

## تخصیص تصادفی مندلی، روشی برای استنباط روابط علیتی در مطالعات مشاهده‌ای و جایگزینی برای مطالعات کارآزمایی بالینی: یک گزارش کوتاه

### چکیده

دریافت: ۱۴۰۱/۰۷/۱۰ ویرایش: ۱۴۰۱/۰۷/۱۷ پذیرش: ۱۴۰۱/۱۰/۲۴ آنلاین: ۱۴۰۱/۱۱/۰۱

**زمینه و هدف:** روش تصادفی سازی مندلی نسل جدیدی در روش تحلیل آماری است که از تغییرات ژنتیکی، به عنوان متغیرهای ابزاری در داده‌های حاصل از مطالعات غیرتجربی به منظور ارزیابی و برآورد تاثیرات علی عوامل خطر استفاده می‌کند.

**روش بررسی:** داده‌های این پژوهش مقطعی از دی ۱۳۸۱ تا مهر ۱۳۹۹ از پایگاه PubMed جمع‌آوری شده است. روش تصادفی سازی مندلی رویکردی بر مبنای روش‌های متآنالیز است.

**یافته‌ها:** در این پژوهش به معرفی روش تصادفی سازی مندلی، طرح‌های مطالعه آن، روش‌های آماری موجود برای برآورد، محدودیت‌های این روش و بسته‌های نرم‌افزاری موجود برای اجرای آن پرداخته شد.

**نتیجه‌گیری:** هنگامی که ما به دنبال یافتن ارتباط علیتی هستیم اما امکان استفاده از کارآزمایی بالینی به عنوان روش استاندارد وجود ندارد، می‌توان از طرح تخصیص تصادفی مندلی به عنوان روشی قدرتمند در مطالعات مشاهده‌ای استفاده کرد.

**کلمات کلیدی:** علیت، ژنتیک، انواع مطالعات علوم پزشکی، تصادفی سازی مندلی، مطالعات مشاهده‌ای.

مهدی اکبرزاده<sup>۱</sup>، دانیال حبیبی<sup>۱</sup>، گودرز کلی‌فرهود<sup>۱</sup>، محمد بیدخوری<sup>۲</sup>، فریدون عزیزی<sup>۳</sup>، مریم السادات دانشپور<sup>۳\*</sup>

۱- مرکز تحقیقات علوم سلولی و ملکولی غدد، دانشگاه علوم پزشکی شهید بهشتی، تهران، ایران.

۲- گروه اپیدمیولوژی و آمار زیستی، دانشکده بهداشت، دانشگاه علوم پزشکی تهران، تهران، ایران.

۳- مرکز تحقیقات غدد درون‌ریز، دانشگاه علوم پزشکی شهید بهشتی، تهران، ایران.

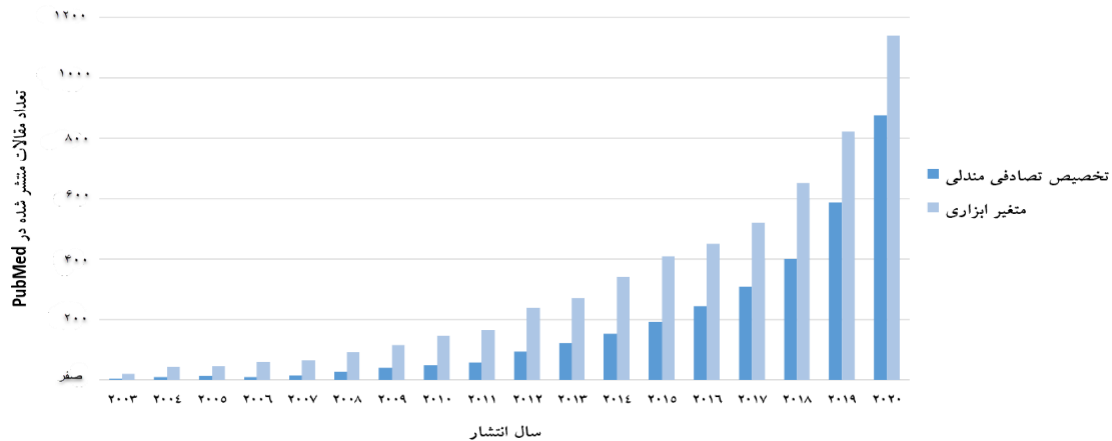
\* نویسنده مسئول: تهران، دانشگاه علوم پزشکی شهید بهشتی، مرکز تحقیقات علوم سلولی و ملکولی غدد، پژوهشگاه علوم غدد درون‌ریز. تلفن: ۰۲۱-۲۲۴۱۶۲۶۴

