیست. نکته مهم دقت پرسشگران در نحوه بیان سؤال‌ها است که باید به طریقی باشد که افراد تحت مطالعه ناخودآگاه تحت تأثیر قرار نگیرند.

ب-۲-۲ مصاحبه نیمه منظم

در این نوع مصاحبه، مصاحبه‌گر از موضوع بررسی آگاهی کامل دارد؛ اما در نحوه جمله‌بندی سؤال‌ها، توالی و ترتیب آن‌ها آزادی عمل دارد. امکان دارد از فرم‌های مخصوص نیز استفاده شود؛ اما پرسشگر مجاز است فقط سؤال‌های خاصی از فرم را بپرسد. پرسشگران در این نوع مصاحبه افراد تحت مطالعه را ترغیب به بیان پاسخ‌های کامل‌تر و

دقیق تر می‌کنند. برای ملاحظه مثال‌هایی از این نوع خوانندگان می‌توانند به منابع [۶] تا [۹] مراجعه کنند.

ب-۲-۳ مصاحبه نامنظم

در این نوع مصاحبه، هیچ مجموعه از پیش تعیین‌شده‌ای برای سؤال‌ها وجود ندارد. افراد تحت مطالعه مجاز هستند احساسات و عقاید خود را بدون هیچ‌گونه محدودیت، کنترل یا جهت‌دهی از سوی پرسشگر بیان نمایند. در مصاحبه نامنظم پرسشگر بعضی نکات مبهم را از فرد مطالعه سؤال می‌کند و سعی می‌کند درباره موضوعات جالب‌تر اطلاعات بیشتری به دست آورد. استفاده عمده مصاحبه نامنظم در مطالعات روانشناسی، جلسات مشاوره و بیماری‌های خاص نظیر ایدز، بیماری‌های روانی خاص و غیره است؛ اما در مطالعات دیگر به‌ویژه در مراحل اولیه پژوهش از این نوع مصاحبه برای جمع‌آوری اطلاعات اولیه استفاده می‌شود. منابع [۱۰]، [۱۱] و [۱۲] مثال‌های موفقی از این موضوع هستند.

ب-۲-۴ مزایا و معایب مصاحبه منظم

- ♦ مصاحبه منظم روشی دقیق، واضح و هدفمند در جمع‌آوری اطلاعات است.
- ♦ مصاحبه منظم نسبتاً کم‌هزینه و کم‌خرج است.
- ♦ استخراج و کدبندی داده‌ها در این روش آسان است.
- ♦ این روش معمولاً دارای روایی و پایایی بالا است.
- ♦ ترتیب پرسیدن سؤال‌ها، نحوه بیان دقیق آن‌ها و الگوی از پیش تعیین‌شده سؤال‌ها ممکن است این ذهنیت را در افراد تحت مطالعه ایجاد کند که خود آن‌ها به‌طور مستقل مهم نیستند بلکه به‌عنوان عضوی از این مطالعه دارای اهمیت هستند. این موضوع می‌تواند سبب کاهش همکاری افراد تحت مطالعه شود [۱۳].

ب-۲-۵ مزایا و معایب مصاحبه نیمه منظم و نامنظم

- ♦ مصاحبه نیمه منظم و نامنظم در مواردی که موضوع مطالعه نسبتاً مبهم است یا هنگامی که مایل به جمع‌آوری اطلاعات عمیق از افراد تحت مطالعه هستیم روشی مفید است.
- ♦ در مصاحبه نیمه منظم پرسشگر می‌تواند نحوه بیان سؤال‌ها را طوری تنظیم کند که برای افراد مختلف مطالعه مناسب باشد. مثلاً سؤال‌ها را برای افراد جوان و مسن با کلمه‌های مختلف بیان کند یا برای افرادی که آشنایی زیادی با زبان پرسشگر ندارند ساده‌تر بیان کند. هم‌چنین بسته به شرایط مصاحبه برخی سؤال‌ها را زودتر و برخی را دیرتر مطرح کند.
- ♦ استاندارد نبودن نحوه بیان سؤال‌ها و یکسان نبودن ترتیب آن‌ها در مصاحبه نیمه منظم و نامنظم می‌تواند به جمع‌آوری داده‌هایی که شدیداً بی‌اعتبار هستند منجر شود.
- ♦ استخراج و کدبندی پاسخ‌ها در این نوع مصاحبه‌ها بسیار مشکل است.
- ♦ این نوع مصاحبه‌ها گران و وقت‌گیر هستند.
- ♦ آموزش افراد برای انجام این نوع مصاحبه‌ها بسیار مشکل است.

- ♦ تورش ناشی از تأثیر و جهت‌دهی پرسشگر در این مصاحبه‌ها می‌تواند بسیار زیاد و اساسی باشد تا حدی که باعث تحریف تمام واقعیت‌ها شود [۴] و [۵].
- مهارت مصاحبه‌کننده روی کیفیت پاسخ‌ها تأثیر می‌گذارد. استاندارد کردن فرآیند مصاحبه از یک مصاحبه به مصاحبه بعدی، اساس بالا بردن تکرارپذیری مصاحبه‌ها است. مصاحبه باید با جمله‌بندی یکنواخت، سؤالات و نشانه‌های غیرکلامی یکسان تنظیم شود. مصاحبه‌کننده باید به‌دقت از تورش مربوط به مصاحبه‌کننده اجتناب کند. گاهی اوقات لازم است پرسش‌های اضافه برای رسیدن به پاسخ مناسب مطرح شود. این موضوع به علت ترغیب افراد تحت مطالعه به جواب دادن واضح و کامل به سؤالات انجام می‌شود. این روش با نوشتن عبارات استاندارد در حاشیه متن هر سؤال استاندارد می‌شود. مثلاً هنگامی که در جواب سؤال ”هرروز شما چند فنجان قهوه می‌نوشید؟“ پاسخ‌دهنده می‌گوید ”مطمئن نیستم“ و یا ”هرروز متفاوت است“ دستورالعمل باید شامل یکسری سؤال‌های اضافی نظیر ”به بهترین نحو ممکن تخمین بزنید“ و یا ”تقریباً چند فنجان قهوه در یک روز عادی می‌نوشید؟“ باشد.
- پدیده‌ها و متغیرهای مورد مطالعه با کمک ابزارهای پژوهش موردسنجش و اندازه‌گیری قرار می‌گیرند و به ”داده“ تبدیل می‌شود. قبل از گردآوری داده‌ها باید به سؤالاتی پاسخ داد. نمونه‌ای از این سؤالات در جدول زیر آورده شده است. پاسخ این پرسش‌ها مهم است، ولی ثابت نیست و بیشتر پاسخ‌ها همراه با پیشرفت پروژه تحقیق تغییر می‌کند و باعث بازنگری در بسیاری از عناصر می‌شود.

مرحله	سؤالاتی که هنگام گردآوری اطلاعات، آنالیز، به اشتراک گذاشتن، نشر و ضبط داده‌ها پرسیده می‌شود.
گردآوری	چه داده‌هایی جمع‌آوری خواهد شد؟ داده‌ها چگونه ثبت می‌شود؟ چه کسی داده‌ها را جمع‌آوری می‌کند؟ معیار رد کردن داده‌ها چیست؟ معیار پذیرفتن داده‌ها چیست؟
آنالیز	چه روشی برای آنالیز داده‌ها استفاده می‌شود؟ کدام نتایج باید منتشر شود؟ چه زمانی داده‌ها منتشر می‌شود؟ داده‌ها در کجا منتشر می‌شوند؟
ذخیره سازی	کدام‌یک از سوابق تحقیق، ضبط و ذخیره می‌شود؟ سوابق را چگونه ذخیره می‌کنند؟ سوابق را چه مدت ذخیره می‌کنند؟ مسئول ذخیره سازی داده‌ها کیست؟
به اشتراک گذاشتن	کدام‌یک از سوابق تحقیق ممکن است به اشتراک گذاشته شود؟

چه زمانی می‌توان سوابق را به اشتراک گذاشت؟

چه کسی مسئولیت تصمیم‌گیری در مورد سوابق به اشتراک گذاشته شده را دارد؟

پایانی ابزار پژوهش

پایانی ابزار پژوهش تابع تکرار پذیری است؛ یعنی اگر متغیر موردنظر را با همان ابزار (تحت شرایط مشابه) دو بار اندازه بگیریم، نتایج مشابه باشد. ثبات و همبستگی درونی و توافق بین مشاهده‌گران در جمع‌آوری اطلاعات نیز نشان دهنده پایایی است. به‌طور کلی عوامل مؤثر بر پایایی ابزار پژوهش تعداد سؤالات، دامنه تغییرات نمرات در جمعیت مورد مطالعه، پیچیدگی و مشکل بودن سؤالات هستند.

انواع روش‌های تعیین پایایی شامل آزمون مجدد^{۲۶}، تعیین ثبات درونی^{۲۷}، ثبات اندازه‌گیری یک آزمون در صورت کاربرد آن توسط افراد متفاوت^{۲۸} و فرم‌های معادل^{۲۹} است.

در روش آزمون مجدد، یک گروه از افراد انتخاب و آزمون به افراد ارائه می‌شود، با در نظر گرفتن فاصله زمانی مناسب، همان آزمون به افراد (با آزمون کننده یکسان) داده می‌شود و سپس همبستگی بین نمرات به دست آمده محاسبه می‌شود. این روش برای متغیرهای نسبتاً^{۳۰} پایدار مناسب خواهد بود. فاصله زمانی دو آزمون باید نه آنقدر کوتاه باشد که پاسخ‌های بار اول در خاطر افراد مانده باشد و نه آنقدر طولانی که متغیر اندازه‌گیری شده طی زمان تغییر کرده باشد. تکنیک‌های آماری مورد استفاده آزمون مجدد پیرسون^{۳۰} یا اسپیرمن^{۳۱} هستند. نکته مهم این است که در اثر گذشت زمان مقدار همبستگی کاهش می‌یابد زیرا وارینانس خطا افزایش پیدا می‌کند و هرچه فاصله دو بار اندازه‌گیری کوتاه‌تر باشد عوامل مؤثر بر وارینانس خطا مشابه‌تر خواهند بود. در این آزمون ضریب پایایی از ۰ تا ۱ می‌باشد و این مقدار هرچه به ۱ نزدیک‌تر باشد نشان دهنده همبستگی بیشتر است. همبستگی بالا تر از ۰/۷۰ برای بیشتر آزمون‌ها و همبستگی ۰/۹۰ برای آزمون‌های حیطة شناختی قابل قبول محسوب می‌شود.

در روش تعیین ثبات درونی، همبستگی بین سؤالات پرسشنامه سنجیده می‌شود و برای اطمینان از این‌که تمام سؤالات (یا مشاهدات) یک مفهوم را می‌سنجند کاربرد دارد. در این روش سوال‌ها فقط یکبار به آزمودنی‌ها ارائه می‌شود. روش‌های آماری متداول آلفا کرونباخ^{۳۲} و کودر ریچاردسون^{۳۳} هستند.

در تعیین ثبات اندازه‌گیری یک آزمون در صورت کاربرد آن توسط افراد متفاوت، دو نفر مشاهده کننده مشاهدات مستقل خود را گزارش و سپس ضریب همبستگی بین نمرات محاسبه خواهد شد. روش آماری مورد

۱. Test-retest reliability

۲. Internal consistency reliability

۳. Inter rater or inter observer

۴. Parallel forms

۵. Pearson r

۶. Spearman

۷. Cronbach's alpha: for non-dichotomous data

۸. Kuder-Richardson-20 (KR-20): for dichotomous data

استفاده برای داده‌های اسمی آزمون کاپا^{۳۴}، برای داده‌های رتبه‌ای آزمون کندال^{۳۵} و برای داده‌های دامنه‌ای یا یا نسبتی ضریب پیرسون^{۳۶} می‌باشد.

در تعیین پایایی ابزار پژوهش از طریق فرم‌های معادل ثبات نتایج در صورت کاربرد دو پرسشنامه مشابه که همان مفهوم را بسنجند تعیین می‌شود. پرسشنامه تنها یک بار به آزمودنی‌ها ارائه می‌شود و در مواردی که استفاده از چند فرم یا پرسشنامه موردنظر باشد، مناسب یا حتی گاه ضروری است. روش آماری مورد استفاده برای این روش ضریب پیرسون یا اسپیرمن خواهد بود.

روایی ابزار پژوهش

روایی بخش لازم و مهمی از یک ابزار پژوهش است، ولی شاخصی کافی برای کامل بودن آن ابزار نیست. متأسفانه در بسیاری از مطالعات منتشر شده روایی محتوای^{۳۷} ابزارهای تحقیق مورد استفاده سنجیده نشده یا گزارش نمی‌شوند. تعیین روایی محتوا برای سنجش واقعی اطلاعات مورد نیاز لازم است و می‌تواند کیفیت مطالعه را ارتقا بخشد. دو روش استاندارد برای کسب اطمینان از روایی محتوا، نمونه‌گیری از آیتم‌ها^{۳۸} و روش ساختار آزمون^{۳۹} هستند. قضاوت مردم، یا قضاوت (۱۰ تا ۱۵ نفر) متخصصان و خبرگان موضوع از مفاد و محتوای پرسشنامه از راه‌های معمول برای تعیین روایی محتوا است. تعیین مرتبط بودن^{۴۰}، وضوح سؤالات^{۴۱} و سادگی^{۴۲} آن‌ها همراه با رتبه بندی (برای مثال غیر مرتبط=۱، مرتبط ولی نیازمند اصلاح اساسی=۲، مرتبط و نیازمند اصلاح جزئی=۳ و مرتبط=۴ امتیاز) و تعیین مجموع امتیاز به دست آمده و ارزیابی هر سؤال از روش‌های مرسوم است. [۱۴]

۹. Kappa Statistics

۱۰. Kendall's test (see if pairs of ranks for each of several individuals are related)

۱. Pearson r (using data obtained from the hyperactivity index)

۲. Content validity

۳. The sampling of items

۴. The method of constructing the test

۵. Relevance

۶. Clarity

۷. Simplicity

منابع

1. Judd C, Smith ER, Kidder LH. Research methods in social relations (6th ed.). Rinehart and Winston, London 1991.
2. Bland JM, Altman DG. Cronbach's alpha. British Medical Journal; 314: 542,1997.
3. May T. Social research: Issues, method and process (2nd ed.) Open University Press, Buckingham 1997.
4. Kvale S. Interviews: An introduction to qualitative research interviewing. Sage, London 1996.
5. Rubin HJ, Rubin IS. Qualitative interviewing: The art of hearing data. Sage. Thousand Oaks 1995.
6. Teram E, Schachter CL, Stalker CA. Opening the doors to disclosure: Childhood sexual abuse survivors reflection telling physiotherapists about their trauma. Physiotherapy 85(2): 88-97 1999.
7. Murray R. Measurement of the effect of participation in a medical humanities group on the practice of physiotherapists. Physiotherapy, 84(10): 473-9 1998.
8. Lister R. Loss of ability to drive following a stroke: The early experiences of three elderly people on discharge from hospital. B J Occup Therap; 62(11): 514-20,1999.
9. Melton J. How do clients with learning disabilities evaluate their experience of cooking with the occupational therapists? British Journal of Occupational Therapy; 61(3): 106-10, 1998.
10. Parr S, Byng S. Talking about Aphasia: Living with loss of language after stroke. Open University Press, Buckingham 1996.
11. Atkinson D, Jackson M, Walmsley J. Forgotten lives: Exploring the history of learning disability. British Institute of Learning Disabilities. Kidderminster 1997.
12. Soori H & Bhopal R.S. Parental permission for children's independent outdoor activities: Implications for injury prevention. European Journal of Public Health. 2002; 12: 104-109.
13. Arksey A, Knight P. Interviewing for Social Scientists. Sage, London 1999.

۱۴. سوری ح و همکاران. نرم افزارهای مورد استفاده در کارگاههای آزمونهای تشخیصی و اعتبارسنجی. دانشگاه علوم پزشکی شهید بهشتی. ۱۳۸۶

(برای دسترسی مراجعه شود به سایت مدیریت پژوهشی دانشگاه www.sbmu.ac.ir)

ضمیمه

ج

محاسبه حداقل حجم نمونه مورد نیاز

یکی از قسمت‌های مهم در اکثر تحقیقات کاربردی تصمیم‌گیری در مورد اندازه حجم نمونه است. حتی با دقیق‌ترین و ماهرانه‌ترین طرح‌های مطالعاتی هنگامی که حجم نمونه ناکافی باشد، قادر به بررسی علمی سؤال پژوهش نخواهیم بود. در مقابل، هنگامی که مطالعه‌ای با حجم نمونه بیش از اندازه مورد نیاز طراحی می‌شود اجرا کردن آن مشکل و پرهزینه می‌شود. هدف از تعیین حجم نمونه برآورد تعداد اعضای تحت مطالعه در مطالعه مورد نظر است. محاسبه حجم نمونه، روشی ریاضی برای یافتن برآوردی تقریباً صحیح از اعضای مورد مطالعه است و باید در ابتدای طراحی مطالعه که امکان اعمال تغییرات کلی وجود دارد برآورد شود. در این فصل ابتدا اصول پایه برای محاسبه حجم نمونه معرفی شده و سپس تکنیک‌های ساده‌ای برای برآورد حجم نمونه در مطالعات ارائه شده است. جداول و روابط متنوعی برای محاسبه حجم نمونه در انواع مختلف مطالعات ارائه شده و راهنمایی‌هایی برای جلوگیری از برخی خطاهای متداول در محاسبه حجم نمونه ذکر شده است. در ادامه کتاب‌ها و مقاله‌های سودمندی برای طرح‌های پیچیده‌تر معرفی شده است.

ج-۱- مقدمه

در مرحله طراحی هر مطالعه بالینی سؤال‌های زیر باید مورد توجه واقع شوند.