E-mail: daneshpour@sbmu.ac.ir

### مقدمه

بودن طرح مطالعه و مضر بودن مداخله برای آزمودنی‌ها، وجود ندارد. بنابراین، در شرایطی که کارآزمایی قابلیت انجام ندارد تنها طرح مطالعه‌ای که می‌تواند به محقق در بررسی ارتباط علیتی مواجهه و پیامد کمک کند، استفاده از مطالعات مشاهده‌ای است. مسئله‌ای که در اینجا ممکن است دقت مطالعه را کاهش دهد این است که در مطالعات مشاهده‌ای به دلیل وجود شرایطی از قبیل متغیرهای مخدوش‌کننده و روابط علی معکوس بین پیامد و عامل خطر، این مطالعات به طور مستقیم نمی‌توانند روابط علی را شناسایی کنند. اگرچه می‌توان تاثیر این عوامل مخرب را با طراحی مناسب مطالعه و همچنین در مرحله تجزیه و تحلیل داده‌ها کاهش داد، اما هنوز هم نمی‌توان اثرات آنها را به طور کامل از بین برد. به دلیل وجود چنین

یکی از اهداف اصلی در مطالعات اپیدمیولوژیک و بالینی بررسی ارتباط علیتی بین مواجهه و پیامد با توجه به حضور انواع متغیرهای مخدوشگر است. روش مرسوم برای رسیدن به این هدف استفاده از کارآزمایی بالینی به عنوان یک روش استاندارد است که می‌تواند برای بررسی ارتباط علیتی بین مواجهه و پیامد به کار گرفته شود. فرآیند تخصیص تصادفی در کارآزمایی‌های بالینی نه تنها موجب مقایسه پذیرش گروه‌ها می‌گردد، بلکه موجب می‌شود تا احتمال اربیی نیز حذف گردد. اما بعضی اوقات امکان انجام یک کارآزمایی بالینی به دلایلی از قبیل هزینه بالا، عملیاتی نبودن یا حتی غیراخلاقی



شکل ۱: روند روبه‌رشد مطالعه تصادفی‌سازی مندلی و تکنیک متغیر ابزاری

در آن هیچ مداخله‌ای از طریق محقق انجام نشده باشد.<sup>۳</sup> روش تصادفی‌سازی مندلی یا Mendelian Randomization, (MR) یک روش تحلیلی است که از تغییرات ژنتیکی به‌عنوان متغیرهای ابزاری در داده‌های حاصل از مطالعات غیرتجربی به‌منظور ارزیابی و برآورد تاثیرات علی عوامل خطر استفاده می‌کند.<sup>۴</sup>

در مطالعات اپیدمیولوژی، تصادفی‌سازی مندلی شیوه جدیدی است که از اطلاعات ژنتیکی در روش‌های اپیدمیولوژی کلاسیک استفاده می‌کند. ایده اصلی روش تصادفی‌سازی مندلی بر این است که با استفاده از روش متغیر ابزاری، علاوه بر کنترل اثر متغیرهای مخدوشگر، رابطه علی بین مواجهه و پیامد را تعیین می‌کند. به‌این‌ترتیب محدودیت‌های موجود در مطالعات مشاهده‌ای تعدیل می‌گردد. به‌عبارت دیگر می‌توان گفت که روش MR یک کاربردی از رویکرد IV است.

تصادفی‌سازی مندلی از لحاظ قدرت ارایه شواهد در بالاترین رتبه در بین مطالعات مشاهده‌ای قرار گرفته است ولی نسبت به مطالعات مداخله‌ای در سطح پایین‌تری قرار دارد (شکل ۲).<sup>۵</sup>

روش تصادفی‌سازی مندلی به‌علت تاثیرپذیری خصوصیات خود از قوانین مندلی به این نام شهرت یافته است. قوانین مدل به‌طور خلاصه به این مطلب اشاره دارد که در مرحله میوز، تقسیم‌بندی سلول‌ها با احتمال برابر و بدون تاثیرپذیری از سایر عوامل (از جمله عوامل مخدوشگر) است. تصادفی‌سازی مندلی مبتنی بر این واقعیت

محدودیت‌هایی لازم است تا رویکردی مناسب اتخاذ شود.<sup>۱</sup> اگر عوامل مخدوش‌کننده شناخته‌شده و قابل اندازه‌گیری باشند، می‌توان با در نظر گرفتن این عوامل در مدل‌های رگرسیونی اثرات آنها را کنترل کرد. اما در اغلب موارد با توجه به ماهیت متغیرهای مخدوشگر و همچنین تعداد زیاد آنها، مدل‌های رگرسیونی برآوردهای نارایب ارایه نمی‌کنند.<sup>۱</sup> در چنین شرایطی، روش متغیر ابزاری Instrumental Variable, (IV) به‌عنوان یک رویکرد آماری جایگزین برای بررسی رابطه علی بین مواجهه (مصرف نمک، چاقی و ...) و پیامد (پرفشاری خون، بیماری قلبی-عروقی و ...) ارایه شده است. روش IV می‌تواند به‌طور همزمان اثرات مخرب متغیر مخدوشگر را کنترل نماید و رابطه علیتی بین مواجهه و پیامد را نیز به‌دست آورد.<sup>۲</sup> در شکل ۱ روند روبه‌رشد استفاده از این نوع مطالعه نشان داده شده است. در دو دهه اخیر استفاده از رویکرد تصادفی‌سازی مندلی به‌شدت فزونی یافته است. لذا در این مطالعه برآنیم تا به معرفی روش تصادفی‌سازی مندلی، طرح‌های مطالعه، روش‌های آماری موجود برای برآورد، محدودیت‌ها و معرفی بسته‌های نرم‌افزاری برای اجرای آن بپردازیم.