۱) حجم نمونه مورد نیاز برای رسیدن به توان آماری مورد نظر چقدر است؟ (برای مثال دستیابی به شانس ۸۰٪ در مشاهده تفاوت معنی‌دار بالینی)

۲) هنگام مواجهه با محدودیت‌هایی نظیر کمبود بودجه، امکانات و بیماران واجد شرایط چگونه می‌توان به حجم نمونه قابل قبولی دست یافت؟ بررسی این سؤال‌ها معمولاً با محاسبه آماری حجم نمونه ارزیابی می‌شود. تعیین حجم نمونه بر اساس روابط آماری با توجه به معیارهایی نظیر دقت و توان مطالعه صورت می‌گیرد. در صورتی که ارزیابی حجم نمونه با معیارهای آماری، به روی حجم نمونه از پیش تعیین شده که امکان دارد به علت محدودیت‌ها کوچک باشد انجام می‌شود. تأکید اصلی به روی محاسبه حجم نمونه استوار است؛ اما کاربرد این روش‌ها در قضاوت به روی حجم نمونه از پیش معین به آسانی امکان‌پذیر است. در هر مطالعه، حجم نمونه بر اساس مشخصه‌هایی نظیر خطاهای نوع اول و دوم محاسبه می‌شود. برای

مثال امکان دارد حجم نمونه را به طریقی محاسبه کرد که دقت مطلوب برای دستیابی به سطح اطمینان موردنظر (و یا خطای نوع اول ثابت) حاصل شود. به این روش تعیین حجم نمونه با دقت مشخص می‌گویند. یکی از معایب این روش کم بودن شانس مشاهده اختلاف‌های واقعی است (درواقع در این روش هیچ کنترلی بر توان آماری آزمون‌ها وجود ندارد). روش جایگزین محاسبه حجم نمونه برای توان آماری (خطای نوع دوم-۱) قابل قبول و دقت (خطای نوع اول) ثابت است. به عبارت دیگر در این روش، حجم نمونه محاسبه شده از توان آماری مناسبی برای تشخیص اختلاف‌های مهم بالینی در سطح ثابت خطای نوع اول برخوردار است. در اکثر مطالعات بالینی انجام مطالعه راهنما برای تعیین حجم نمونه متداول است. تعیین حجم نمونه بر اساس مطالعه راهنما با بررسی تابع توان که با استفاده از اطلاعات بدست آمده در این مرحله محاسبه می‌شود، آغاز می‌شود. در عمل وجود تناقض در اهداف (فرضیه‌ها) مطالعه، طرح مطالعه، تحلیل‌های آماری و محاسبه حجم نمونه غیرمنطقی است و می‌تواند باعث مشاهده نتایج غیرواقعی شود. تناقض‌های مهم در این زمینه می‌توانند ناشی از:

- ۱) آزمون نادرست برای فرضیه درست
 - ۲) آزمون درست برای فرضیه نادرست
 - ۳) آزمون نادرست برای فرضیه نادرست
 - ۴) آزمون درست برای فرضیه درست ولی با توان نامناسب
- باشد. به همین علت توصیه می‌شود قبل از محاسبه حجم نمونه مراحل زیر به دقت بررسی شود:
- ۱) اهداف و فرضیه‌های مطالعه به دقت مشخص شوند.
 - ۲) طراحی مطالعه و آزمون‌های آماری به کاررفته مناسب باشند.
 - ۳) حجم نمونه بر اساس فرضیه‌ها و آزمون‌های موردنظر محاسبه شود.

ج-۲- اساس محاسبه حجم نمونه

برای محاسبه حجم نمونه موردنیاز، لازم است صفات و پارامترهای مهم از لحاظ اهداف مطالعه را در افرادی که واجد شرایط برای ورود به مطالعه هستند ارزیابی کنیم و مراحل زیر باید انجام شود:

- ۱- تعیین اندازه اثر خطر که مایل به آشکار کردن آن‌ها در مطالعه هستیم. برای مثال ممکن است در یک مطالعه محقق مایل باشد اگر نسبت شانس ابتلا به بیماری در افراد دارای مواجهه ۴ برابر افراد بدون مواجهه باشد، آزمون‌های آماری قادر به نشان دادن این اختلاف باشد.
- ۲- سطح معنی‌داری، قدرت شواهد برای رد کردن فرضیه صفر است. مقدار احتمال کوچک‌تر به معنی قدرت بیشتر مدارک و شواهد برای رد فرضیه صفر و در نتیجه بزرگ‌تر شدن حجم نمونه می‌شود.
- ۳- احتمال دست یافتن به سطح معنی‌داری موردنظر (توان آزمون).

ج-۳- انواع خطاها در آزمون‌های آماری

از یک آزمون آماری هرگز نمی‌توان نتیجه گرفت که فرضیه صفر درست است یا نادرست. بلکه فقط نشانه‌ای از قدرت شواهد مخالف فرضیه صفر آشکار می‌شود. هنگام استفاده از یک آزمون آماری برای تصمیم‌گیری جهت رد یا قبول فرضیه صفر دو نوع خطا آشکار می‌شود. رد فرضیه صفر وقتی واقعاً درست است و رد نکردن فرضیه صفر وقتی واقعاً نادرست است. اولی خطای نوع اول و دومی خطای نوع دوم نامیده می‌شود. بنابراین در یک آزمون فرضیه دو طرفه هنگامی که مقدار خطای نوع اول را برابر $0/05$ در نظر می‌گیریم اگر نتیجه معنی‌دار شود (فرضیه صفر رد شود)، 5% احتمال دارد که خطا کرده باشیم (شکل ج-۱-الف).

خطای نوع دوم عبارت است از رد نکردن فرضیه صفر هنگامی که واقعاً نادرست است. این موضوع در شکل ج-۱-ب نشان داده شده است. ناحیه خاکستری، نسبتی $(b\%)$ از توزیع نمونه‌ای واقعی است که داخل ناحیه قبول فرضیه صفر قرار می‌گیرد. احتمال این که مرتکب خطای نوع دوم $(b\% - 100)$ نشویم برابر با توان آزمون است. جدول ج-۱-۱ حالت‌های وقوع این دو خطا را نمایش می‌دهد.

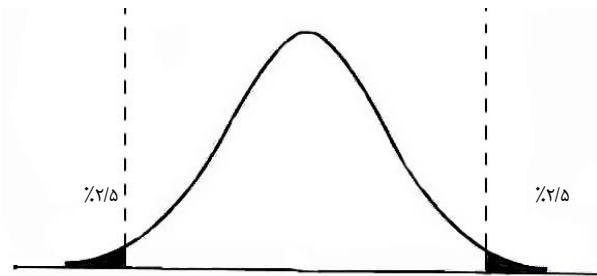
جدول ج-۱-۱: انواع خطا در آزمون فرضیه

واقعیت		
نتیجه آزمون معنی‌داری: فرضیه صفر رد می‌شود.	فرضیه صفر درست است. خطای نوع اول احتمال = سطح معنی‌داری	فرضیه صفر درست نیست. نتیجه درست احتمال = توان
فرضیه صفر رد نمی‌شود.	نتیجه درست احتمال = سطح معنی‌داری ۱-	خطای نوع دوم احتمال = توان ۱-

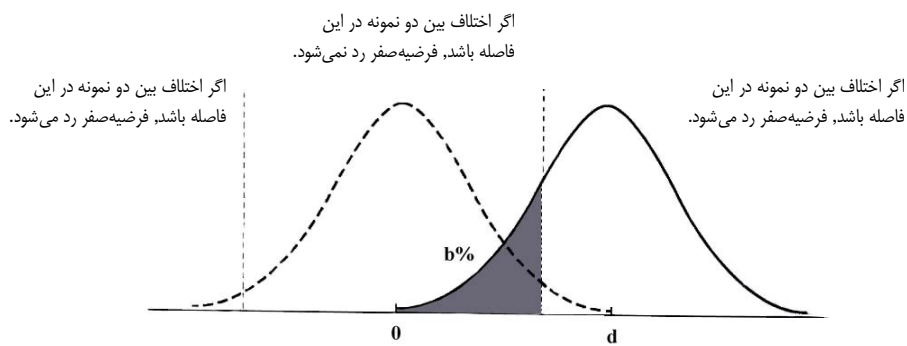
ج-۴- روابطی برای محاسبه حجم نمونه

فرمول‌های مناسب برای محاسبه حجم نمونه در حالت‌های کلی در جدول ج-۲ آمده است. این فهرست شامل دو قسمت است. جدول ج-۲-الف برای مواردی است که هدف مطالعه نشان دادن یک تفاوت معین است (مطالعات تحلیلی) و جدول ج-۲-ب برای مواردی که هدف برآورد کمیت مورد علاقه با دقت معین است (مطالعات توصیفی، تعیین شیوع یا بروز)، به کار می‌رود.

توجه به این نکته ضروری است که در حالت‌هایی که مقایسه دو میانگین، نسبت یا میزان موردنظر است، حجم نمونه موردنیاز، برای هر گروه محاسبه می‌شود و با دو برابر کردن این مقدار، حجم نمونه کلی بدست می‌آید.



الف) خطای نوع اول: فرضیه صفر درست است و تفاوت بین دو جامعه صفر است. در این شکل منحنی توزیع نمونه‌های تفاوت نمونه‌ها نشان داده شده است. نواحی خاکستری (۵٪ مساحت زیر منحنی) احتمال این که فرضیه صفر به اشتباه رد شود را نشان می‌دهد.



ب) خطای نوع دوم: فرضیه صفر نادرست است و تفاوت بین دو جامعه (d) صفر نیست. در این شکل منحنی با خط پیوسته توزیع نمونه‌های واقعی تفاوت نمونه‌ها را نشان می‌دهد و منحنی خط چین توزیع نمونه‌ای تحت فرضیه صفر را نشان می‌دهد. ناحیه خاکستری، احتمالی (b) است که فرضیه صفر، نادرست رد می‌شود.

شکل ج-۱: احتمال رخداد دو نوع خطا (خطای نوع اول و دوم) در آزمون فرضیه با سطح ۵٪

جدول ج-۲: محاسبه حجم نمونه الف) مطالعاتی که هدف پیدا کردن یک تفاوت معنی دار است. ب) مطالعاتی که هدف برآورد کمیت موردنظر با دقت است.

نحوه محاسبه کمترین حجم نمونه	اطلاعات مورد نیاز	الف) نتیجه معنی دار
$\frac{(v+v)^2 \sigma^2}{(\mu - \mu_0)^2}$	تفاوت بین میانگین μ و مقدار تحت فرضیه μ_0	۱- تعیین میانگین
$\frac{(v+v)^2 \mu}{(\mu - \mu_0)^2}$	نرخ μ	۲- تعیین نرخ *
$\frac{\left\{ v\sqrt{[\pi(1-\pi)]} + v\sqrt{[\pi_0(1-\pi_0)]} \right\}^2}{(\pi - \pi_0)^2}$	مقدار تحت فرضیه صفر	μ_0
	نسبت	π
	مقدار نسبت تحت فرضیه صفر	π_0
$\frac{(v+v)^2 (\sigma_1^2 + \sigma_0^2)}{(\mu_1 - \mu_0)^2}$	تفاوت بین میانگینها	۴- مقایسه دو میانگین (حجم نمونه برای هر گروه)
$\frac{(v+v)^2 (\mu_1 + \mu_0)}{(\mu_1 - \mu_0)^2}$	انحراف معیار	۵- مقایسه دو نرخ * (حجم نمونه برای هر گروه)
	نرخها	μ_1 و μ_0
$\frac{\left\{ v\sqrt{[\pi_1(1-\pi_1)] + \pi_0(1-\pi_0)} + v\sqrt{[2\pi(1-\pi)]} \right\}^2}{(\pi_0 - \pi_1)^2}$	نسبتها	π_2 و π_1
		نسبت $\pi = \frac{\pi_1 + \pi_0}{2}$ و
	نسبت گروه شاهد دارای مواجهه	π_0
	نسبت گروه مورد دارای مواجهه که از رابطه $\pi_1 = \frac{\pi_0 \cdot OR}{1 + \pi_0 \cdot (OR - 1)}$	π_1
$\frac{\left\{ v\sqrt{[\pi_0(1-\pi_0)] + \pi_1(1-\pi_1)} + v\sqrt{[2\pi(1-\pi)]} \right\}^2}{(\pi_1 - \pi_0)^2}$	محاسبه می شود	۷- مطالعه مورد- شاهد (حجم نمونه برای هر گروه)
	نسبت شانس	OR
		$\pi = \frac{\pi_0 + \pi_1}{2}$

نوع محاسبه کمترین حجم نمونه	اطلاعات مورد نیاز	در همه حالت‌های ذکر شده
	<p>۷ درصدی از توزیع نرمال استاندارد که منتظر با نقطه "توان" ۱۰۰٪ در آزمون یک طرفه است.</p> <p>۷ درصدی از توزیع نرمال استاندارد که منتظر با سطح معنی داری آزمون دو طرفه است.</p> <p>برای اطلاعات بیشتر به جدول ج-۳ مراجعه شود.</p>	<p>در همه حالت‌های ذکر شده</p> <p>ب) وقت معلوم</p>
$\frac{\sigma^2}{e^2}$	انحراف معیار (استاندارد)	۸- تعیین میانگین
$\frac{\mu}{e^2}$	نرخ μ	۹- تعیین نرخ*
$\frac{\pi(1-\pi)}{e^2}$	نسبت π	۱۰- تعیین نسبت
$\frac{\sigma_1^2 + \sigma_2^2}{e^2}$	انحراف معیار σ_1 و σ_2	۱۱- تفاوت بین دو میانگین (حجم نمونه برای هر گروه)
$\frac{\mu_1 + \mu_2}{e^2}$	نرخها μ_1 و μ_2	۱۲- تفاوت بین دو نرخ* (حجم نمونه برای هر گروه)
$\frac{\pi_1(1-\pi_1) + \pi_2(1-\pi_2)}{e^2}$	نسبت‌ها π_1 و π_2	۱۳- تفاوت بین دو نسبت (حجم نمونه برای هر گروه)

در همه حالت‌های ذکر شده e حداکثر خطای معیار قابل قبول است.

* در این حالت‌ها واحد حجم نمونه همان واحدی است که نرخ بر حسب آن بیان شده است. برای مثال اگر نرخ بر حسب شخص سال باشد، عدد محاسبه شده تعداد شخص سال مورد نیاز را بیان می‌کند.

جدول ج-۳- مقادیر مختلف v به ازای مقادیر معین خطای نوع اول (α) و v به ازای مقادیر معین توان $(1-\beta)$ و خطای نوع دوم (β)

الف) برای محاسبه حجم نمونه بر اساس مقادیر منتخب خطای نوع اول:

مقدار v		α
آزمون دو دامنه	آزمون یک دامنه	
۱/۶۵	۱/۲۸	۰/۱
۱/۹۶	۱/۶۵	۰/۰۵
۲/۳۴	۱/۹۶	۰/۰۲۵
۲/۵۸	۲/۳۳	۰/۰۱

ب) v برای محاسبه حجم نمونه بر اساس مقادیر منتخب توان آزمون

β	$1-\beta$	v
۰/۵۰	۰/۵	۰/۰۰
۰/۴۰	۰/۶۰	۰/۲۵
۰/۳۰	۰/۷۰	۰/۵۳
۰/۲۰	۰/۸۰	۰/۸۴
۰/۱۵	۰/۸۵	۱/۰۳
۰/۱۰	۰/۹۰	۱/۲۸
۰/۰۵	۰/۹۵	۱/۶۵
۰/۰۲۵	۰/۹۷۵	۱/۹۶
۰/۰۱	۰/۹۹	۲/۳۳

ج-۵- تعیین حجم نمونه برای گروه‌های نامتعادل

اگر به جای انتخاب گروه‌های مساوی تصمیم بگیریم نسبت‌های مختلفی از بیماران را در هر گروه (مثلاً در یک مطالعه مورد-شاهدی برای هر مورد، ۳ شاهد) بررسی کنیم حجم کلی نمونه تغییر می‌کند. جدول ج-۴ عامل تعدیل برای استفاده در مطالعاتی که حجم دو گروه مورد مقایسه نامساوی است، کاربرد دارد. در این جدول عامل (f) برای تصحیح حجم نمونه گروه کوچک‌تر به کار می‌رود و برابر $f = \frac{c+1}{c}$ است. تعداد نمونه گروه بزرگ‌تر، c برابر حجم نمونه گروه کوچک‌تر است. بنابراین حجم نمونه کوچک‌تر $n \times f$ است (n حجم نمونه در حالت تساوی حجم دو گروه است) و حجم نمونه گروه بزرگ‌تر $n \times f \times c$ است [۲].

جدول ج-۴: عامل تعدیل برای محاسبه حجم نمونه هنگامی که حجم نمونه در گروه‌ها برابر نیست.