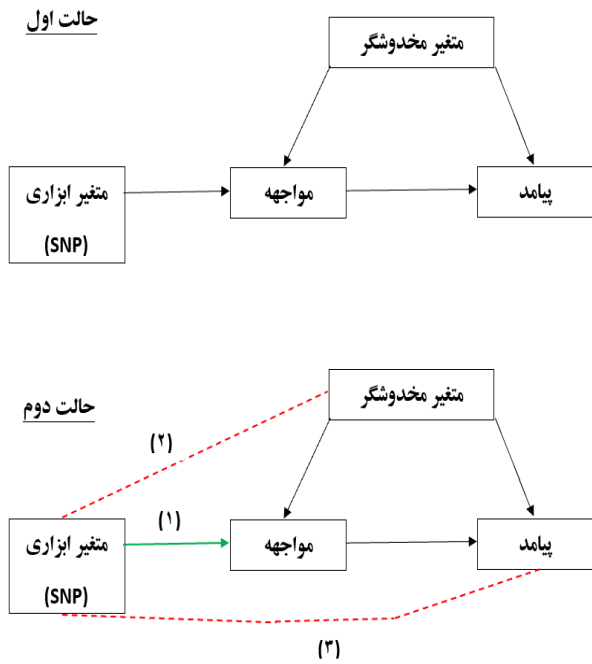
## روش بررسی

اطلاعات مطالعه حاضر در بازه زمانی سال‌های ۲۰۰۳ تا ۲۰۲۰ جمع‌آوری شده است. مطالعات غیرتجربی در واقع مطالعاتی هستند که

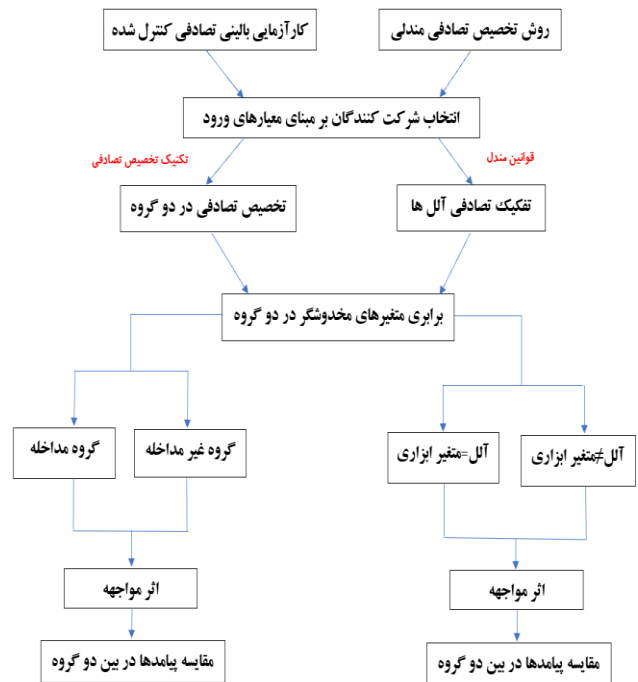
در شکل ۲، میزان شباهت بین دو طرح به دلیل وجود قوانین مندلی در روش تصادفی سازی مندلی است. این قوانین بیان می‌دارد که در واقع، ژن صرفاً بر روی ریسک فاکتور تاثیر بگذارد و ژن تاثیرگذاری بر روی پیامد از طریق ریسک فاکتور داشته باشد. هنگامی که چنین ژنی یافت شود ما تصادفی سازی خواهیم داشت. با این وجود، روش تصادفی سازی مندلی و کارآزمایی با هم تفاوت‌هایی دارند. هدف تصادفی سازی مندلی برآورد کردن اندازه یک اثر ژنتیکی نیست بلکه می‌خواهد اثر علیتی مواجهه بر روی پیامد را برآورد کند.

مفروضات مدل تخصیص تصادفی مندلی: انتخاب متغیر ابزاری مناسب برای انجام موفقیت‌آمیز یک مطالعه تصادفی سازی مندلی ضروری است (شکل ۳- حالت اول). برای یک برآورد ناریب سه فرضیه اصلی می‌بایست مدنظر قرار بگیرد، متغیر ابزاری با مواجهه (عامل خطر) مورد نظر ارتباط قوی داشته باشد، متغیر ابزاری با متغیر مخدوشگر بین مواجهه و پیامد ارتباط نداشته باشد، (شکل ۳- حالت دوم، خط نقطه‌چین ۲) و متغیر ابزاری به‌طور مستقیم با پیامد ارتباط نداشته باشد (شکل ۳- حالت دوم، خط نقطه‌چین ۳). در شکل ۳، حالت اول مفهوم اصلی مدل تصادفی سازی مندلی نشان می‌دهد.