تعدیل حجم نمونه برای گروه کوچک‌تر (f)	نسبت حجم نمونه (گروه بزرگ‌تر به گروه کوچک‌تر)
۱	۱
۳/۴	۲
۲/۳	۳
۵/۸	۴
۳/۵	۵
۷/۱۲	۶
۴/۷	۷
۹/۱۶	۸
۵/۹	۹
۱۱/۲۰	۱۰

ج-۶- تعدیل حجم نمونه برای موارد گم‌شده از پیگیری‌ها، مخدوش‌کننده‌ها و اثر متقابل

در محاسبه حجم نمونه باید احتمال وجود موارد بدون پاسخ یا گم‌شدن اعضای مطالعه را نیز در نظر گرفت و بر این اساس حجم نمونه را افزایش داد. اگر در تحلیل نهایی اثر متغیرهای مخدوش‌کننده در نظر گرفته می‌شود و یا اگر آزمون اثرات زیرگروه‌ها مدنظر است، این موارد در محاسبه حجم نمونه نیز باید لحاظ شود. توجه به این نکته مهم است که افرادی که از ابتدا وارد مطالعه می‌شوند ممکن است در مرحله نتیجه‌گیری یا تحلیل نهایی مفید نباشد. مثلاً ممکن است اعضای نمونه از دسترس خارج شوند، یا پیگیری کافی نداشته باشند و یا اطلاعات متغیرهای کلیدی آن گم شود. حجم نمونه موردنیاز باید برای این حالات تعدیل شوند. اگر فرض کنیم که X درصد بیماری است که برای تحلیل نهایی مفید نیستند، آنگاه حجم نمونه باید در عامل $\frac{1}{1-X}$ ضرب شود. مثلاً اگر $x = 20\%$ باشد آنگاه عدد $\frac{1}{1-0.2} = 1.25$ در حجم نمونه نهایی ضرب می‌شود. عامل تعدیل برای $x\%$ داده گم‌شده است.

اثر کنترل متغیرهای مخدوش‌کننده در طراحی مطالعات مورد-شاهدی بررسی شده است. به ازای هر متغیر مخدوش‌کننده حجم نمونه بیشتر از ۱۰٪ افزایش می‌یابد [۳]. شواهد حاکی از آن است که می‌توان در حالت وجود چند مخدوش‌کننده هم‌زمان، افزایش حجم نمونه را به همان اندازه‌ای که متغیرها تک‌تک در نظر گرفته می‌شوند، محاسبه کرد [۴].

در بعضی حالات علاقمند به طراحی مطالعه‌ای هستیم که تفاوت‌های بین گروه‌ها را در زیرگروه‌های مختلف نشان دهیم. به عبارت دیگر علاقمند هستیم، اثر متقابل بین درمان و اثر دارو و یا برخی از مشخصه‌های بیماری که زیرگروه‌ها را تشکیل می‌دهند را پیدا کنیم. در چنین حالاتی حجم نمونه موردنیاز حداقل ۴ برابر بزرگ‌تر از وقتی

است که هدف بررسی دو گروه است [۳] و [۴].

توجه کنید که حجم نمونه بر اساس بهترین حدس از پارامتر مورد مطالعه محاسبه می‌شود. این عدد فقط در مورد تعداد افراد مورد مطالعه پیشنهادی ارائه می‌دهد. به عبارت دیگر این عدد برای تشخیص بین ۵۰ و ۱۰ مفید است اما برای تشخیص بین ۵۱ و ۵۲ کاربردی ندارد. گاهی اوقات حجم نمونه برای چند هدف مختلف محاسبه می‌شود. بدین ترتیب تصویر واضحی از فرصت‌های ممکن در مطالعه بدست می‌آید و بین عدد مورد علاقه و چیزی که واقعاً امکان پذیر است تعادلی به وجود می‌آید.

ج-۲- تعدیل حجم نمونه برای طرح‌های خوشه‌ای

مطالعاتی که در آن‌ها از طراحی خوشه‌ای استفاده می‌شود، نظیر کارآزمایی‌های تصادفی خوشه‌ای که انتساب‌های تصادفی به جای اشخاص به خوشه‌ها منتسب می‌شود، نظیر مطالعات فامیلی و مطالعاتی که دارای طرح نمونه‌گیری خوشه‌ای هستند، موارد استفاده زیادی در حیطه علوم پزشکی دارا هستند. طرح‌های خوشه‌ای به علت تشابه اعضای داخل هر خوشه نسبت به سایر اعضای جامعه، در به حجم نمونه بیشتری نسبت به طرح‌های معمولی نیاز است.

عاملی که باید در حجم نمونه معمولی ضرب شود تا حجم نمونه با دقت مورد نظر در مطالعه به دست آید، اثر طرح (Deff) خوانده می‌شود. اثر طرح به ضریب همبستگی داخل خوشه‌ای (Icc) وابسته است. این شاخص نسبت واریانس بین خوشه‌ها به واریانس کل است. ضریب همبستگی داخل خوشه‌ای:

$$Icc = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_e^2}$$

در این رابطه σ_u^2 واریانس بین خوشه‌ای، σ_e^2 واریانس درون خوشه‌ای (خطاها) و واریانس کل مجموع واریانس بین خوشه‌ها و درون خوشه‌ها است. محاسبه اثر طرح به وسیله رابطه:

$$\text{Deff} = 1 + (n_c - 1) \times Icc$$

که در آن n_c برابر میانگین حجم هر خوشه است، محاسبه می‌شود.
 به طور کلی دو عامل بر اندازه اثر طرح مؤثر هستند.
 - هر قدر Icc بزرگ‌تر باشد، اثر طرح بزرگ‌تر می‌شود.
 - هر قدر تعداد اعضای درون خوشه‌ها بیشتر باشد، اثر طرح بزرگ‌تر می‌شود.
 تعداد خوشه‌های مورد نیاز از رابطه زیر محاسبه می‌شود:

$$\text{تعداد خوشه‌ها} = \frac{n}{n_c} [1 + (n_c - 1) \times Icc]$$

در این رابطه n حجم نمونه در حالت عادی (بدون اثر خوشه‌ها) و n_e میانگین حجم اعضای داخل هر خوشه است. برآورد حجم نمونه، هنگامی که مطالعه در مرحله طراحی است معمولاً بسیار مشکل است. هرچند که تلاش‌هایی برای محاسبه ICC در شرایط مختلف انجام شده [۵]، اما این مقادیر به تعداد خوشه‌ها و حجم اعضای داخل هر خوشه وابسته است. نمایش نتایج در یک نمودار که تعداد خوشه‌ها در برابر اعضای داخل هر گروه را برای مقادیر مختلف ICC نشان می‌دهد، می‌تواند بسیار مفید باشد.

برای بررسی جزئیات بیشتر در محاسبه حجم نمونه در طرح‌های کارآزمایی تصادفی خوشه‌ای می‌توانید به منابع [۶] تا [۹] مراجعه نمایید.

ج-۸- مثال‌های واقعی

ج-۸-۱- مطالعات توصیفی

الف) مطالعه توصیفی برای تعیین شیوع

سازمان بهداشت جهانی طرح Vision 2020 را جهت پیشگیری، درمان و کاهش ۸۰ درصدی بیماری‌های منجر به کم بینایی و نابینایی تا سال ۲۰۲۰ را در دست اجرا دارد و این طرح در وزارت بهداشت، درمان و آموزش پزشکی نیز به‌عنوان یک اولویت مورد پذیرش قرار گرفته است. اولین گام در کنترل و پیشگیری بیماری‌ها وجود اطلاعاتی در باره شیوع و یا به عبارتی بار بیماری در منطقه تحت بررسی است. از آنجاکه در کشور ما اطلاعات بهنگام و مستند درباره شیوع و علل نابینایی و کم بینایی وجود ندارد. یک طرح تحقیقاتی در مرکز تحقیقات چشم دانشگاه علوم پزشکی شهید بهشتی برای پاسخ به این سؤال طراحی اجرا شد. در این طرح، شیوع تقریبی موارد نابینایی و کم بینایی بر اساس اطلاعات مطالعه راهنما که در سال ۱۳۸۱ در شهرستان کرج انجام گرفت، در حدود ۱۵ درصد در نظر گرفته شد و با توجه به آن حجم نمونه به‌صورت زیر محاسبه شد:

با استفاده از رابطه ۱۰ و با در نظر گرفتن $\pi = 0/15$ و $e = 0/004$ (حداکثر خطای معیار قابل قبول در مطالعه) داریم:

$$n_{\text{simple}} = \frac{0/15(0/85)}{(0/004)^2} = 7032$$

اثر طرح را با توجه به مطالعه راهنما برابر ۱/۵ در نظر گرفته شد. با توجه به نوع طرح نمونه‌گیری که یک مطالعه نمونه‌گیری خوشه‌ای با حجم ۲۰ خانوار در هر خوشه (تقریباً ۱۰۰ نفر در هر خوشه) است، حجم کلی نمونه برابر:

$$n_{\text{cluster}} = (\text{Design effect}) \times (n_{\text{simple}})$$

$$n_{\text{cluster}} = 1/5 \times 7032 = 10548$$

محاسبه می‌شود؛ اما نمونه محاسبه‌شده تنها افراد دارای همکاری کامل که اطلاعات آن‌ها در دسترس است را شامل می‌شود. معمولاً در مطالعات میدانی درصد همکاری، میزان مراجعه اعضای نمونه و افراد ارجاع شده به کلینیک کامل (۱۰۰ درصد) نیست و برای تعدیل این موضوع با توجه با در نظر گرفتن درصد عدم همکاری برابر

۲۵ درصد حجم نمونه نهایی برابر $n = \left(\frac{100}{100-25}\right) \times 10548 = 14064$ است. در نتیجه تقریباً به ۱۴۱ خوشه ۲۰ خانواری برای دستیابی به حجم نمونه مفید ۱۰۵۴۸ نیاز داریم [۹].

ب) مطالعه توصیفی برای تعیین بروز

قبل از تدارک یک مخزن بهداشتی آب آشامیدنی مایل هستیم میانگین تعداد کودکان زیر ۵ سال که در طول سال دچار اسهال می‌شوند را برآورد نماییم. حدس می‌زنیم بروز این عارضه در حدود ۳ باشد؛ اما می‌خواهیم این میزان را با حداکثر خطای معیار ۰/۲ برآورد کنیم. این موضوع به این معنی است که اگر برای مثال ۲/۶ عارضه/کودک/سال مشاهده شود بتوانیم استنباط کنیم مقدار واقعی میزان بروز بین ۲/۸-۲/۴ است. به بیان آماری، می‌خواهیم اندازه فاصله اطمینان ۹۵٪ میزان بروز از ۰/۲ \pm بزرگ‌تر نباشد. از آنجاکه طول فاصله اطمینان تقریباً برابر $2 \times se \pm$ است پس انحراف معیار برآورد نباید بزرگ‌تر از ۰/۱ باشد. با به‌کارگیری رابطه ۹ از جدول ج-۱ داریم:

$$n > \frac{3}{0.1^2} = 300$$

توجه به این نکته مهم است که رابطه‌های محاسبه حجم نمونه برای بروز (روابط ۲، ۵، ۹ و ۱۲) حجم نمونه را بر اساس واحد میزان به‌کاررفته محاسبه می‌کند. بنابراین در مثال بروز، نیاز است ۳۰۰ کودک/سال مشاهده شود. این موضوع با مشاهده ۳۰۰ کودک در یک سال و یا مثلاً ۱۲۰۰ کودک در سه ماه حاصل می‌شود.

ج-۸-۲ کارآزمایی بالینی تصادفی شده

الف) حجم نمونه برای مقایسه دو نسبت

مقایسه نتایج آناتومیک و ویتروکتومی اولیه و باکل اسکلارا در درمان جداشدگی شبکه در درمان دارای سابقه جراحی آب مروارید [۱۰].

در این مطالعه بیمارانی که دچار جداشدگی شبکه بودند و سابقه عمل جراحی کاتاراکت داشتند به‌صورت تصادفی در دو گروه قرار می‌گرفتند. در یک گروه عمل ویتروکتومی انجام می‌شد و گروه دیگر تحت عمل جراحی باکل اسکلارا قرار می‌گرفتند. در پایان وضعیت آناتومیک شبکه و دید بیماران در فواصل زمانی پیگیری (۱ و ۲ و ۴ و ۶ ماه بعد از عمل) مورد مقایسه قرار گرفتند.

برای محاسبه حجم نمونه این مطالعه از رابطه ۶ استفاده شد. در این رابطه با توجه به مطالعات قبلی مقادیر تقریبی $\pi_{vitx} = 30\%$ (درصد جداشدگی شبکه بعد از شش ماه پیگیری در گروه ویتروکتومی اولیه) $\pi_{S,B} = 40\%$ (درصد جداشدگی شبکه بعد از شش ماه پیگیری در گروه باکل اسکلارا) در نظر گرفته شد. در نتیجه برای سطح معنی‌داری $\alpha = 0.05$ و توان آزمون $\gamma = 80\%$ داریم:

$$n = \frac{\left\{ v\sqrt{[\pi_{vitx}(1-\pi_{vitx}) + \pi_{S,B}(1-\pi_{S,B})]} + v\sqrt{2\pi(1-\pi)}} \right\}^2}{(\pi_{vitx} - \pi_{S,B})^2}$$

$$n = \frac{\left\{ 0/84\sqrt{[0/3 \times (0/7) + 0/4 \times (0/6)]} + 1/96\sqrt{2 \times 0/35 \times (0/65)} \right\}^2}{(0/3 - 0/4)^2} = 125$$

یعنی تعداد ۱۲۵ نمونه در هر گروه لازم بود مطالعه واقعی با ۱۲۶ چشم در گروه باکل و ۹۹ چشم در گروه ویتراکتومی کامل شد که در این صورت توان واقعی آزمون اندکی کمتر از ۸۰ درصد است.

ب) حجم نمونه برای مقایسه دو میانگین

مقایسه کوتریموکسازول و روش کلاسیک در درمان توکسوپلاسموز چشمی [۱۱]

در یک کارآزمایی بالینی بیماران مبتلا به توکسوپلاسموز چشمی به صورت تصادفی در دو گروه درمانی قرار می‌گرفتند در یک گروه درمان کلاسیک توکسوپلاسموز (پریمیتامین) و در گروه دوم کوتریموکسازول تجویز شد. کلیه بیماران به مدت شش هفته تحت درمان گرفتند و در پایان بعد از ۶ ماه پیگیری بیماران از نظر تغییر در اندازه ضایعه چشمی، مورد مقایسه قرار گرفتند.

محاسبه حجم نمونه با استفاده از رابطه ۴ انجام پذیرفت در این رابطه $\mu_1 = 50$ (میانگین کاهش ضایعه با درمان کلاسیک) و $\mu_2 = 55$ (میانگین کاهش ضایعه با درمان به وسیله کوتریموکسازول) و واریانس هر دو گروه برابر ۲۱ در نظر گرفته شد. با سطح معنی‌داری $\alpha = 0/05$ و توان آزمون ۹۵ درصد حجم نمونه به صورت زیر محاسبه شد:

$$n = \frac{(v + v)^2 (\sigma_1^2 + \sigma_2^2)}{(\mu_1 - \mu_2)^2}$$

$$n = \frac{(1/65 + 2/57)^2 (21 + 21)}{(22 - 50)^2}$$

پس در هر گروه ۳۰ نمونه لازم است. مطالعه واقعی با تعداد ۲۹ نفر در گروه درمان کلاسیک و ۳۰ نفر در گروه درمان کوتریموکسازول انجام شد.

ج-۸-۳ مطالعه مورد-شاهدی

برای بررسی افزایش خطر مرگ پس از بیماری‌های عفونت حاد تنفسی در نوزادانی که از شیر خشک تغذیه می‌کنند نسبت به نوزادانی که از شیر مادر تغذیه می‌کنند، مطالعه‌ای طراحی می‌شود. با مادران گروه مورد (مادرانی که نوزادان آن‌ها مطابق ثبت در گواهی مرگ بر اثر بیماری عفونت تنفسی فوت کرده) از لحاظ نحوه

تغذیه نوزادان خود (شیر خشک یا شیر مادر) مصاحبه می‌شود. نتایج حاصله با داده‌های بدست آمده از مصاحبه با مادرانی که دارای نوزاد سالم هستند (گروه شاهد) مقایسه می‌شوند. برآورد می‌شود در حدود ۴۰٪ از گروه کنترل ($\pi_1 = 0.4$) از شیر خشک تغذیه می‌کنند. مایل هستیم اندازه اثری برابر $OR = 2$ در افزایش خطر مرگومیر قابل تشخیص و معنی‌دار شدن باشد. اندازه حجم نمونه از موردها و شاهد‌ها برای دستیابی به توان ۹۰٪ مدنظر است. برای این منظور مراحل زیر باید محاسبه شود.