است که تغییرات ژنتیکی، به‌طور تصادفی در طی فرآیند میوز مطابق قوانین اول و دوم وراثت مندلی (به‌ترتیب قوانین تفکیک و جورشدن تصادفی ژن‌ها) از نسلی به نسل بعد تخصیص داده می‌شوند. لذا نگرانی‌های مربوط به روابط علی معکوس را به‌جز در موارد نادر برطرف می‌نمایند.<sup>۵</sup> علاوه بر این، می‌توان به این نکته اعتماد کرد که می‌تواند فارغ از مخدوشگرهای محیطی باشد.<sup>۶</sup> در نهایت می‌توان به این نکته کلی توجه داشت که خارج از محدوده مطالعات کارآزمایی، مشخص کردن اینکه یک عامل صرفاً با پیامد در ارتباط است و یا اینکه آن عامل اثر علیتی بر روی پیامد دارد، مشکل است. بنابراین ما با استفاده از رویکرد تخصیص تصادفی مندلی به‌دنبال یافتن روابط علیتی در مطالعات مشاهده‌ای هستیم. اگر بتوان واریانت ژنتیکی‌ای یافت که مفروضات تخصیص تصادفی مندلی برای آن صدق کند، می‌توانیم یک ارتباط علیتی بین مواجهه و پیامد بدون تاثیرپذیری از نقش متغیر مخدوشگر برآورد کنیم. ژن‌های وراثتی مستقل از عوامل خارجی از قبیل مخدوشگرها می‌باشد. این شباهت زمانی به‌طور کامل خواهد بود که ژن، ریسک فاکتور و پیامد در یک راستا، مشابه همان رفتار کارآزمایی را داشته باشند.



شکل ۳: بررسی فرضیات طرح تخصیص تصادفی مندلی



شکل ۲: مقایسه طرح کارآزمایی بالینی و تخصیص تصادفی سازی مندلی

درحالی‌که در داده‌های فردی، داده‌های هر فرد قابل دسترسی است. داده‌های دیگری هم تحت عنوان متآنالیز داده‌های فردی، متآنالیز خلاصه داده و متآنالیز براساس داده‌های حاصل از مطالعات دیگر است. طرح‌های متفاوتی برای اجرای یک مطالعه تصادفی‌سازی مندلی وجود دارد (جدول ۱).

محدودیت‌های روش تصادفی‌سازی مندلی: اگرچه روش تصادفی‌سازی مندلی کاربرد گسترده‌ای در حوزه‌های مختلف علوم دارد، اما این رویکرد حتی با برقرار بودن مفروضات اصلی متغیر ابزاری بازهم محدودیت‌هایی دارد. در جدول ۲ به صورت خلاصه این محدودیت‌ها تعریف شده و راه‌کارهایی برای غلبه بر آن ارائه گردیده است.

روش‌های برآورد اثرات علی در مدل تصادفی‌سازی مندلی در جدول ۳ به نمایش آمده است.

در این روش از رویکرد از متغیر ابزاری استفاده می‌کند. این متغیر ابزاری می‌بایست دارای سه فرض باشد که در حالت دوم به نمایش درآمده است. فرض اول بیان می‌کند که متغیر ابزاری با مواجهه ارتباط قوی داشته باشد. فرض دوم نشان می‌دهد که متغیر ابزاری مستقل از متغیر مخدوشگر می‌باشد. فرض سوم به این مطلب اشاره می‌کند که متغیر ابزاری می‌بایست تنها از طریق مواجهه با پیامد در ارتباط باشد.

## یافته‌ها

داده‌های مورد استفاده در تصادفی‌سازی مندلی مشتمل بر داده‌های خلاصه‌شده، سطح فردی و متآنالیز است. داده‌های خلاصه‌شده به داده‌های حاصل از بررسی ارتباطات میان متغیر ابزاری-مواجهه و متغیر ابزاری-پیامد، یعنی ضرایب  $\beta$  و خطاهای استاندارد اطلاق می‌گردد.