۱- محاسبه π_1 ، یعنی نسبت موردهایی که با شیر خشک تغذیه می‌شوند.

$$\pi_1 = \frac{\pi_0 \times OR}{1 + \pi_0(OR - 1)} = \frac{0.4 \times 2}{1 + 0.4(2 - 1)} = \frac{0.8}{1.4} = 0.57$$

۲- محاسبه $\bar{\pi}$ ، یعنی میانگین π_1 و π_0

$$\bar{\pi} = \frac{0.4 + 0.57}{2} = 0.485$$

۳- محاسبه حداقل حجم نمونه موردنیاز

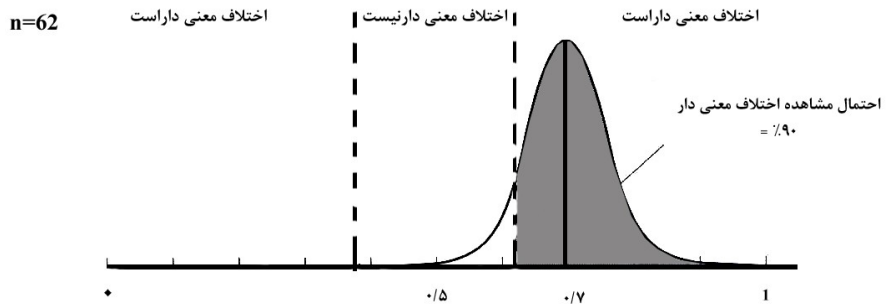
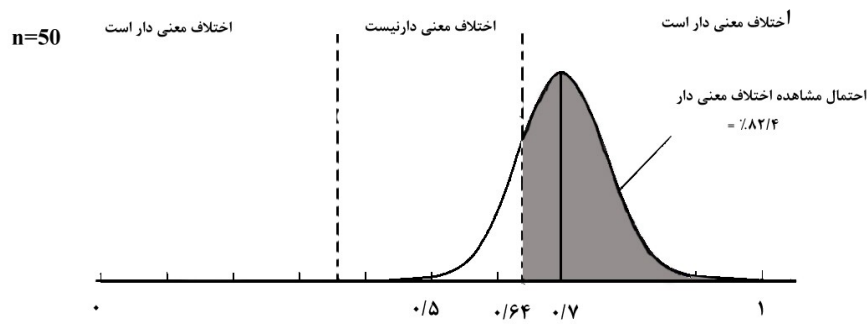
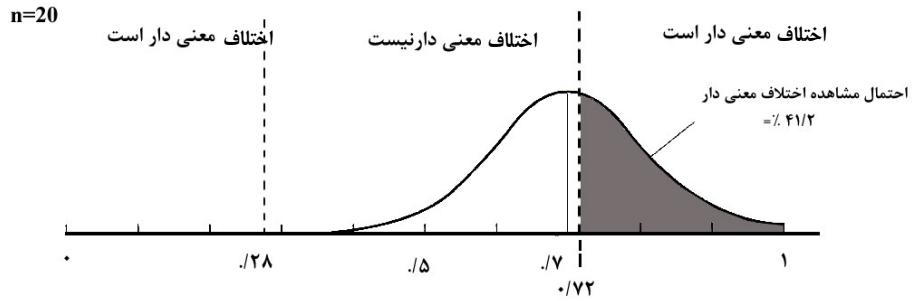
$$n > \frac{(1/28 \sqrt{(0.4 \times 0.6)} + 0.57 \times 0.43 + 1/96 \sqrt{2 \times 0.485 \times 0.515})^2}{(0.57 - 0.4)^2}$$

$$= \frac{(1/28 \sqrt{0.485} + 1/96 \sqrt{0.4996})^2}{0.17^2} = \frac{2/2769}{0.17^2} = 179/4$$

پس حداقل به ۱۸۰ مورد و ۱۸۰ شاهد یعنی حجم نمونه کلی ۳۶۰ نفر نیاز داریم.

ج-۹- کاربرد توان مطالعه در تفسیر آزمون‌های آماری

ابتدا به شکل ج-۲ مراجعه کنید. در این شکل احتمال به دست آوردن نتایج معنی‌دار در سطح ۵٪ با حجم نمونه‌های مختلف برای مقایسه دو نسبت، با توزیع فرضی نرمال مقایسه شده‌اند.



شکل ج-۲: احتمال به دست آوردن نتایج معنی دار در سطح ۵٪ با حجم نمونه‌های مختلف وقتی فرضیه " $\pi_1 = ۰/۷$ " در مقابل فرضیه صفر

" $H_0 = \frac{\pi_1}{\pi_0} = ۰/۵$ " آزمون می‌شود.

متأسفانه هنگامی که محققان سقف ۵٪ را برای معنی‌داری آماری در نظر می‌گیرند و استنباط می‌کنند که هر نتیجه‌ای که معنی‌دار نیست دلالت بر درست بودن فرضیه صفر دارد، از آزمون‌های معنی‌داری به‌درستی استفاده نشده است. یک تفسیر نادرست متداول دیگر این است مقدار احتمال را معادل احتمال درست بودن فرضیه صفر تفسیر کنیم.

جدول ج-۵ نشان می‌دهد که چرا چنین تصویری نادرست است. این جدول بر این اساس تنظیم شده که اگر ۱۰۰۰ فرضیه صفر متفاوت با سطح معنی‌داری ۵٪ تحت مفروضات زیر آزمون شود چه اتفاقی می‌افتد. ۱-۱۰٪ فرضیه‌های صفر آزمون شده واقعاً نادرست هستند (یعنی اثر عامل تحت بررسی واقعی است) و ۹۰٪ فرضیه‌های صفر آزمون شده درست هستند. تعداد زیادی از عوامل مورد بررسی در مطالعات اپیدمیولوژی از این نوع هستند. مثلاً در سال ۱۹۸۵ تقریباً ۳۰۰ عامل خطر برای بیماری کرونری قلبی مشخص شد اما واقعاً تعداد کمی از این عوامل خطر بیماری را افزایش می‌دهند.

۲- توان آزمون ۵۰٪ است که با مطالعات منتشر شده در مورد اندازه کارآزمایی‌های بالینی مطابقت دارد [۱۲]. نسبت زیادی از این نوع مطالعات با حجم نمونه ناکافی برای بررسی سؤال تحقیق انجام می‌شود. فرض اول در ستون مجموع جدول خلاصه شده است. در ۹۰۰ آزمون فرضیه صفر درست و در ۱۰۰ آزمون نادرست است. نسبت خطای نوع اول با سطح معنی‌داری مورد استفاده یعنی ۵٪ یکسان است. یعنی ۴۵ فرضیه از ۹۰۰ فرضیه صفر درست، به اشتباه رد شده است.

فرض دوم یعنی نسبت خطای نوع دوم برابر ۵۰٪ (توان -۱۰۰٪) است را در نظر می‌گیریم. پس برای رد کردن ۵۰ فرضیه از ۱۰۰ فرضیه صفر که نادرست هستند دچار خطا می‌شویم.

همان‌طوری که در جدول ج-۵-الف مشاهده می‌شود از ۹۵ آزمونی که نتیجه معنی‌دار دارند، ۵ آزمون درست است و ۴۵ (۴۷/۴٪) آزمون دارای خطای نوع اول هستند (نتیجه مثبت نادرست).

جدول ج-۵-ب جزئیات جدول ج-۵-الف را نشان می‌دهد. در این جدول درصد نتایج مثبت کاذب برای مقدار احتمال‌های مختلف و فرض‌های مختلف در مورد توان آزمون و نسبت صحیح فرضیه صفر ارائه شده است. برای هر سطح معنی‌داری با افزایش توان، نسبت نتایج معنی‌دار که همان مثبت کاذب است به شدت کاهش می‌یابد. با توجه به این جدول مقدار احتمال کمتر از ۰/۰۰۱ به‌جز مواردی که نسبت واقعه تحت بررسی بسیار کوچک است، دلیل محکمی بر ضد فرضیه صفر فراهم می‌آورد [۱۳] و [۱۴].

جدول ج-۵: کاربردهای توان مطالعه برای تفسیر آزمون‌های آماری

الف) نتایج آزمون‌های آماری برای ۱۰۰۰ فرضیه، هنگامی که ۱۰٪ فرضیه‌ها نادرست باشد، با استفاده از $\alpha = 0/05$ (سطح معنی‌داری) و توان ۵۰٪ نشان داده شده است.

نتیجه آزمون آماری	واقعیت	
	فرضیه صفر درست	فرضیه صفر نادرست
فرضیه صفر رد می‌شود ($\alpha < 0/05$)	۴۵ (خطای نوع اول)	۵۰
فرضیه صفر رد نمی‌شود ($\alpha > 0/05$)	۸۵۵	۵۰ (خطای نوع دوم)
کل	۹۰۰	۱۰۰

ب) نسبت مثبت کاذب در نتایج معنی‌دار شده متناظر با مقدار احتمال مورد استفاده برای معنی‌داری. توان مطالعه و نسبت مطالعاتی که فرضیه صفر واقعاً نادرست است.

درصد نتایج معنی‌داری که مثبت کاذب هستند				نسبت مطالعاتی که در آن‌ها فرضیه صفر نادرست است
$\alpha = 0/001$	$\alpha = 0/01$	$\alpha = 0/05$	توان مطالعه	
۰/۱	۱/۲	۵/۹	٪۲۰	٪۸۰
۰/۰	۰/۵	۲/۴	٪۵۰	
۰/۰	۰/۳	۱/۵	٪۸۰	
۰/۵	۴/۸	۲۰/۰	٪۲۰	٪۵۰
۰/۲	۲/۰	۹/۱	٪۵۰	
۰/۱	۱/۲	۵/۹	٪۸۰	
۴/۳	۳۱/۰	۶۹/۲	٪۲۰	٪۱۰
۱/۸	۱۵/۳	۴۷/۴	٪۵۰	
۱/۱	۱۰/۱	۳۶/۰	٪۸۰	
۳۳/۱	۸۳/۲	۹۶/۱	٪۲۰	٪۱
۱۶/۵	۶۶/۴	۹۰/۸	٪۵۰	
۱۱/۰	۵۵/۳	۸۶/۱	٪۸۰	

ج-۱۰- راهکارها و توصیه‌ها برای محاسبه حجم نمونه

♦ هنگام محاسبه حجم نمونه در مطالعات تحلیلی مراحل زیر باید انجام شود:

- ۱) تنظیم فرضیه‌های صفر و مقابل
- ۲) انتخاب آزمون‌های آماری مناسب
- ۳) برآورد اندازه اثر و میزان تغییرات آن از نتایج مطالعات قبلی یا مطالعه راهنما
- ۴) تعیین مقادیر مناسب α و β بر اساس اهمیت اجتناب از خطاهای نوع اول و دوم

- ♦ نکات مهم دیگری که باید در محاسبه حجم نمونه در مطالعات تحلیلی به آن‌ها توجه شود عبارت‌اند از:
 - تصحیح برای موارد بی پاسخی، وجود اثر متقابل و مخدوش‌کننده‌های بالقوه در مطالعه
 - ♦ مطالعاتی که در آن‌ها متغیرهای اسمی وجود دارند، تحلیل بقا، نمونه‌گیری خوشه‌ای، همسان‌سازی، مدل‌سازی و تحلیل چند متغیره و مطالعات مخصوص بررسی همانندی آزمودنی‌ها^{۴۳} مواردی هستند که محاسبه حجم نمونه برای آن‌ها با حالت‌های ساده‌ای که در این فصل به آن‌ها اشاره شد تفاوت دارند و توجه به راهبردهای خاص در این موارد ضروری است.
 - ♦ مراحل محاسبه حجم نمونه در مطالعات توصیفی که دارای فرضیه نیستند عبارت‌اند از:
 - (۱) برآورد نسبت اعضای دارای واقعه در متغیرهای اسمی و برآورد انحراف معیار در متغیرهای پیوسته
 - (۲) تعیین دقت (طول فاصله اطمینان) موردنظر
 - (۳) تعیین سطح اطمینان مطلوب (برای مثال سطح اطمینان ۹۵ درصد)
 - ♦ هنگامی که حجم نمونه با توجه به شرایط و محدودیت‌ها از قبل مشخص است، امکان محاسبه حداقل اندازه اثر قابل تشخیص در مطالعه و یا توان آزمون وجود دارد.
 - ♦ برخی راهکارها برای حداقل کردن حجم نمونه عبارت‌اند از استفاده از متغیرهای پیوسته، انجام اندازه‌گیری‌های دقیق‌تر، استفاده از طراحی همسان شده، محاسبه حجم نمونه برای گروه‌های نامتعادل و در نظر گرفتن پیشامدهای شایع‌تر در مطالعه
 - ♦ هنگامی که به نظر می‌رسد اطلاعات کافی برای برآورد حجم نمونه در دسترس نیست، مرور بیشتر و دقیق‌تر منابع، بررسی امکان اجرای یک مطالعه راهنما و یا تحلیل مجدد داده‌ها، امکان دسته‌بندی متغیرها به متغیرهای دوحالتی گزینه‌هایی هستند که امکان انجام آن‌ها باید به دقت بررسی شود. اگر هیچ کدام از این موارد عملی نباشند حدس‌های واقع بینانه توسط محققان با تجربه می‌تواند بسیار راهگشا باشد.
 - برخی از محققان کم تجربه اشتباهات شایعی هنگام محاسبه حجم نمونه مرتکب می‌شوند که برخی از آن‌ها عبارت‌اند از:
 - ♦ متغیرهای اسمی هنگامی که به صورت درصد یا میزان بیان می‌شوند ممکن است به اشتباه پیوسته در نظر گرفته شوند. برای مثال متغیری نظیر وضعیت زنده بودن (زنده یا مرده) ممکن است هنگامی که به صورت درصد زنده‌ها یا تعداد مرده‌ها در سال بیان می‌شود، پیوسته تعبیر شود. به‌طور مشابه، در تحلیل بقا متغیر دوحالتی موردبررسی ممکن است پیوسته تعبیر شود (نظیر میانگین بقا در ماه‌های مختلف). در تمام این موارد، پاسخ موردبررسی دوحالتی است و در حالت‌های ساده آزمون‌های مناسب برای این موارد از نوع کای-دو است.
 - ♦ حجم نمونه محاسبه شده باید برای تعداد افرادی که پیگیری‌ها را کامل می‌کنند و پاسخ‌های نهایی برای آن‌ها تکمیل می‌شود محاسبه شود نه برای تعداد افرادی که در ابتدا به مطالعه وارد می‌شوند.

ج-۱۱- بحث و نتیجه‌گیری

حالت‌هایی که برای محاسبه حجم نمونه در این فصل بررسی شده‌اند، موارد نسبتاً ساده‌ای هستند که در آن‌ها مقایسه دو گروه بدون هیچ‌گونه مفروضات پیچیده (نظیر وجود مخدوش‌کننده‌ها و غیره) مدنظر است. تعیین حجم نمونه در اغلب موارد به‌وسیله به‌کارگیری جداول از پیش تهیه شده بسیار راحت‌تر است. منابع [۱۵] و [۱۶] حاوی این مجموعه جداول هستند. حتی درچنین موقعیت‌های ساده‌ای نیز داشتن ایده‌های واقع بینانه نسبت به پارامترهای جامعه موردبررسی، اجتناب‌ناپذیر است. برای مثال هنگامی که هدف مقایسه میانگین‌ها است برآورد قابل‌قبولی از انحراف‌معیار و یا در مقایسه نسبت‌ها برآورد مناسبی از اندازه یکی از نسبت‌ها در جامعه موردنیاز است. این اطلاعات را می‌توان از مطالعات مشابه انجام شده و یا از مطالعه راهنما بدست آورد. در شرایط پیچیده‌تر به اطلاعات تکمیلی بیشتری احتیاج است.

مواردی نظیر محاسبه حجم نمونه برای داده‌های زوج شده، روش‌های بیز^{۴۴} برای یافتن حجم نمونه مناسب، کارآزمایی‌هایی که هدف آن‌ها نشان دادن تشابه بین درمان‌ها است^{۴۵}، محاسبه حجم نمونه برای داده‌های بقا^{۴۶}، محاسبه حجم نمونه هنگام وجود متغیرهای مخدوش‌کننده یا اثرمتقابل مواردی هستند که به مفروضات، اطلاعات تکمیلی و محاسبات پیچیده‌تری نیاز دارند که خارج از حوصله این فصل است. خوانندگان برای کسب اطلاعات بیشتر می‌توانند به منابع [۱۷] و [۱۸] و [۱۹] مراجعه نمایند.

در بسیاری از مطالعات لازم است پارامترهای مختلفی از اعضای نمونه اندازه‌گیری و تحلیل شوند. برای مثال در بررسی یک بیماری، محدوده وسیعی از سندرم‌ها، نتایج آزمایش‌های تشخیصی و معاینات بالینی در هر بیمار جمع‌آوری می‌شود. تعداد نمونه‌ای که برای یک هدف خاص محاسبه می‌شود ممکن است در شرایط دیگر و برای اهداف دیگر مناسب نباشد. در بسیاری از مطالعات حجم نمونه در شرایط مختلف و برای تمام اهداف محاسبه می‌شود و حجم نمونه نهایی، بزرگ‌ترین حجم نمونه محاسبه‌شده در حالت‌های مختلف در نظر گرفته می‌شود؛ اما در بعضی از مطالعات این روش به‌علت هزینه بعضی آزمایش‌ها (اندازه‌گیری‌ها) و یا مشکلات عملی امکان‌پذیر نیست. در چنین مواقعی استفاده از طرح نمونه‌گیری چندفازی^{۴۷} بسیار مفید است. در فاز اول پارامترهای معینی برای تمام افراد نمونه اندازه‌گیری می‌شود. در فاز دوم تنها زیرگروهی از نمونه اولیه انتخاب (به‌وسیله روش تصادفی ساده و یا سایر روش‌های انتخاب تصادفی) و سایر پارامترها اندازه‌گیری می‌شود. این روند می‌تواند برای فازهای مختلفی توسعه یابد.

-
۱. Bayesian method
 ۲. Equivalence trials
 ۳. Survival data
 ۴. Multiphase sampling

منابع

۱. محبی م ر، شاگو م، کاظم‌نژاد ل نکاتی در مورد کاربرد مقدار احتمال در آزمون فرض‌ها. اندیشه آماری، ۱۳۸۱.
2. Kirkwood RB, Jonathan A.C.S. Essential medical statistics 2nd ed. Blackwell Science, Cornwall 2003.
3. Smith PG, Day NE. The design of case-control studies: The influence of confounding and interaction effects. International Journal of Epidemiology 1984, 13: 356-365.
4. Breslow NE, Day NE. Statistical methods in cancer research. Volume1- The analysis of cohort studies. International Agency for Research on Cancer, Lyon 1987.
5. Gulliford MC, Ukoumunne OC, Chinn S. Components of variance and intraclass correlations for the design of community-based surveys and intervention studies: Data from Health Survey for England 1994. A M J Epidemiology 1999, 149: 876-883.
6. Donner A, Klar N. Design and analysis of cluster randomization trials in health research. Arnold, London 2000.
7. Ukoumunne OC, et al. Methods for evaluating area-wide and organization-based interventions in health and health care: A systematic review. Health Technology Assessment 3:1-92, 1999. Available at <http://www.hta.nhsweb.nhs.uk> (تاریخ دسترسی ۸۶/۱۲۰).
8. Hayes RJ, Bennet S. Simple sample size calculation for cluster-randomized trials. Int J Epidemiology 1999. 28: 319-326.
۹. سوری ح، جوادی م ع، رفعتی ن، ربانی‌خواه ز. بررسی میزان شیوع و علل نابینایی و کم بینایی در استان تهران. مجله چشم پزشکی بینا. شماره ۱۱، شماره ۲، ۱۳۸۴: ۱۶۳-۱۵۱.
۱۰. احمدیه ح، مرادیان س، فقیهی ه و همکاران مقایسه ویتراکتومی اولیه و باکل اسکلرا در درمان جداشدگی شبکیه در بیماران دارای سابقه جراحی آب مروارید. مجله چشم پزشکی بینا، شماره ۱، سال ۹، ۱۳۸۲.
۱۱. صدوقی م م، سهیلیان م، دهقان فیروزآبادی و همکاران. مقایسه کوتریموکسازول و روش کلاسیک در درمان توکسوپلاسماوز چشمی. م مجله چشم پزشکی بینا، شماره ۱، سال ۹، ۱۳۸۲.
12. Moher D, Dulberg CS, Wells GA. Statistical power, sample size, and their reporting in randomized controlled trials. JAMA 1994, 272:122-124.
13. Oakes M. Statistical inference. Wiley, Chichester 1986.
14. Sterne JAC, Davey-Smith G. Sifting the evidence-what's wrong with significance tests? BMJ 2001,322: 226-231.
15. Lemeshow S, Hosmer DW, Klar J, Lwanga SK. Adequacy of sample size in health studies. Wiley, Chichester 1990.
16. Machine D, Campbell MJ, Fayers PM, Pinol ASPY. Sample size tables for clinical studies, 2nd ed. Blackwell, Oxford 1997.
17. Armitage P, Berry G, Matthews JNS. Statistical methods in medical research (4th ed.) Blackwell, Cornwall 2002.
18. Chow SC, Liu JP. Design and analysis of clinical trials: Concepts and methodologies 2nd ed. Wiley, New York 2004.
۱۹. چانگ چاو ش و همکاران. حجم نمونه در پژوهش‌های بالینی. ترجمه: علوی مجد ح. دانشگاه علوم پزشکی شهید بهشتی. زمستان ۱۳۸۵.

نمایه

اشتباه، ۲۴، ۳۵، ۳۶، ۵۴، ۹۰، ۹۸، ۹۹، ۱۱۳، ۱۵۱،	الف
۲۳۸، ۲۳۹، ۲۴۵، ۲۶۱، ۲۷۲، ۲۷۴	ابزار اندازه‌گیری، ۹۵، ۹۸، ۱۰۰، ۱۰۱، ۱۰۳، ۱۰۶،
اعتبار، ۲۷، ۳۰، ۳۴، ۳۵، ۳۶، ۳۹، ۵۳، ۵۵، ۷۵، ۷۷،	۲۰۸
۷۹، ۸۹، ۹۹، ۱۰۰، ۱۰۳، ۱۰۴، ۱۰۵، ۱۰۷،	ابزار جمع‌آوری داده‌ها، ۱۱، ۲۴۲
۱۰۹، ۱۱۶، ۲۲۳، ۲۲۶	اثر انتقالی، ۸۵
اعتبار تجربی، ۱۰۷	اثر متقابل، ۱۳، ۵۳، ۸۳، ۱۳۹، ۱۵۲، ۱۵۳
اعتبار خارجی، ۱۰۷	اجرای مطالعه، ۲۱، ۳۴، ۳۶، ۹۳، ۱۳۲، ۱۳۵، ۱۹۸،
اعتبار داخلی، ۱۰۷	۲۱۱، ۲۱۸، ۲۱۹، ۲۲۰، ۲۲۹
اعتبار ساختاری، ۱۰۷	اختصاصی بودن همبستگی، ۱۳۹
اعتبار ظاهری، ۱۰۷، ۲۵۱	ارزش اخباری آزمون مثبت، ۱۱۴
اعتبار محتوا، ۱۰۷	ارزش اخباری آزمون منفی، ۱۱۴
اعضای موردبررسی، ۱۹، ۱۱۴	ارزش اخباری مثبت و منفی، ۱۰۸، ۱۱۵، ۱۱۸
آمارهای توصیفی، ۱۷۰، ۱۷۱	ارزیابی آزمون تشخیصی، ۱۱۵
انتخاب بیماران، ۵۸، ۷۱، ۷۸، ۸۸، ۱۹۹، ۲۱۱	ارزیابی اعتبار، ۱۰۴
انتخاب روش مطالعه، ۳۹	ارزیابی پاسخ بیمار، ۷۳
انتخاب شاهدها، ۵۴، ۵۷، ۵۹	ارزیابی پیشرفت کار آزمایی، ۸۶
انتخاب موردها، ۵۴، ۵۵، ۵۶، ۵۷، ۲۰۱	ارزیابی دقت، ۳۶، ۱۰۱، ۲۲۷
انتخاب موردهای جدید، ۵۵، ۵۶	آزمون آلفای کرونباخ، ۲۴۹
انتخاب موردهای موجود، ۵۵، ۵۶	آزمون لگ-رتبه، ۱۸۹، ۱۹۱
انتساب تصادفی بیماران، ۷۴	آزمون منفی، ۱۱۴، ۱۲۳
انحراف معیار، ۳۲، ۱۰۸، ۱۱۱، ۱۷۱، ۱۷۲، ۱۷۳،	آزمون‌های تشخیصی، ۱۴، ۹۸، ۱۰۸، ۱۰۹، ۱۱۳،
۱۷۴، ۱۸۰، ۱۸۲، ۱۸۹، ۲۲۶، ۲۷۴	۱۱۴، ۱۱۵، ۱۳۰
اندازه اثر: ۲۸، ۲۹، ۳۰، ۳۱، ۴۱، ۷۶، ۸۳، ۱۲۳،	استاندارد کردن غیرمستقیم، ۱۶۵
۱۲۴، ۱۳۳، ۱۳۶، ۱۴۹، ۱۷۷، ۲۱۹، ۲۵۷، ۲۶۴،	استاندارد کردن مستقیم، ۱۶۳، ۱۶۴
۲۶۸، ۲۷۱، ۲۷۲	استنباط سببی، ۱۳۴، ۱۴۰

اندازه حجم نمونه، ۲۵۸	پیگیری کوتاه مدت، ۲۰۴
اندازه کارآزمایی بالینی، ۸۶	ت
اندازه گیری متغیرهای پایه، ۷۲	تابع بقا، ۱۴، ۱۸۷، ۱۸۸، ۱۸۹، ۲۸۷
اندازه های خلاصه، ۴۹	تجزیه و تحلیل چند متغیره، ۱۵۷
اندازه های کلان، ۴۹	تحلیل چند متغیره، ۸۰، ۱۴۰، ۱۵۷، ۱۶۰، ۱۸۰،
اندازه های محیطی، ۴۹	۱۸۵، ۲۲۴، ۲۷۵، ۲۸۷
ب	تحلیل داده ها، ۱۱، ۱۶، ۲۱، ۳۳، ۹۲، ۹۳، ۱۵۷،
بازبینی منظم، ۹۰، ۹۳	۱۵۹، ۱۶۰، ۱۶۹، ۱۸۰، ۱۹۰، ۱۹۳، ۱۹۴، ۱۹۵،
بانک های اطلاعاتی، ۵۰، ۹۱، ۱۷۰	۱۹۶، ۱۹۹، ۲۱۱، ۲۱۳، ۲۲۲، ۲۲۶، ۲۲۹، ۲۳۱،
برآورد اثر، ۹۳، ۱۸۰، ۱۹۰	۲۳۷، ۲۵۲
برآورد خطر، ۵۳، ۱۴۶، ۱۵۹، ۱۷۸	تحلیل داده های بقا، ۱۸۰، ۱۸۷
برچسب گذاری، ۲۲۷، ۲۲۸	تحلیل مجدد داده، ۹۲
بروز، ۱۴، ۲۳، ۳۶، ۴۰، ۴۴، ۴۹، ۵۰، ۵۲، ۶۰، ۶۱	تحلیل مجدد داده ها، ۳۹، ۴۱، ۸۸، ۸۹، ۲۷۱
۶۲، ۶۴، ۶۵، ۶۷، ۸۳، ۸۹، ۱۰۲، ۱۰۸، ۱۳۵	ترکیب اثر مواجهه ها، ۱۴۸
۱۳۷، ۱۴۰، ۱۴۱، ۱۴۴، ۱۵۱، ۱۵۲، ۱۵۳، ۱۶۲	تصادف، ۱۷، ۲۲، ۳۰، ۴۳، ۷۸، ۸۰، ۸۱، ۸۲، ۸۷
۱۶۳، ۱۶۴، ۱۶۵، ۱۷۰، ۱۷۵، ۱۷۹، ۱۸۳، ۲۰۴	۱۰۷، ۱۳۲، ۲۰۲، ۲۴۱
۲۰۵، ۲۰۸، ۲۰۹، ۲۱۰، ۲۱۱، ۲۲۶، ۲۳۹، ۲۶۰	تصادفی کردن، ۴۵، ۷۴، ۷۸، ۸۳، ۹۳، ۱۳۸، ۱۵۰،
۲۶۸	۱۵۶، ۲۱۲
بلوک های جایگشتی تصادفی، ۷۵	تصادفی سازی با جایگذاری، ۷۲
بهبود پروتکل، ۲۳۰	تصادفی سازی طبقه بندی شده، ۷۲
پ	تصادفی سازی طبقه شده، ۷۲
پایش بیماران، ۲۰۸	تضعیف اثرها در مدل رگرسیون، ۱۲۳
پرسشنامه، ۱۱، ۲۱، ۲۴، ۲۵، ۳۵، ۳۶، ۴۳، ۶۹، ۹۲	تغییرات ذاتی اعضای مورد مطالعه، ۹۷
۱۰۱، ۱۰۲، ۱۱۱، ۱۱۲، ۲۰۴، ۲۰۵، ۲۲۰، ۲۲۹	تغییرات متغیر پاسخ، ۳۰
۲۴۱، ۲۴۲، ۲۴۳، ۲۴۴، ۲۴۵، ۲۴۶، ۲۴۹، ۲۵۰	تفاضل خطر، ۱۷۲، ۱۷۳، ۱۷۵
۲۵۱، ۲۵۲، ۲۵۳	تفسیر آزمون های آماری، ۱۴، ۲۶۵، ۲۶۸
پروپوزال، ۱۱، ۱۴، ۱۵، ۱۹، ۲۰، ۲۵، ۲۹، ۳۰، ۳۳	تکرارپذیری، ۷۴، ۹۶، ۹۸، ۱۰۶، ۱۳۶، ۱۳۸، ۲۲۳،
۳۶، ۷۳، ۹۳، ۱۷۰، ۲۱۳، ۲۱۹، ۲۳۳، ۲۳۴	۲۳۸، ۲۴۵، ۲۴۷، ۲۵۱
۲۳۵، ۲۳۶، ۲۳۷، ۲۳۸، ۲۳۹، ۲۴۲	تکمیل فرآیند پذیرش بیماران، ۷۷
پروتکل، ۲۱، ۷۰، ۷۱، ۸۰، ۸۷، ۸۸، ۱۰۲، ۲۱۴	تنظیم پروپوزال، ۱۹، ۳۴، ۲۲۹، ۲۳۰
۲۱۹، ۲۲۰، ۲۲۲، ۲۲۵، ۲۳۱، ۲۳۴	توالی زمانی، ۵۰، ۲۰۷، ۲۳۳
پیش آزمون، ۱۱، ۲۱۹، ۲۲۱، ۲۲۲، ۲۵۲	توالی سؤال ها، ۲۴۷

تورش زبان، ۲۱۰	توان آزمون، ۲۹، ۲۵۶، ۲۵۷، ۲۶۱، ۲۶۵، ۲۶۶
تورش سوء طبقه‌بندی، ۲۰۵	۲۶۹، ۲۷۰، ۲۷۱، ۲۷۵
تورش طبقه‌بندی، ۲۰۵	توان آماری، ۳۱، ۵۰، ۸۱، ۹۷، ۱۰۱، ۱۳۵، ۱۵۶
تورش کارآزمایی بزرگ، ۲۱۲	۲۵۶، ۲۵۱
تورش کارآزمایی چندمرکزی، ۲۱۲	توان مطالعه، ۷۲، ۲۵۵، ۲۶۷، ۲۷۰
تورش کارآزمایی کوچک، ۲۱۲	توجیه بیولوژیکی، ۱۳۶، ۱۳۸
تورش مجله خوشنام، ۲۱۱	تورش، ۱۴، ۱۷، ۱۸، ۳۳، ۳۸، ۴۷، ۴۹، ۵۰، ۵۲
تورش مجله غیر خوشنام، ۲۱۱	۵۴، ۵۵، ۵۷، ۵۸، ۶۰، ۷۶، ۸۶، ۹۰، ۹۵، ۹۶
تورش مشاهده‌گر، ۵۷، ۱۰۰، ۲۰۱، ۲۰۲	۱۰۰، ۱۰۱، ۱۰۲، ۱۰۵، ۱۰۶، ۱۲۳، ۱۳۱، ۱۳۲
تورش مصاحبه‌گر، ۲۰۰، ۲۰۱	۱۳۳، ۱۳۴، ۱۳۷، ۱۴۹، ۱۹۵، ۱۹۶، ۱۹۷، ۱۹۸
تورش موسسه مشهور، ۲۱۱	۱۹۹، ۲۰۰، ۲۰۱، ۲۰۲، ۲۰۳، ۲۰۴، ۲۰۵، ۲۰۶
تورش موسسه ناشناخته، ۲۱۱	۲۰۷، ۲۰۸، ۲۰۹، ۲۱۰، ۲۱۱، ۲۱۲، ۲۲۰، ۲۲۱
تورش ناشی از تأخیر زمانی، ۲۱۰	۲۲۴، ۲۵۰، ۲۵۱
تورش نسبت دوره بیماری، ۲۰۶	تورش ابزار اندازه‌گیری، ۱۰۱
تورش نسبت مکمل شیوع، ۲۰۶	تورش اخلاقی، ۲۱۱
تورش نشر، ۹۰، ۲۱۰، ۲۱۹	تورش ارجاع به، ۲۰۰
تورش نویسنده سرشناس، ۲۱۱	تورش اطلاع، ۱۹۵، ۱۹۷، ۲۰۰
تورش نویسنده ناشناس یا غیر سرشناس، ۲۱۱	تورش اعضای نمونه، ۱۰۰
تورش‌های متداول در کارآزمایی‌های بالینی، ۲۰۷	تورش اکولوژیک، ۴۷
تورش‌های مرتبط با مطالعات مقطعی، ۲۰۵	تورش انتخاب، ۵۲، ۵۴، ۱۹۵، ۱۹۶، ۱۹۷، ۱۹۸
توزیع نرمال، ۹۷، ۱۱۷، ۱۶۹، ۱۷۰، ۱۷۱، ۱۸۶	۱۹۹، ۲۰۶، ۲۰۸
توزیع نرمال استاندارد، ۱۷۰	تورش برکسون، ۱۹۶
ث	تورش بقا، ۲۰۶
ثبات (تکرارپذیری و سازگاری) یافته‌ها، ۱۳۶	تورش پاسخگویی، ۲۰۲
ج	تورش پایش، ۱۹۶، ۲۰۲، ۲۰۳
جبران سوگرایی، ۱۹۹	تورش پی‌بردن، ۲۰۹
جدول توافقی، ۱۱۱، ۱۴۳، ۱۴۴، ۱۷۱	تورش تقدم پیامد بر مواجهه، ۲۰۵، ۲۰۷
جدول عمر، ۱۸۳، ۱۸۸	تورش توان خارق‌العاده، ۲۱۱
جمع‌آوری داده‌ها، ۴۶، ۴۷، ۸۸، ۱۴۴، ۲۲۰، ۲۲۱	تورش جغرافی، ۲۱۲
حجم نمونه، ۱۷، ۱۸، ۱۹، ۲۰، ۲۱، ۲۳، ۲۸، ۲۹	تورش خاطره، ۲۰۰، ۲۰۱، ۲۰۲، ۲۰۳
۳۰، ۳۱، ۳۲، ۳۳، ۴۳، ۵۰، ۵۷، ۶۲، ۶۸، ۶۹	تورش داوطلب‌ها، ۲۰۰
۷۵، ۷۹، ۸۰، ۸۳، ۸۴، ۸۶، ۹۰، ۹۵، ۱۳۲، ۱۳۵	تورش دوره بیماری، ۲۰۶