جدول ۱: طرح‌های مطالعات تصادفی‌سازی مندلی

تعریف	نوع طرح مطالعه
در این طرح داده‌های مواجهه و پیامد از یک نمونه به دست می‌آیند. به عنوان مثال، مطالعه ژن (ALDH2)، مواجهه (مصرف الکل) و پیامد (فشارخون) از نمونه‌های یکسان استخراج شده‌اند.	طرح مطالعه تک نمونه‌ای
در این طرح اطلاعات مواجهه و پیامد از نمونه‌های مجزا جمع‌آوری می‌شوند. برای مثال، داده‌های مربوط به ژن ALDH2 و مصرف الکل از یک نمونه جمع‌آوری می‌شوند و داده‌های ژن ALDH2 و فشارخون از نمونه دیگر اندازه‌گیری می‌شود.	طرح مطالعه دو نمونه‌ای
در این طرح داده‌های مواجهه مربوط به زیرمجموعه‌ای از شرکت‌کنندگان در دسترس می‌باشد و در مقابل داده‌های پیامد کل شرکت‌کنندگان را می‌توان جمع‌آوری نمود. همچنین، حالت دیگر طرح زیرنمونه استخراج داده‌های پیامد از زیرمجموعه‌ای از شرکت‌کنندگان و داده‌های مواجهه از کل شرکت‌کنندگان می‌باشد.	طرح زیرنمونه
در مواردی که ابهام در زمینه جهت رابطه بین مواجهه و پیامد وجود دارد (برای مثال بررسی رابطه دیابت با کلسترول) می‌توان از طرح دو طرفه استفاده نمود. روش اجرا به این صورت است که ابتدا بررسی می‌شود که آیا مواجهه سبب پیامد شده است و به دنبال آن به بررسی پیامد به سمت مواجهه پرداخته می‌شود.	طرح دو طرفه
در این طرح هدف ارزیابی نقش یک صفت به عنوان متغیر میانجی بین مواجهه و پیامد می‌باشد. پیش‌فرض انجام چنین طرحی وجود یک ارتباط خطی و همگن بین مواجهه-میانجی و مواجهه-پیامد می‌باشد و همچنین نیازی به اثر متقابل بین مواجهه و میانجی نیست. در مرحله اول، ارتباط میان مواجهه و میانجی و در مرحله دوم ارتباط بین میانجی و پیامد محاسبه می‌شود. اگر شواهدی دال بر وجود ارتباط در هر دو مرحله وجود داشته باشد می‌توان گفت نقش متغیر میانجی وجود دارد.	طرح دو مرحله‌ای
این نوع طرح در واقع بسط طرح تک‌متغیره روش تصادفی‌سازی مندلی است. این طرح اجازه می‌دهد تا چندین مواجهه به‌طور همزمان مورد بررسی قرار بگیرند و از طرف دیگر می‌تواند به مسئله همبستگی مواجهه‌ها در شرایط پلی‌تروپیک بپردازد. یک تفاوت کلی میان این روش و روش تصادفی‌سازی مندلی وجود دارد. در روش تصادفی‌سازی مندلی اثر کلی مواجهه و پیامد را می‌توان به دست آورد این درحالی است که در روش چندمتغیره اثر هر یک از مواجهه‌ها بر روی پیامد را می‌توان مشاهده کرد.	طرح چندمتغیره
از این طرح می‌توان برای پاسخ به سؤالاتی در مورد اثر متقابل استفاده کرد. در واقع این طرح مشابه طرح فاکتوریل کارآزمایی بالینی است.	طرح فاکتوریل

جدول ۲: محدودیت‌های روش تصادفی‌سازی مندلی

محدودیت	راه حل پیشنهادی