- داده‌های گم شده، ۶۰، ۹۰، ۱۶۷، ۱۶۸، ۲۱۰، ۲۴۱
 داده‌های نادرست، ۲۲۰، ۲۲۱
 درصد مثبت، ۱۱۲
 دستورالعمل CONSORT، ۷۴، ۷۵، ۷۷
 دستورالعمل عملیاتی تحقیق، ۲۲۱
 دقت، ۲۱، ۲۴، ۲۵، ۳۲، ۳۳، ۳۴، ۵۷، ۶۸، ۷۶، ۸۴، ۹۰، ۹۶، ۹۷، ۹۸، ۹۹، ۱۰۰، ۱۰۱، ۱۰۳، ۱۰۶، ۱۳۲، ۱۳۵، ۱۳۵، ۱۷۱، ۱۷۷، ۲۱۶، ۲۲۱، ۲۲۴، ۲۳۰، ۲۳۳، ۲۳۴، ۲۳۵، ۲۴۶، ۲۴۷، ۲۴۹، ۲۵۱، ۲۵۵، ۲۵۶، ۲۵۷، ۲۵۹، ۲۶۳، ۲۷۱
- ر
 رابطه اثر- علت، ۱۳۴، ۱۳۴
 رابطه تصنعی، ۱۳۲، ۲۰۷
 رابطه زمانی، ۱۳۶، ۱۳۷، ۱۳۸
 رابطه زمانی مواجهه، ۱۳۷
 رابطه غیر سببی، ۱۴۴
 رگرسیون پوآسن، ۱۷۶، ۱۷۹، ۱۸۱، ۱۸۷، ۱۸۸
 رگرسیون حول میانگین، ۱۲۳، ۱۲۴
 رگرسیون خطی، ۱۲۳، ۱۷۶، ۱۷۷، ۱۷۹، ۱۸۶، ۱۸۷
 رگرسیون کاکس، ۱۲۴، ۱۷۶، ۱۸۶، ۱۸۸
 رگرسیون لجستیک، ۱۷۶، ۱۸۱، ۱۸۷، ۱۸۸، ۱۹۲
 روابط تصنعی، ۱۳۱، ۱۳۲
 روابط حقیقی، ۱۳۱، ۱۳۴، ۱۳۷
 روابط سببی، ۵۰، ۱۳۶، ۱۳۷
 روش استاندارد مرجع، ۱۰۱، ۱۰۵
 روش تحلیل داده‌ها در مطالعات کارآزمایی تصادفی، ۱۹۰
 روش تحلیل داده‌ها در مطالعات متقاطع و طول، ۱۹۱
 روش تصادفی ساده، ۷۱، ۲۷۲
 روش سکه اریب، ۷۲
- ح
 حداقل اختلاف مهم از نظر بالینی، ۲۸
 حذف اثر تورش، ۱۹۸
 حساسیت، ۱۰۱، ۱۰۲، ۱۰۵، ۱۰۶، ۱۱۰، ۱۱۱
 ۱۱۲، ۱۱۳، ۱۱۶، ۱۱۷، ۱۱۹، ۱۲۰، ۱۲۴، ۱۲۵، ۲۱۸
- خ
 خطاها در اندازه‌گیری، ۹۵
 خطاهای تصادفی، ۳۲، ۳۳، ۳۴، ۹۵، ۹۷، ۹۸، ۹۹، ۱۳۲، ۱۹۵
 خطاهای منظم، ۳۲، ۳۳، ۷۴، ۹۵، ۱۰۰، ۱۰۲، ۱۹۵
 خطاهای نمونه‌گیری، ۹۵
 خطای اندازه‌گیری، ۱۰۸، ۱۲۳، ۱۲۴
 خطای تصادفی، ۱۷، ۳۲، ۳۳، ۹۵، ۹۹، ۱۰۰، ۱۰۱
 خطای منظم، ۱۷، ۳۲، ۳۳، ۹۵، ۹۶، ۹۶، ۱۰۰، ۱۰۱، ۱۰۳، ۱۳۱
 خطای نوع اول، ۲۹، ۱۳۲، ۲۵۶، ۲۵۷، ۲۵۸، ۲۶۱، ۲۶۹
 خطای نوع دوم، ۲۹، ۲۵۶، ۲۵۷، ۲۵۸، ۲۶۱، ۲۵۹، ۲۷۰
 خطر قابل انتساب، ۱۶۰، ۱۷۲
 خطر قابل انتساب تعدیل یافته، ۱۶۰
 خطر نسبی، ۵۰، ۱۳۵، ۱۴۱، ۱۴۲، ۱۴۳، ۱۴۴، ۱۶۰، ۱۶۲، ۱۶۳، ۱۷۲، ۱۷۳، ۱۷۵، ۱۷۶، ۲۰۵
 خطر نسبی تعدیل یافته، ۱۶۰
- د
 داخل رده‌ای، ۱۰۳، ۱۰۸
 داده‌های دور افتاده، ۱۶۷، ۱۶۸

شیوع استاندارد، ۱۶۲	روش طبقه‌بندی، ۱۳۶، ۱۵۶، ۱۵۷، ۱۵۹
ص	روش مانتل - هانزل، ۱۶۳، ۱۶۴، ۱۸۷، ۱۸۸
صحت آزمون‌ها، ۱۰۹، ۱۱۰	۱۹۲
صحت داده‌ها، ۱۶۷، ۱۶۸، ۲۲۶	روش مرجع استاندارد، ۱۰۱
ض	روش مک - نما، ۱۶۴
ضریب کاپا، ۱۰۵، ۱۰۷	روش‌های تصادفی کردن، ۷۱
ضریب همبستگی، ۱۰۸، ۱۲۱، ۱۲۳، ۱۲۵، ۱۸۷	روش‌های نمونه‌گیری، ۲۳
۲۶۳، ۲۷۸	ز
ضریب همبستگی داخل رده‌ای، ۱۰۸، ۱۲۳	زوج‌های همسان، ۸۰
ط	س
طبقه‌بندی، ۲۲، ۲۵، ۴۱، ۴۵، ۴۷، ۵۱، ۶۳، ۷۲، ۷۵	ساختار پرسشنامه، ۲۴۰
۷۶، ۹۵، ۱۲۳، ۱۳۶، ۱۴۹، ۱۵۳، ۱۵۶، ۱۵۷	ساختار پژوهش، ۱۸، ۱۹
۱۵۹، ۱۶۴، ۱۶۵، ۱۸۱، ۱۹۲، ۱۹۵، ۲۰۰، ۲۰۳	سازگاری اطلاعات، ۱۶۸
۲۰۴، ۲۰۵	سازگاری داده‌ها، ۲۲۰
طراحی با شاهد‌های متوالی، ۸۴	سوء طبقه‌بندی بدون جهت، ۲۰۳، ۲۰۴، ۲۰۵
طراحی پرسشنامه، ۲۳۸، ۲۴۶	سوء طبقه‌بندی جهت‌دار، ۲۰۴، ۲۰۵
طراحی خوشه‌ای، ۲۶۳	ش
طراحی متقاطع، ۸۱، ۸۲	شانس، ۱۷، ۲۸، ۲۹، ۳۰، ۳۳، ۴۹، ۵۰، ۵۲، ۵۴
طراحی مربع لاتین، ۸۱	۹۷، ۱۰۵، ۱۰۷، ۱۳۱، ۱۳۲، ۱۴۱، ۱۴۳، ۱۴۴
طراحی مطالعه، ۲۰، ۲۱، ۲۸، ۳۳، ۴۵، ۴۷، ۸۹	۱۴۹، ۱۵۰، ۱۵۳، ۱۵۷، ۱۵۹، ۱۶۲، ۱۶۳، ۱۶۴
۱۰۸، ۱۱۰، ۱۳۳، ۱۴۴، ۱۵۳، ۲۰۱، ۲۱۱، ۲۱۹	۱۷۱، ۱۷۲، ۱۷۳، ۱۷۴، ۱۷۶، ۱۷۷، ۱۸۱، ۱۸۲
۲۲۵، ۲۵۵، ۲۶۲	۱۸۷، ۱۹۲، ۱۹۷، ۱۹۸، ۱۹۹، ۲۰۲، ۲۰۳، ۲۰۴
طراحی موازی، ۷۸	۲۰۵، ۲۱۲، ۲۲۷، ۲۲۹، ۲۵۵، ۲۵۶
طرح مربع لاتین، ۸۱	شاهد‌ها، ۱۸، ۲۲، ۳۹، ۴۲، ۴۶، ۴۹، ۵۰، ۵۱، ۵۴
طرح مطالعه، ۱۸، ۱۹، ۲۱، ۲۲، ۲۶، ۲۸، ۳۲، ۴۹	۵۵، ۵۶، ۵۷، ۷۸، ۸۱، ۸۴، ۸۵، ۱۳۳، ۱۳۴
۶۰، ۶۲، ۱۰۵، ۱۳۲، ۱۳۳، ۱۸۹، ۱۹۵، ۲۳۱	۱۳۵، ۱۵۴، ۱۵۵، ۱۵۶، ۱۶۳، ۱۶۴، ۱۹۲، ۱۹۶
۲۳۲، ۲۳۴، ۲۵۶	۱۹۸، ۱۹۹، ۲۰۱، ۲۰۴، ۲۰۵، ۲۶۷
طیف مقیاس پیوسته، ۲۳۹	شاهد‌های هم‌زمان موازی، ۷۸
ع	شیوع، ۲۲، ۲۳، ۲۶، ۳۳، ۳۴، ۳۸، ۴۰، ۴۸، ۴۹، ۵۵
علت - معلول، ۱۳۱	۸۷، ۱۱۳، ۱۱۴، ۱۱۵، ۱۱۶، ۱۲۹، ۱۳۷، ۱۳۸
ف	۱۴۱، ۱۴۴، ۱۶۲، ۱۷۲، ۱۸۱، ۱۸۲، ۲۰۵، ۲۰۶
فاصله اطمینان، ۷۶، ۱۲۳، ۱۲۷، ۱۳۲، ۱۷۱، ۱۷۳	۲۰۷، ۲۵۷، ۲۶۴، ۲۷۳

- ۱۷۴، ۱۷۵، ۱۷۶، ۱۷۹، ۱۸۱، ۱۸۲، ۱۸۶، ۱۸۷، ۱۸۸، ۲۶۵، ۲۷۱
 فرضیه پژوهش، ۲۶
 فرضیه دوطرفه، ۲۷، ۲۸
 فرضیه ساده، ۲۶
 فرضیه صفر، ۲۷، ۲۹، ۳۰، ۱۲۳، ۲۵۶، ۲۵۷، ۲۵۸، ۲۶۸، ۲۶۹، ۲۷۰
 فرضیه مرکب، ۲۶
 فرضیه مقابل، ۲۷، ۳۰
 فرضیه یک طرفه، ۲۷، ۲۸
 ق
 قابل اجرا بودن، ۲۰
 قدرت همبستگی، ۱۳۶
 قطع مواجهه، ۱۳۶، ۱۳۸
 ک
 کپلان- مایر، ۱۸۳، ۱۸۴، ۱۸۵، ۱۸۸
 کارآزمایی با شاهد های خارجی، ۸۵
 کارآزمایی با شاهد های هم زمان، ۸۷
 کارآزمایی بالینی، ۲۴، ۲۶، ۶۶، ۵۷، ۶۸، ۶۹، ۷۰، ۷۴، ۷۸، ۸۳، ۸۴، ۸۵، ۲۰۸، ۲۶۵، ۲۶۶
 کارآزمایی تصادفی، ۲۲، ۷۱، ۸۵، ۸۶، ۱۸۳، ۱۹۰، ۲۶۴
 کارآزمایی تصادفی نشده، ۸۵، ۸۶
 کارآزمایی های بالینی، ۲۳، ۴۳، ۶۶، ۶۷، ۶۸، ۷۰، ۷۴، ۷۵، ۷۷، ۷۸، ۸۳، ۸۴، ۸۶، ۹۲، ۱۸۹، ۱۹۹، ۲۰۷، ۲۰۸، ۲۱۰، ۲۲۳، ۲۶۹
 کارآزمایی های بالینی تصادفی شده، ۶۶، ۷۵، ۱۸۹، ۲۰۸، ۲۱۰
 کارآزمایی های تصادفی، ۲۸، ۴۳، ۸۶، ۸۹، ۱۰۵، ۱۶۷، ۱۸۹، ۲۶۳
 کارآزمایی های تصادفی خوشه ای، ۸۶، ۲۶۳
 کارآزمایی های چندمرکزی، ۸۴
- کارآزمایی های میدانی، ۶۷، ۸۶
 کارایی، ۲۳، ۲۷، ۶۱، ۶۲، ۶۹، ۷۱، ۸۷، ۱۱۰، ۱۴۴، ۲۱۰، ۲۲۰، ۲۲۵
 کنترل خطاهای تحقیق، ۳۲
 کنترل کیفیت، ۳۲، ۸۴، ۱۶۷، ۲۱۵، ۲۲۰، ۲۲۱
 ۲۲۲، ۲۲۳، ۲۲۴، ۲۲۵، ۲۲۶، ۲۲۷، ۲۷۹
 کنترل کیفیت فرآیندهای آزمایشگاهی، ۲۲۳
 کور کردن، ۶۸، ۷۲، ۷۵، ۹۰، ۱۰۶، ۲۰۹، ۲۲۲، ۲۲۴
 کوه یخ بیماری، ۵۱
 کیفیت داده ها، ۴۷، ۶۰، ۱۶۷، ۱۶۸، ۲۲۵، ۲۲۶
 گ
 گروه شاهد بیمارستانی، ۵۴، ۵۵، ۵۶، ۵۷
 گروه های شاهد چندگانه، ۵۴، ۵۶
 گزارش استاندارد کارآزمایی های بالینی، ۷۴
 گزارش علمی اندازه ها، ۹۵
 گزارش مورد، ۴۵
 گم شدن اعضای، ۲۰، ۲۶۲
 گم شده، ۲۳، ۸۳، ۸۷، ۱۶۸، ۲۰۰، ۲۱۰، ۲۱۷، ۲۲۰، ۲۲۳، ۲۲۵، ۲۲۶، ۲۶۲
 م
 ماهیت پژوهش، ۱۸، ۳۱
 متآنالیز، ۳۹، ۹۰، ۹۴
 متغیر پاسخ، ۱۹، ۲۵، ۲۶، ۲۷، ۲۸، ۳۰، ۴۸، ۶۹، ۱۰۵، ۱۱۰، ۱۳۳، ۱۳۶، ۱۶۸، ۱۷۱، ۱۸۶، ۱۸۹، ۲۰۵
 متغیر مخدوش کننده، ۲۶، ۱۳۶، ۱۴۲، ۱۴۳، ۱۵۴، ۱۶۰، ۱۶۴، ۲۶۲
 متغیر مستقل، ۲۵، ۲۶، ۲۷، ۲۸، ۱۷۷، ۲۲۰
 متغیرهای اسمی، ۲۵، ۱۰۷، ۱۲۰، ۱۲۳، ۲۷۱
 متغیرهای رتبه ای، ۲۵
 متغیرهای کمی، ۲۵

مدل رگرسیون، ۱۲۳، ۱۶۵، ۱۷۷، ۱۷۹، ۱۸۱	متغیرهای کیفی، ۲۵، ۲۴۴
مدل رگرسیون خطی، ۱۲۳	متغیرهای محدودش‌کننده، ۴۶، ۶۷، ۶۹، ۸۱، ۱۳۵
مدل‌سازی، ۱۳۶	۱۳۶، ۱۵۳، ۱۵۴، ۱۵۹، ۱۶۴، ۱۶۷، ۱۶۸، ۱۸۶
مراحل تدوین پرسشنامه، ۲۴۵	۱۸۷، ۱۸۸، ۱۹۲، ۲۶۲، ۲۷۲
مرور منظم، ۳۹	متغیرهای نسبتی، ۲۵
مرور موارد، ۳۹، ۴۵	متناسب بودن مخاطره، ۱۸۵
مشاهده، ۱۵، ۱۷، ۲۵، ۲۶، ۳۰، ۳۱، ۴۱، ۴۲، ۴۷	مثبت کاذب، ۱۱۱، ۱۱۲، ۱۱۷، ۲۶۹، ۲۷۰
۴۸، ۵۷، ۶۳، ۶۴، ۶۷، ۷۰، ۸۱، ۹۶، ۹۷	مثبت واقعی، ۱۱۰، ۱۱۱، ۱۱۷
۱۰۰، ۱۰۴، ۱۰۷، ۱۰۸، ۱۱۰، ۱۱۱، ۱۲۳، ۱۲۴	محاسبه حجم نمونه، ۲۸، ۳۰، ۷۵، ۲۵۵، ۲۵۶
۱۲۶، ۱۳۱، ۱۳۲، ۱۳۴، ۱۴۴، ۱۴۷، ۱۴۸، ۱۵۰	۲۵۷، ۲۵۹، ۲۶۱، ۲۶۲، ۲۶۴، ۲۶۵، ۲۶۶، ۲۷۰
۱۵۶، ۱۶۲، ۱۶۴، ۱۹۳، ۱۹۹، ۲۰۰، ۲۰۱، ۲۰۲	۲۷۱، ۲۷۲
۲۰۳، ۲۰۵، ۲۰۷، ۲۲۲، ۲۲۳، ۲۲۴، ۲۳۷، ۲۵۲	محدود کردن، ۲۱، ۲۸، ۶۸، ۶۹، ۱۳۵، ۱۳۶، ۱۵۳
۲۵۵، ۲۵۶، ۲۶۵، ۲۶۹	۱۵۴
مشاهده‌گر، ۵۷، ۶۹، ۹۸، ۹۹، ۱۰۰، ۱۰۳، ۱۰۵	محدودسازی، ۱۳۵
۱۰۶، ۱۰۷، ۱۰۱، ۲۰۱، ۲۰۲، ۲۳۷	مخاطره، ۱۰۵، ۱۳۸، ۱۷۷، ۱۸۳، ۱۸۵، ۱۸۶
مشخصات جمعیت‌شناختی، ۲۳، ۲۴	محدوش شدن، ۱۷، ۱۸، ۸۱، ۱۳۱، ۱۳۴، ۱۳۵
مصاحبه، ۲۳، ۲۴، ۴۱، ۵۷، ۶۹، ۹۷، ۹۹، ۲۰۰	۱۳۷، ۱۴۴، ۱۴۸، ۱۵۳، ۱۵۴، ۱۵۹
۲۰۱، ۲۱۶، ۲۱۸، ۲۳۷، ۲۳۸، ۲۴۰، ۲۴۲، ۲۴۵	محدوش‌کنندگی، ۱۸، ۱۳۵، ۱۳۶، ۱۴۱، ۱۴۳
۲۴۶، ۲۴۸، ۲۴۹، ۲۵۰، ۲۶۷	۱۴۴، ۱۴۷، ۱۴۸، ۱۴۹، ۱۵۳، ۱۵۵، ۱۵۷
مصاحبه منظم، ۲۴۹، ۲۵۰	محدوش‌کننده، ۱۸، ۲۶، ۴۲، ۴۶، ۵۴، ۵۶، ۶۴، ۶۵
مصاحبه نامنظم، ۲۴۹، ۲۵۰	۶۷، ۶۹، ۷۵، ۸۰، ۸۱، ۸۸، ۱۰۴، ۱۳۲، ۱۳۵
مصاحبه نیمه منظم، ۲۴۹، ۲۵۰	۱۳۶، ۱۳۷، ۱۴۱، ۱۴۲، ۱۴۳، ۱۴۴، ۱۴۵، ۱۴۶
مطالعات KAP، ۴۱	۱۴۷، ۱۴۸، ۱۴۹، ۱۵۳، ۱۵۴، ۱۵۵، ۱۵۶، ۱۵۹
مطالعات بدون شاهد، ۸۵	۱۶۰، ۱۶۲، ۱۶۳، ۱۶۴، ۱۶۵، ۱۶۷، ۱۶۸، ۱۷۶
مطالعات بر مبنای داده‌های موجود، ۸۷	۱۷۹، ۱۸۶، ۱۸۷، ۱۸۸، ۱۹۲، ۱۹۷، ۲۶۲، ۲۷۲
مطالعات تاریخی، ۴۰، ۶۰	محدوش‌کننده‌ها، ۷۵، ۱۲۳، ۱۳۵، ۱۵۴، ۱۶۵
مطالعات تجربی، ۲۵، ۴۲، ۴۳، ۶۶، ۶۷، ۸۶، ۱۳۵	۱۶۸، ۲۴۶، ۲۷۲
۱۴۷	محدوش‌کننده‌های پایه، ۸۰
مطالعات تحلیلی، ۲۶، ۳۸، ۴۰، ۴۱، ۴۲، ۱۳۲، ۱۵۶	مداخله، ۲۵، ۳۸، ۴۳، ۵۸، ۶۶، ۶۷، ۶۸، ۶۹، ۷۰
۲۵۷، ۲۷۰	۷۵، ۷۹، ۸۱، ۸۲، ۸۵، ۸۶، ۸۷، ۹۰، ۱۰۴، ۲۰۷
مطالعات تصادفی، ۸۷، ۱۴۷	۲۰۹، ۲۱۱، ۲۱۸، ۲۱۹، ۲۲۰، ۲۲۷، ۲۳۲
مطالعات توصیفی، ۲۶، ۳۸، ۳۹، ۴۰، ۴۵، ۱۰۵	مداخله‌های اجتماعی، ۶۷، ۸۶