در دسترس نبودن متغیر ابزاری ژنتیکی	مطالعات ارتباط گسترده ژنومی (GWAS), Genome-wide association study, می‌تواند تعداد واریانت‌های ژنتیکی را افزایش دهد. در مواردی که واریانت ژنتیکی مرتبط با یک مواجهه شناسایی نشده است، می‌توان تصادفی‌سازی مندلی را بر روی واریانت‌های ژنتیکی مرتبط با مواجهه مشابه در خلاصه داده‌های GWAS انجام داد. به‌عنوان مثال از واریانت‌های ژنتیکی مرتبط با شاخص توده بدنی (Body mass index, BMI) به‌جای چاقی استفاده نمود. همچنین می‌توان از رویکرد نمره خطر پلی‌ژنیک استفاده کرد، البته نتایج حاصل از این رویکرد به‌دلیل فقدان ویژگی خاصیت آنها و پلی‌تروپی باید با احتیاط تفسیر شود.
متغیر ابزاری ضعیف	افزایش حجم نمونه با استفاده از خلاصه داده‌های در دسترس GWAS، استفاده از نمرات آلی و روش‌های برآورد نیرومند مثل درست‌نمایی محدود شده، استفاده از برآورد Continuous updating estimator, (CUE) هنگامی که تعداد ابزارهای ضعیف زیاد باشد، وارد کردن کووریت‌ها در مدل، خارج کردن ابزارهایی که موجب تضعیف ارتباط واریانت ژنتیکی با مواجهه می‌گردد، استفاده از روش متآنالیز در روش ایگر اگر آماره $I^2$ (شاخص ناهمگونی) برابر یک باشد نشان می‌دهد ابزار ضعیف وجود ندارد، استفاده از نمرات پلی‌ژنیک، استفاده از طرح دو نمونه‌ای.
توان پایین	افزایش حجم نمونه و ترکیب کردن واریانت‌های ژنتیکی در درون نمرات آلی
پلی‌تروپی افقی	استفاده از نظر افراد خبره جهت مشخص نمودن روابط حاکم بر بیولوژیک واریانت ژنتیکی، استفاده از روش‌هایی مثل رگرسیون ایگر و رویکرد میانه وزنی.
طبقه‌بندی جمعیت	استفاده کردن از جوامع همگن برای انتخاب ژنتیکی واریانت، استفاده کردن از روش مولفه‌های اصلی اصلی استفاده کردن از داده‌های خانوادگی.
عدم تعادل پیوستگی	فهم بهتری از روابط حاکم بر بیولوژیک واریانت ژنتیکی داشته باشیم. از روش‌هایی مثل رگرسیون ایگر و رویکرد میانه وزنی استفاده کنیم.
کانال‌سازی کردن	روش مشخصی وجود ندارد البته داشتن دانش بیولوژی در اینجا به‌کمک خواهد آمد
پیچیدگی ارتباطات	افزایش دانش بیولوژیکی واریانت ژنتیکی، مواجهه و پیامد برای فهم بهتر روابط
اثرات دینامیکی	روش‌های درون خانوادگی با استفاده از داده‌های سطح فردی یا خلاصه‌شده ممکن است برخی از تورش‌های ناشی از اثرات دینامیکی را از بین ببرد.
اثرات دوره بحرانی	می‌توان از چندین روش اپیدمیولوژیک مختلف در دوره‌هایی که مواجهه احتمالاً منجر به پیامد می‌شود، استفاده کرد. یا استفاده از چندین روش اپیدمیولوژیک مختلف در دوره‌هایی که مواجهه احتمالاً منجر به پیامد می‌شود.
پدیده نفرین برنده	یک راه حل پیشنهادی برای این مسئله استفاده کردن از طرح دو نمونه‌ای است.
اریبی انتخاب یا اریبی بر خوردکننده	استفاده از هر پارامتر شناخته‌شده‌ای مثل شیوع بیماری و همچنین می‌توان از وزن‌دهی احتمالی معکوس استفاده کرد. روش پیشنهادی دیگر بررسی ارتباط بین واریانت ژنتیکی و پیشرفت بیماری است. هرگونه ارتباط بین واریانت مشابه و بروز بیماری باید مورد توجه قرار گیرد، زیرا ممکن است تورش برخوردارکننده را نشان دهد.