- ۲۵۷، ۲۶۴، ۲۷۱ مطالعه مقطعی، ۲۲، ۴۸، ۴۹، ۱۹۹، ۲۰۶
 مطالعات چندمرکزی، ۲۲۶
 مطالعات طولی، ۴۰، ۶۸، ۱۸۳، ۱۸۹
 مطالعات غربالگری، ۲۰۴، ۲۰۵
 مطالعات فامیلی، ۲۶۳
 مطالعات فرعی، ۸۷، ۸۹، ۹۰
 مطالعات مداخله‌ای، ۳۸، ۳۹، ۴۲، ۶۶، ۱۵۳
 مطالعات مشاهده‌ای، ۲۲، ۲۵، ۲۶، ۳۸، ۳۹، ۴۲
 ۴۹، ۶۷، ۱۰۲، ۱۶۷، ۱۶۸، ۱۸۹
 مطالعات مقطعی، ۳۹، ۴۸، ۴۹، ۵۱، ۸۹، ۱۰۵
 ۱۰۶، ۱۱۰، ۱۳۴، ۲۰۵، ۲۰۶
 مطالعات مورد-شاهدی، ۳۹، ۴۲، ۴۹، ۵۰، ۵۱، ۵۴
 ۸۹، ۱۱۰، ۱۳۷، ۱۴۷، ۱۵۴، ۱۵۵، ۱۶۴، ۱۹۲، ۱۹۵، ۱۹۶، ۱۹۸، ۱۹۹، ۲۰۰، ۲۰۱، ۲۰۲
 ۲۰۴، ۲۰۳
 مطالعات همبستگی، ۴۰، ۴۶، ۴۷
 مطالعات هم‌گروهی، ۲۶، ۳۹، ۴۲، ۴۹، ۵۰، ۵۷، ۵۸
 ۵۹، ۶۰، ۶۱، ۶۵، ۶۶، ۸۹، ۱۰۵، ۱۱۰، ۱۳۴، ۱۳۷، ۱۴۷، ۱۹۶، ۱۹۹، ۲۰۱، ۲۰۳
 مطالعات هم‌گروهی آینده‌نگر، ۵۷، ۵۸، ۵۹، ۸۹، ۱۱۰، ۱۲۷
 مطالعات هم‌گروهی گذشته‌نگر، ۶۱
 مطالعات هنجاریابی، ۴۰
 مطالعه KAP، ۳۹، ۴۱
 مطالعه با استفاده از داده‌های موجود، ۴۱
 مطالعه تجربی، ۶۷، ۸۶، ۹۸، ۲۳۲
 مطالعه راهنما، ۲۰، ۲۱۵، ۲۱۶، ۲۱۷، ۲۱۸، ۲۱۹، ۲۲۰، ۲۳۳، ۲۳۴، ۲۵۶، ۲۶۴، ۲۷۰، ۲۷۱، ۲۷۲
 مطالعه روش‌های مناسب برای تحلیل داده‌های مورد-شاهدی، ۱۹۲
 مطالعه فاکتوریل، ۷۹
 مطالعه فرعی، ۳۹، ۴۱، ۸۹، ۹۰
- مطالعه مورد، ۲۴، ۴۵، ۲۵۵
 مطالعه مورد-شاهدی، ۲۲، ۴۹، ۵۰، ۵۱، ۵۳، ۵۴، ۵۵، ۶۳، ۱۰۸، ۱۱۰، ۱۵۴، ۱۵۷، ۱۹۵، ۱۹۶، ۱۹۷، ۱۹۸، ۲۰۰، ۲۰۱، ۲۰۳، ۲۰۵، ۲۶۱
 مطالعه نیمه تجربی، ۶۷، ۸۴
 مطالعه هم‌گروهی، ۲۲، ۳۹، ۵۸، ۶۰، ۶۱، ۶۲، ۶۳، ۶۶، ۱۴۷، ۲۰۱، ۲۰۲
 مطالعه هم‌گروهی گذشته‌نگر، ۶۰
 مطالعه‌های تحلیلی، ۳۸، ۱۵۳
 مطالعه‌های توصیفی، ۳۸
 معتبرسازی، ۲۴۷
 معلول-علت، ۱۳۱، ۱۳۴
 معیارهای خروج، ۲۱، ۲۲، ۲۳، ۲۴، ۶۹، ۲۲۷
 معیارهای ورود، ۲۱، ۲۲، ۲۴، ۶۹، ۱۵۳
 معیارهای ورود و خروج، ۵۲، ۵۴، ۶۸، ۹۰، ۲۲۲، ۲۲۷
 مقدار احتمال، ۳۰، ۳۳، ۱۳۷، ۱۳۸، ۱۷۹، ۲۵۶، ۲۶۹، ۲۷۳
 مقیاس عددی، ۲۵
 ملاحظات اخلاقی، ۴۵، ۶۷
 منابع تورش، ۳۳، ۵۰، ۷۶، ۱۳۴، ۱۹۹
 منبع بالقوه تورش، ۱۳۳
 منحنی‌های راک، ۱۱۷، ۱۱۹
 منفی کاذب، ۱۱۰، ۱۱۱، ۱۱۲
 منفی واقعی، ۱۱۱
 مهار اثر شانس، ۱۳۲
 مهار اثر مخدوش‌کنندگی، ۱۵۳
 موارد گم شده، ۸۳، ۸۷، ۱۶۸، ۲۰۰، ۲۱۷، ۲۶۲
 مورد-شاهدی، ۱۷، ۲۲، ۲۶، ۳۹، ۴۲، ۴۹، ۵۰، ۵۱، ۵۳، ۵۴، ۵۵، ۵۷، ۶۳، ۸۹، ۱۰۵، ۱۰۸، ۱۱۰، ۱۳۴، ۱۳۷، ۱۴۷، ۱۵۴، ۱۵۵، ۱۵۶، ۱۵۷، ۱۶۴

نمونه‌گیری خوشه‌ای، ۲۵، ۲۱۸، ۲۶۳، ۲۶۴، ۲۷۱	۱۹۲، ۱۹۵، ۲۰۰، ۲۰۱، ۲۰۲، ۲۰۴، ۲۰۵، ۲۶۱
نمونه‌گیری سهمیه‌ای، ۲۵	۲۶۲، ۲۶۶
نمونه‌گیری طبقه‌ای، ۲۵	موردها، ۱۸، ۲۲، ۴۲، ۴۹، ۵۰، ۵۱، ۵۲، ۵۳، ۵۴
نمونه‌گیری غیر احتمالاتی، ۲۴	۵۵، ۵۶، ۵۷، ۶۲، ۱۳۳، ۱۳۴، ۱۳۵، ۱۵۴، ۱۵۵
نمونه‌گیری مبتنی بر قضاوت، ۲۵	۱۵۶، ۱۶۳، ۱۶۴، ۱۹۲، ۱۹۶، ۱۹۷، ۱۹۸، ۱۹۹
نمونه‌گیری منظم، ۲۵	۲۰۰، ۲۰۱، ۲۰۴، ۲۶۷
نهایی کردن پروتکل، ۲۱۹	میانگین، ۱۸، ۲۸، ۳۰، ۴۰، ۴۶، ۹۷، ۹۹، ۱۰۰
و	۱۰۱، ۱۰۵، ۱۰۸، ۱۱۷، ۱۲۲، ۱۲۳، ۱۲۴، ۱۲۵
واریانس، ۹۰، ۹۵، ۱۰۸، ۱۰۹، ۱۲۳، ۱۲۴، ۱۶۸	۱۵۷، ۱۵۹، ۱۶۸، ۱۷۰، ۱۷۱، ۱۷۷، ۱۸۶، ۱۹۵
۱۶۹، ۱۸۶، ۱۸۷، ۲۶۳، ۲۶۶	۱۹۶، ۱۹۹، ۲۲۳، ۲۴۳، ۲۴۵، ۲۵۷، ۲۶۳، ۲۶۴
ویژگی، ۱۶، ۲۰، ۲۱، ۲۲، ۲۴، ۲۶، ۳۳، ۴۰، ۴۵	۲۶۵، ۲۶۶، ۲۶۷، ۲۷۱، ۲۷۲
۴۶، ۴۸، ۵۲، ۵۸، ۵۹، ۶۳، ۷۹، ۸۲، ۸۸، ۹۸	میانگین وزنی، ۱۶۰
۱۰۰، ۱۰۱، ۱۰۲، ۱۰۴، ۱۰۵، ۱۰۶، ۱۱۱، ۱۱۲	ن
۱۱۳، ۱۱۶، ۱۱۷، ۱۱۹، ۱۲۰، ۱۴۱، ۱۴۶، ۱۴۸	میزان، ۲۳، ۴۸، ۵۸، ۶۹، ۱۰۵، ۱۴۸، ۱۴۹، ۱۵۹
۱۷۰، ۱۸۹، ۲۰۴، ۲۰۵، ۲۱۸، ۲۲۵، ۲۳۴	۱۶۰، ۱۶۱، ۱۶۲، ۱۶۳، ۱۷۷، ۱۷۹، ۱۸۰، ۱۸۱
ه	۱۸۳، ۱۸۸، ۲۰۴، ۲۰۶، ۲۱۵، ۲۴۹، ۲۵۷، ۲۶۵
هدایت مطالعه، ۸۶، ۸۹	۲۷۱
همسان کردن، ۵۴، ۵۶، ۱۳۶، ۱۵۳، ۱۵۵	میزان بروز استاندارد، ۱۶۲
همسان‌سازی، ۵۶، ۱۳۵، ۱۳۶، ۱۵۴، ۱۵۵، ۱۵۶	میزان پاسخگویی، ۲۴۸
۱۹۲	میزان توافق، ۱۰۷
همسان‌سازی فردی، ۵۶، ۱۵۴، ۱۹۲	نسبت درست‌نمایی، ۱۰۵، ۱۲۰، ۱۲۱
همسان‌سازی گروهی، ۵۶، ۱۵۴	نسبت درست‌نمایی آزمون، ۱۲۰
هم‌گروهی، ۲۲، ۲۶، ۳۹، ۴۲، ۴۹، ۵۰، ۵۱، ۵۷	نسبت شانس، ۴۹، ۵۰، ۱۰۵، ۱۴۳، ۱۵۹، ۱۶۳
۵۸، ۵۹، ۶۰، ۶۱، ۶۲، ۶۳، ۶۴، ۶۵، ۶۶، ۸۹	۱۶۴، ۱۷۲، ۱۷۳، ۱۷۴، ۱۸۱، ۱۸۲، ۱۸۷، ۱۹۲
۱۰۵، ۱۱۰، ۱۳۷، ۱۴۷، ۱۹۵، ۱۹۶، ۱۹۹، ۲۰۱	۱۹۷، ۱۹۸، ۱۹۹، ۲۰۳، ۲۰۴، ۲۵۶
۲۰۲، ۲۰۳	نسبت شانس تعدیل شده، ۱۵۷، ۱۶۲، ۱۶۳
هم‌گروهی گذشته‌نگر، ۵۸، ۶۰، ۶۱، ۱۳۷	نسبت مرگ‌ومیر استاندارد، ۱۶۲
	نلسون - آلن، ۱۸۳
	نمودار بلاند و آلتمن، ۱۲۲، ۱۲۳
	نمونه‌گیری احتمالاتی، ۲۴، ۲۵
	نمونه‌گیری با احتمال متغیر، ۲۵
	نمونه‌گیری تصادفی، ۲۵، ۴۹

واژه‌نامه فارسی به انگلیسی

Over matching	بیش همسان‌سازی		
Non response	بی‌پاسخی		
		آ	
		Diagnostic test	آزمون تشخیصی
		t-student test	آزمون t-استیودنت
		Log-rank test	آزمون لگ-رتبه
		Descriptive statistics	آمار توصیفی
		Carry-over effect	اثر انتقالی
		Design effect	اثر طرح
		Interaction	اثر متقابل
		Positive predictive value	ارزش اخباری آزمون مثبت
		Negative predictive value	ارزش اخباری آزمون منفی
		predictive value	ارزش اخباری
		Indirect standardization	استاندارد کردن غیرمستقیم
		Direct standardization	استاندارد کردن مستقیم
		Mistake	اشتباه
		Proposal revision	اصلاح پروپوزال
		Validity	اعتبار
		Experimental validity	اعتبار تجربی
		External validity	اعتبار خارجی
		Internal validity	اعتبار داخلی
		Structural validity	اعتبار ساختاری
		Face validity	اعتبار ظاهری
		Content validity	اعتبار محتوا
		Study subjects	اعضای موردبررسی
		Standard deviation	انحراف معیار
		Effect size	اندازه اثر
		Global measure	اندازه خلاصه
		Environmental measure	اندازه محیطی
		Aggregate measure	اندازه کلان
			ب
		Estimation	برآورد
		Risk estimation	برآورد خطر
		Labeling	برچسب‌گذاری
		Incidence	بروز
			بلوک‌های جایگشتی تصادفی
		Random permutation blocks	
		Protocol revision	بهبود پروتکل
	پ		
Questionnaire	پرسشنامه		
Proposal	پروپوزال/پیشنهادیه		
Protocol	پروتکل		
Pre-test	پیش آزمون		
Follow up	پیگیری		
		ت	
Survival function	تابع بقا		
Multivariate analysis	تحلیل چند متغیره		
Data analysis	تحلیل داده‌ها		
Secondary data analysis	تحلیل مجدد داده‌ها		
Random	تصادف		
Randomization	تصادفی کردن/تصادفی سازی تغییرات ذاتی اعضای تحت مطالعه		
Within subject variability			
Power of a test	توان آزمون		
Biological plausibility	توجیه بیولوژیکی		
Instrument bias	تورش ابزار اندازه‌گیری تورش استفاده از بروز به جای شیوع		
Incidence-prevalence bias			
Information bias	تورش اطلاع		
Subject bias	تورش اعضای نمونه		
Selection bias	تورش انتخاب		
Ecologic bias	تورش اکولوژیک		
Berksonian bias	تورش برکسون		
Respondent bias	تورش پاسخگویی		
Surveillance bias	تورش پایش		
Temporal bias	تورش تقدم پیامد بر مواجهه		
Recall bias	تورش خاطره		
Observer bias	تورش مشاهده‌گر		
Interviewer bias	تورش مصاحبه‌گر تورش مکمل شیوع		
Point prevalence complement ratio bias			
Duration ratio bias	تورش نسبت دوره بیماری		