## جدول 3: روش‌های برآورد اثرات علی در مدل تصادفی سازی مندلی

تعریف	برآورد اثرات علی
این رویکرد به‌عنوان ساده‌ترین روش برای بررسی ارتباط علی مواجهه و پیامد مورد استفاده قرار می‌گیرد. به این رویکرد روش والد نیز اطلاق می‌شود. هنگامی که بیش از یک متغیر ابزاری وجود داشته باشد می‌توان به‌صورت مجزا برای هر کدام یک روش نسبت استفاده کرد و یا همه متغیرهای ابزاری را در یک متغیر ابزاری ادغام کرده و از رویکرد نمره آلی (وزنی و غیروزنی) استفاده نمود.	رویکرد نسبت
در صورتی که برای پیامد پیوسته، پیش‌فرض رابطه خطی بین مواجهه و پیامد برقرار باشد و همچنین اثر متقابل وجود نداشته باشد، به این رویکرد روش دومرحله‌ای حداقل مربعات نیز گفته می‌شود. این روش برای یک یا چند متغیر ابزاری در سطح داده فردی مورد استفاده قرار می‌گیرد. در مرحله اول، رگرسیون مواجهه بر روی متغیر ابزاری انجام می‌شود و در مرحله دوم رگرسیون پیامد بر روی مواجهه برآورد شده‌ای که در مرحله اول به‌دست آمده است (باقیمانده‌های رگرسیون مرحله اول)، برازش داده می‌شود. <sup>۷</sup> در حالت پیامد دو حالت، مرحله دوم آنالیز با استفاده از رگرسیون لجستیک یا لگ خطی انجام می‌گیرد.	رویکرد دو مرحله‌ای
روش حداکثر درست‌نمایی می‌تواند به‌طور همزمان پارامترهای مدل را برآورد کند. در عمل جهت انجام آنالیز تنها نیاز به یک پارامتر می‌باشد بنابراین این روش کارایی مناسبی نداشته و از روش درست‌نمایی محدود شده استفاده می‌شود. <sup>۸</sup> روش درست‌نمایی محدود شده می‌تواند برای یک یا چندین متغیر ابزاری و برای داده‌ها در سطح فردی مورد استفاده قرار بگیرد و در حالت تک‌متغیر ابزاری نتایج مشابهی با روش نسبت و رگرسیون دومرحله‌ای حداقل مربعات دارد.	رویکرد درست‌نمایی
رویکرد بیز به‌عنوان یک تکنیک آماری برای برآورد رابطه علیتی مورد استفاده قرار می‌گیرد. این روش نتایج مشابهی با رویکرد درست‌نمایی دارد. رویکرد بیز یک روش انعطاف‌پذیری برای استنباط از طریق Markov chain monte carlo, (MCMC) به‌منظور به‌دست آوردن توزیع پسین است. این رویکرد برای برازش مدل‌هایی با روابط غیرخطی ابزار یا اثرات غیرخطی مخدوشگر می‌تواند مورد استفاده قرار بگیرد. <sup>۹</sup>	رویکرد بیز
روش گشتاور تعمیم یافته (GMM), Generalized method of moments, و مدل‌های ساختاری میانگین (SMM), Structural mean model, و برآوردگر بروزرسانی مداوم (CUE) را می‌توان به‌عنوان روش‌های برآورد نیمه پارامتری ذکر کرد. مدل نیمه پارامتریک، متشکل از دو بخش پارامتری و ناپارامتری می‌باشد، به‌طوری‌که بخش پارامتری مربوط به ارتباط مواجهه و پیامد بوده و هیچ فرضی در مورد توزیع خطاها وجود ندارد. برآوردگر GMM هم برای رگرسیون خطی و هم غیرخطی می‌تواند مورد استفاده قرار بگیرد. زمانی که ناهمسانی در توزیع خطا و معادلات ساختاری دومرحله‌ای غیرخطی وجود داشته باشد، فرم انعطاف‌پذیرتری در مقایسه با روش رگرسیون دومرحله‌ای حداقل مربعات، Two-stage least squares, (2SLS) دارد. <sup>۱۰</sup> در مواقعی که در مطالعات مشاهده‌ای، امکان ارزیابی و یا کنترل کامل مواجهه وجود نداشته باشد (مشابه کارآزمایی‌هایی که تمکین ناقص از درمان دارند) روش SMM مورد استفاده قرار می‌گیرد. <sup>۱۱</sup> روش برآوردگر بروزرسانی مداوم (CUE) یک راه حل برای مشکلات ارزیابی در نمونه‌های با حجم نمونه محدود در روش دو مرحله‌ای GMM می‌باشد. یک نقطه‌ضعف مدل‌های نیمه‌پارامتریک این است که در هنگام برخورد با پیامد دو حالت ممکن است نتواند رابطه علی را به‌درستی برآورد نماید.	رویکرد نیمه پارامتری
علاوه بر روش نسبت، روش‌های مبتنی بر درست‌نمایی، وزن‌دهی معکوس واریانس، روش ایگر و رگرسیون میانه وزنی نیز وجود دارند. در روش وزن‌دهی معکوس واریانس، میانگین وزنی اثرات علی هر متغیر ابزاری مورد استفاده قرار می‌گیرد و معادل با روش متآنالیز برای برآورد کردن علیت است. <sup>۱۱</sup> در هنگام روبه‌رو شدن با چندین متغیر ابزاری، فارغ از روش‌های مبتنی بر درست‌نمایی و وزن‌دهی معکوس واریانس می‌توان از روش نمره‌دهی آلی (وزن داده شده، وزن داده نشده) نیز استفاده کرد. <sup>۱۲</sup> روش ایگر یک روش تجزیه و تحلیل در مورد داده‌های ژنتیکی خلاصه شده است و شامل یک تست برای بررسی پلیوتروپی افقی، یک اثر علی و برآورد اثر علی است. به‌طور کلی، روش ایگر یک تکنیک آماری است که می‌توان از آن در مواقعی که برخی از متغیرهای ابزاری معتبر نیستند و یا در صورت وجود یک فرض ضعیف مورد استفاده قرار گیرد. <sup>۱۳</sup> رگرسیون میانه وزنی نیز روشی در حضور پلیوتروپی در داده‌های خلاصه شده است. این روش برآورد علی سازگاری را حتی در صورت وجود حداکثر ۵۰٪ از متغیرهای ابزاری نامعتبر ارایه می‌دهد. <sup>۱۴</sup> در نتیجه، این روش متغیرهای ابزاری را مجاز می‌داند که می‌تواند با عوامل مخدوشگر بین مواجهه و پیامد در ارتباط باشند. در سال‌های اخیر برای اجرای تصادفی سازی مندلی در دو بسته Mendelian randomization و Two sample MR نرم‌افزار R پرداخته شده است.	رویکرد آماری در استفاده از داده‌های خلاصه شده

## بحث

مزایای اصلی یک مطالعه تصادفی مندلی این است که واریانت ژنتیکی، غیرقابل اصلاح هستند، بنابراین به علّیت معکوس حساس نیستند، همچنین به دلیل قانون دوم مندلی نمی‌بایست تحت تاثیر مخدوشگرها قرار بگیرد و نیز بادقت اندازه‌گیری می‌شوند، در نتیجه سوگیری کاهش می‌یابد.<sup>۱۵</sup> در واقع، رویکرد متغیر ابزاری، به یک متغیر خارجی اطلاق می‌شود که با مواجهه در ارتباط است و همچنین این متغیر خارجی به‌طور مستقیم با پیامد ارتباطی ندارد و مستقل از متغیرهای مخدوشگر است. این نگرش در تصادفی‌سازی مندلی به‌کار گرفته می‌شود.