	ذ	Publication bias	تورش نشر
	ر	Bias	تورش / سوگرایی
Effect-cause relation	رابطه اثر-علت	Normal distribution	توزیع نرمال
Spurious relation	رابطه تصنعی	Standard normal distribution	توزیع نرمال استاندارد
Real relation	رابطه حقیقی	Reproducibility	تکرارپذیری
Causal relation	رابطه علت-معلول		
Cause-effect relation	رابطه علت-معلول	ث	
Effect-cause relation	رابطه معلول-علت	Consistency	ثبات
Casual relation	رابطه سببی		
Group sequential method	روش دنباله‌ای گروهی	ج	
Gold standard method	روش استاندارد مرجع	Cross-tabulation	جدول توافقی
Biased coin method	روش سکه اریب	Life table	جدول عمر
Mantel-Haenszel method	روش مانتل - هانزل	چ	
Mc Nemar's method	روش مک - نمار	ح	
Nelson-Aalen method	روش نلسون - آلن	Sample size	حجم نمونه
Sample technique	روش نمونه‌گیری	Sensitivity	حساسیت
Kaplan-Meier method	روش کاپلان - مایر		
	ز		
	س	خ	
Consistency	سازگاری	Error	خطا
Censoring	سانسور شدن	Measurement error	خطای اندازه‌گیری
Misclassification	سوء طبقه‌بندی	Random error	خطای تصادفی
	سوء طبقه‌بندی بدون جهت	Systematic error	خطای منظم
Non differential misclassification		Sampling error	خطای نمونه‌گیری
Differential misclassification	سوء طبقه‌بندی جهت‌دار	Type I error	خطای نوع اول
Open question	سؤال باز	Type II error	خطای نوع دوم
Close question	سؤال بسته	Risk	خطر
Compensating bias	سوگرایی جبران شده	Attributable risk	خطر قابل انتساب
	ش	Adjusted attributable risk	خطر قابل انتساب تعدیل‌یافته
		Relative risk	خطر نسبی
		Adjusted relative risk	خطر نسبی تعدیل‌یافته
		د	
Odds	شانس	Data	داده
Control	شاهد	Outlier	داده پرت
Concurrent parallel control	شاهد هم‌زمان موازی	Outlier	داده دورافتاده
Prevalence	شیوع	Missing data	داده گم‌شده
	ص	Operation research manual	دستورالعمل عملیاتی
Accuracy	صحت	Precision	دقت
	ض	Washout period	دوره شستشو
		Dose-response	دوز - پاسخ

ل		Correlation coefficient	ضریب همبستگی
م		Kappa coefficient	ضریب کاپا
		Regression coefficient	ضریب رگرسیونی
Meta analysis	متا آنالیز		
Nominal variable	متغیر اسمی	ط	
Response variable	متغیر پاسخ		
Ordinal variable	متغیر رتبه‌ای	Stratification	طبقه‌بندی
Numerical variable	متغیر عددی	Latin square design	طرح مربع لاتین
Independent variable	متغیر مستقل	Study design	طرح مطالعه
Ratio variable	متغیر نسبی	Parallel design	طرح موازی
Quantitative variable	متغیر کمی	ظ	
Qualitative variable	متغیر کیفی	ع	
Dependent variable	متغیر وابسته	غ	
True positive	مثبت واقعی	ف	
False positive	مثبت کاذب		
Sample size calculation	محاسبه حجم نمونه	Confidence interval	فاصله اطمینان
Restriction	محدودسازی	Hypothesis	فرضیه
Hazard	مخاطره	Two-sided hypothesis	فرضیه دوطرفه
Proportional hazard	مخاطره متناسب	Simple hypothesis	فرضیه ساده
Confounding	مخدوش شدن	Null hypothesis	فرضیه صفر
Confounder	مخدوش کننده	Multiple Hypothesis	فرضیه مرکب
Community intervention	مداخله اجتماعی	Alternative hypothesis	فرضیه مقابل
Intervention	مداخله	One-sided hypothesis	فرضیه یک طرفه
Poisson regression model	مدل رگرسیون پواسن	ق	
Linear regression model	مدل رگرسیون خطی	Feasibility	قابل اجرا بودن
Logistic regression model	مدل رگرسیون لجستیک	Strength of association	قدرت همبستگی
Cox regression model	مدل رگرسیون کاکس	ک	
Regression model	مدل رگرسیون		
Modeling	مدل سازی	Clinical trial	کارآزمایی بالینی
Systematic review	مرور منظم	Randomized trial	کارآزمایی تصادفی
Case series study	مرور موارد	Cluster randomized trial	کارآزمایی تصادفی خوشه‌ای
Observer	مشاهده گر	Non-randomized clinical trial	کارآزمایی تصادفی نشده
Observation	مشاهده	Multi-center clinical trial	کارآزمایی چندمرکزی
Demographic characteristic	مشخصات جمعیت شناسی	Field trial	کارآزمایی میدانی
Interveiw	مصاحبه	Quality control	کنترل کیفیت
Structured interveiw	مصاحبه منظم	Blinding	کور کردن
Semi-structured interveiw	مصاحبه نامنظم	Quality of data	کیفیت داده‌ها
Unstructured interveiw	مصاحبه نیمه منظم	گ	
Screening study	مطالعه غربالگری		
KAP study	مطالعه KAP	Case report	گزارش مورد
Historical study	مطالعه تاریخی	Lost to follow up	گم شدن در پیگیری
Experimental study	مطالعه تجربی		

	ن		Analytical study	مطالعه تحلیلی
Rate		میزان	Descriptive study	مطالعه توصیفی
Standardized rate		میزان استاندارد	Multi-center study	مطالعه چندمرکزی
Standardized incidence rate		میزان بروز استاندارد	Pilot study	مطالعه راه‌نما
Proportion		نسبت	Longitudinal study	مطالعه طولی
Proportion of agreement		نسبت توافق	Familial study	مطالعه فامیلی
Odds ratio		نسبت شانس	Factorial study	مطالعه فاکتوریل
Adjusted odds ratio		نسبت شانس تعدیل‌یافته	Ancillary study	مطالعه فرعی
Standardized prevalence ratio		نسبت شیوع استاندارد		مطالعه مبنی بر داده‌های موجود
Plot		نمودار	Research using existing data	
Sample		نمونه	Field study	مطالعه مبنی بر روش‌های میدانی
Sampling		نمونه‌گیری	Cross over study	مطالعه متقاطع
Systematic sampling		نمونه‌گیری منظم	Interventional study	مطالعه مداخله‌ای
Probability sampling		نمونه‌گیری احتمالاتی	Observational study	مطالعه مشاهده‌ای
Simple random sampling		نمونه‌گیری تصادفی ساده	Cross-sectional study	مطالعه مقطعی
Multi-phase sampling		نمونه‌گیری چندفازی	Case study	مطالعه مورد
Cluster sampling		نمونه‌گیری خوشه‌ای	Case-control study	مطالعه مورد-شاهدی
Quota sampling		نمونه‌گیری سهمیه‌ای	Quasi-experimental study	مطالعه نیمه‌تجربی
Stratification sampling		نمونه‌گیری طبقه‌ای	Ecologic study	مطالعه همبستگی
Non-probability sampling		نمونه‌گیری غیر احتمالاتی	Prospective study	مطالعه هم‌گروهی آینده‌نگر
	و		Retrospective study	مطالعه هم‌گروهی گذشته‌نگر
Variance		واریانس	Cohort study	مطالعه هم‌گروهی
Specificity		ویژگی	Normative study	مطالعه هنجاریابی
	ه		Exclusion criteria	معیارهای خروج
Matching		همسان‌سازی	Eligibility criteria	معیارهای ورود و خروج
Individual matching		همسان‌سازی فردی	Inclusion criteria	معیارهای ورود
Frequency matching		همسان‌سازی گروهی	P-value	مقدار احتمال
	ی		Ethical considerations	ملاحظات اخلاقی
			Potential source of bias	منابع تورش بالقوه
			ROC curve	منحنی راک
			False negative	منفی کاذب
			True negative	منفی واقعی
			Case	مورد
			Mean	میانگین
			Weighted mean	میانگین وزنی

واژه‌نامه انگلیسی به فارسی

A		Concurrent parallel control	شاهد هم‌زمان موازی
Accuracy	صحت	Confidence interval	فاصله اطمینان
Adjusted attributable risk	خطر قابل انتساب تعدیل‌یافته	Confounder	مخدوش کننده
Adjusted odds ratio	نسبت شانسی تعدیل‌یافته	Confounding	مخدوش شدن
Adjusted relative risk	خطر نسبی تعدیل‌یافته	Consistency	ثبات
Aggregate measure	اندازه کلان	Consistency	سازگاری
Alternative hypothesis	فرضیه مقابل	Content validity	اعتبار محتوا
Analytical study	مطالعه تحلیلی	Control	شاهد
Ancillary study	مطالعه فرعی	Cox regression model	مدل رگرسیون کاکس
Attributable risk	خطر قابل انتساب	Cross over study	مطالعه متقاطع
		Cross-sectional study	مطالعه مقطعی
		Cross-tabulation	جدول توافقی
B		D	
Berksonian bias	تورش برکسون	Data analysis	تحلیل داده‌ها
Bias	تورش	Data	داده
Biased coin method	روش سکه اریب	Demographic characteristic	مشخصات جمعیت شناسی
Biological plausibility	توجیه بیولوژیکی	Dependent variable	متغیر وابسته
Blinding	کور کردن	Descriptive statistics	آمار توصیفی
		Descriptive study	مطالعه توصیفی
		Design effect	اثر طرح
		Diagnostic test	آزمون تشخیصی
		Differential misclassification	سو طبقه بندی جهت دار
		Direct standardization	استاندارد کردن مستقیم
		Dose-response	دوز-پاسخ
		Duration ratio bias	تورش نسبت دوره بیماری
C		E	
Carry-over effect	اثر انتقالی	Ecologic bias	تورش اکولوژیک
Case	مورد	Ecologic study	مطالعه همبستگی
Case report	گزارش مورد	Effect size	اندازه اثر
Case series study	مرور موارد	Effect-cause relation	رابطه اثر - علت
Case study	مطالعه مورد	Effect-cause relation	رابطه معلول - علت
Case-control study	مطالعه مورد-شاهدی	Eligibility criteria	معیارهای ورود و خروج
Casual relation	روابط سببی	Environmental measure	اندازه محیطی
Casual relation	رابطه علت-معلول	Error	خطا
Cause-effect relation	رابطه علت-معلول	Estimation	برآورد
Censoring	سانسور شدن	Ethical considerations	ملاحظات اخلاقی
Clinical trial	کارآزمایی بالینی		
Close question	سؤال بسته		
Cluster randomized trial	کارآزمایی تصادفی خوشه ای		
Cluster sampling	نمونه‌گیری خوشه‌ای		
Cohort study	مطالعه همگروهی		
Community intervention	مداخله اجتماعی		
Compensating bias	سوگرایی جبران شده		

Exclusion criteria	معیارهای خروج	K	
Experimental study	مطالعه تجربی	KAP study	مطالعه KAP
External validity	اعتبار خارجی	Kappa coefficient	ضریب کاپا
		Kaplan-Meier method	روش کاپلان-مایر
		Correlation coefficient	ضریب همبستگی
		L	
F		Labeling	برچسب گذاری
Face validity	اعتبار ظاهری	Latin square design	طرح مربع لاتین
Factorial study	مطالعه فاکتوریل	Life table	جدول عمر
False positive	مثبت کاذب	Likelihood ratio for a negative test	نسبت درستی منفی
True negative	منفی واقعی	Likelihood ratio for a positive test	نسبت درستی مثبت
Familial study	مطالعه فامیلی	Likelihood ratio	نسبت درستی
Feasibility	قابل اجرا بودن	Linear regression model	مدل رگرسیون خطی
Field study	مطالعه مبنی بر روش های میدانی	Logistic regression model	مدل رگرسیون لجستیک
Field trial	کارآزمایی میدانی	Log-rank test	آزمون لگ-رتبه
Follow up	پیگیری	Longitudinal study	مطالعه طولی
Frequency matching	همسان سازی گروهی	Lost to follow up	گم شدن در پیگیری
		M	
G		Mantel-Haenszel test	روش مانتل-هانزل
Global measure	اندازه خلاصه	Matching	همسان سازی
Gold standard method	روش استاندارد مرجع	Mc.Nemar's test	روش مک-نمار
		Mean	میانگین
H		Measurement error	خطای اندازه گیری
Hazard	مخاطره	Meta analysis	فرا تحلیل/متا آنالیز
Historical study	مطالعه تاریخی	Misclassification	سوء طبقه بندی
Hypothesis	فرضیه	Missing data	داده های گم شده
		Mistake	اشتباه
I		Modeling	مدل سازی
Incidence	بروز	Multi center clinical trial	کارآزمایی چندمرکزی
Incidence-prevalence bias	تورش استفاده از بروز به جای شیوع	Multi center study	مطالعه چندمرکزی
Inclusion criteria	معیارهای ورود	Multi-phase sampling	نمونه گیری چندفازی
Independent variable	متغیر مستقل	Multiple Hypotheses	فرضیه مرکب
Indirect standardization	استاندارد کردن غیرمستقیم	Multivariate analysis	تحلیل چندمتغیره
Individual matching	همسان سازی فردی	N	
Information bias	تورش اطلاع	Nelson-Alan method	روش نلسون-آلن
Instrument bias	تورش ابزار اندازه گیری	Nominal variable	متغیر اسمی
Interaction	اثر متقابل	Non-differential misclassification	
Internal validity	اعتبار داخلی		
Intervention	مداخله		
Interventional study	مطالعه مداخله ای		
Interview	مصاحبه		
Interviewer bias	تورش مصاحبه گر		
		J	

Non-probability sampling	سو طبقه بندی بدون جهت نمونه‌گیری غیر احتمالاتی	Proposal revision	اصلاح پروپوزال
Non-randomized clinical trial	کارآزمایی تصادفی نشده	Prospective study	مطالعه هم‌گروهی آینده نگر
Non-response	بی پاسخی	Protocol revision	بهبود پروتکل
Normal distribution	توزیع نرمال	Protocol	پروتکل
Normative study	مطالعه هنجاریابی	Publication bias	تورش نشر
Null hypothesis	فرضیه صفر	P-value	مقدار احتمال
Numerical variable	متغیر عددی		
O			
Observation	مشاهده	Qualitative variable	متغیر کیفی
Observational study	مطالعه مشاهده‌ای	Quality control	کنترل کیفیت
Observer bias	تورش مشاهده‌گر	Quality of data	کیفیت داده‌ها
Observer	مشاهده‌گر	Quantitative variable	متغیر کمی
Odds ratio	نسبت شانس	Quasi-experimental study	مطالعه نیمه تجربی
Odds	شانس	Questionnaire	پرسشنامه
One-sided hypothesis	فرضیه یک طرفه	Quota sample	نمونه‌گیری سهمیه‌ای
Open question	سؤال باز		
Operation research manual	دستورالعمل عملیاتی	R	
Ordinal variable	متغیر رتبه‌ای	Random error	خطای تصادفی
Outlier data	داده‌های پرت	Random permutation blocks	بلوک‌های جایگشتی تصادفی
Outlier data	داده‌های دورافتاده	Random	تصادف
Over matching	بیش همسان‌سازی	Randomization	تصادفی کردن
		Randomized trial	کارآزمایی تصادفی
P			
Parallel design	طرح موازی	Rate	میزان
Pilot study	مطالعه راهنما	Ratio variable	متغیر نسبتی
Plot	نمودار	Real relation	رابطه حقیقی
Point prevalence complement ratio bias	تورش مکمل شیوع	Recall bias	تورش خاطره
Poisson regression model	مدل رگرسیون پواسن	Regression coefficient	ضریب همبستگی
Potential source of bias	منابع تورش بالقوه	Regression model	مدل رگرسیون
Power of test	توان آزمون	Relative risk	خطر نسبی
Precision	دقت	Reproducibility	تکرارپذیری
Predictive value	ارزش اخباری	Research using existing data	مطالعه مبنی بر داده‌های موجود
Pre-test	پیش آزمون	Respondent bias	تورش پاسخگویی
Prevalence	شیوع	Response variable	متغیر پاسخ
Probability sampling	نمونه‌گیری احتمالات	Restriction	محدودسازی
Proportion	نسبت	Retrospective study	مطالعه هم‌گروهی گذشته‌نگر
Proportion of agreement	نسبت توافق	Risk	خطر
Proportional hazard	متناسب بودن مخاطر	Risk estimation	برآورد خطر
Proposal	پروپوزال	ROC curve	منحنی راک

S			
Sample	نمونه	Survival function	تابع بقا
Sample size	حجم نمونه	Systematic error	خطای منظم
Sample size calculation	محاسبه حجم نمونه	Systematic review	مرور منظم
Sample technique	روش نمونه‌گیری	Systematic sampling	نمونه‌گیری منظم
Sampling	نمونه‌گیری	T	
Sampling error	خطای نمونه‌گیری	t-student test	آزمون t-استودنت
Screening study	مطالعات غربالگری	Temporal bias	تورش تقدم پیامد بر مواجهه
Secondary data analysis	تحلیل مجدد داده‌ها	True negative	منفی واقعی
Selection bias	تورش انتخاب	True positive	مثبت واقعی
Semi structure interview	مصاحبه نیمه منظم	Two-sided hypothesis	فرضیه دوطرفه
Sensitivity	حساسیت	Type I error	خطای نوع اول
Simple hypothesis	فرضیه ساده	Type II error	خطای نوع دوم
Simple random sampling	نمونه‌گیری تصادفی ساده	U	
Specificity	ویژگی	Unstructured interview	مصاحبه نامنظم
Spurious relation	رابطه تصنعی	V	
Standard deviation	انحراف معیار	Validity	اعتبار
Standard normal distribution	توزیع نرمان استاندارد	Variance	واریانس
Standardized incidence rate	میزان بروز استاندارد	W	
Standardized prevalence ratio	نسبت شیوع استاندارد	Washout period	دوره شستشو
Standardized rate	میزان استاندارد	Weighted mean	میانگین وزنی
Stratification	طبقه‌بندی	Within subject variability	تغییرات ذاتی اعضای تحت مطالعه
Stratification sampling	نمونه‌گیری طبقه‌ای	X	
Strength of association	قدرت همبستگی	Y	
Structural validity	اعتبار ساختاری	Z	
Structure interview	مصاحبه منظم		
Study design	طرح مطالعه		
Study subjects	اعضای موردبررسی		
Subject bias	تورش اعضای نمونه		
Surveillance bias	تورش پایش		

Common Errors in Biomedical Research & their Control

Authors:

***Dr. Hamid Soori (Ph.D, Associate Professor of Epidemiology)
Masoomeh Sanagoo (MSc of Biostatistics)***

Shahid Beheshti University of Medical Sciences, Tehran- I.R. Iran

(2007)