متغیر ابزاری که در اینجا مورد استفاده قرار می‌گیرد واریانت‌های ژنتیکی از نوع SNP ها هستند. بنابراین، با شناسایی واریانت ژنتیکی که مفروضات متغیر ابزاری برای آن صدق کند، می‌توان ارتباط علّیتی بین مواجهه و پیامد را بدون تاثیرپذیری از نقش متغیر مخدوشگر برآورد نمود. بنابراین، هنگامی که ما به دنبال یافتن ارتباط علّیتی هستیم اما امکان استفاده از کارآزمایی بالینی به‌عنوان روش استاندارد وجود ندارد، می‌توان از طرح تخصیص تصادفی مندلی در مطالعات مشاهده‌ای استفاده کرد و در نهایت نتایج حاصل فصل‌الخطاب برای تعیین روابط علّیتی بین مواجهه و پیامد خواهد بود.

یکی از اهداف اساسی در تحقیقات اپیدمیولوژی برآورد صحیح اثرات عوامل خطر بر بیماری می‌باشد. با این حال، رابطه علّیتی بین مواجهه و پیامد در مطالعات اپیدمیولوژی مشاهده‌ای می‌تواند از طریق متغیرهای مخدوش‌کننده اندازه‌گیری نشده یا اندازه‌گیری شده، علّیت معکوس و سوگیری مختل شود. اگرچه این اثرات مخدوش‌کننده را می‌توان با طراحی مناسب مطالعه و همچنین در مرحله تجزیه و تحلیل کاهش داد، اما نمی‌توان اثرات همه آنها را با روش‌های یادشده کنترل نمود. از طرف دیگر، همبستگی بین مواجهه و پیامد را نمی‌توان به‌عنوان یک اثر علّی در نظر گرفت و همچنین در بسیاری از حالت‌ها محدودیت‌هایی در استفاده از کارآزمایی بالینی وجود دارد. در چنین شرایطی رویکرد تخصیص تصادفی مندلی مورد استفاده قرار می‌گیرد. تصادفی‌سازی مندلی یک نوع طرح مطالعه‌ای است که از اطلاعات ژنتیکی استفاده می‌نماید. ایده اصلی روش تصادفی‌سازی مندلی بر این اصل استوار است که با کمک روش متغیر ابزاری فارغ از غلبه بر اثرات مخربی که در مطالعات مشاهده‌ای موجود است، بتواند به‌طور همزمان نیز رابطه علّیتی بین مواجهه و پیامد را کشف کند.

## References

1. Sekula P, Fabiola Del Greco M, Pattaro C, Köttgen A. Mendelian randomization as an approach to assess causality using observational data. *Journal of the American Society of Nephrology* 2016;27(11):3253-65.
2. Davies NM, Holmes MV, Smith GD. Reading Mendelian randomisation studies: a guide, glossary, and checklist for clinicians. *bmj* 2018;362.
3. Rafiei M, Feizi A, Shariatpanahi SP, Hamraz I, Habibi D. Assessment of the Quality of Randomized Clinical Trials Published by Iranian Researchers in Persian Internal Journals in 2014. *Applied Clinical Research, Clinical Trials and Regulatory Affairs* 2020;7(1):65-70.
4. Emdin CA, Khera AV, Kathiresan S. Mendelian randomization. *Jama* 2017;318(19):1925-6.
5. Burgess S, Swanson SA, Labrecque JA. Are Mendelian randomization investigations immune from bias due to reverse causation?. *European Journal of Epidemiology* 2021;36:253-7.
6. Habibi D, Daneshpour MS, Asgarian S, Kohansal K, Hadaegh F, Mansourian M, Akbarzadeh M. Effect of C-reactive protein on the risk of Heart failure: a mendelian randomization study. *BMC Cardiovascular Disorders* 2023;23(1):1-7.
7. Nagelkerke N, Fidler V, Bernsen R, Borgdorff M. Estimating treatment effects in randomized clinical trials in the presence of non-compliance. *Statistics in medicine* 2000; 19(14): 1849-64.
8. Anderson TW, Rubin H. Estimation of the parameters of a single equation in a complete system of stochastic equations. *The Annals of mathematical statistics* 1949;20(1):46-63.
9. Conley TG, Hansen CB, McCulloch RE, Rossi PE. A semi-parametric Bayesian approach to the instrumental variable problem. *Journal of Econometrics* 2008;144(1):276-305.
10. Foster EM. Instrumental variables for logistic regression: an illustration. *Social Science Research* 1997;26(4):487-504.
11. Burgess S, Foley CN, Allara E, Staley JR, Howson JM. A robust and efficient method for Mendelian randomization with hundreds of genetic variants. *Nature communications* 2020;11(1):376.
12. Burgess S, Thompson SG. Use of allele scores as instrumental variables for Mendelian randomization. *International journal of epidemiology* 2013;42(4):1134-44.
13. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. *International journal of epidemiology* 2015;44(2):512-25.
14. Bowden J, Davey Smith G, Haycock PC, Burgess S. Consistent estimation in Mendelian randomization with some invalid instruments using a weighted median estimator. *Genetic epidemiology* 2016;40(4):304-14.
15. Bennett DA, Holmes MV. Mendelian randomisation in cardiovascular research: an introduction for clinicians. *Heart* 2017 ;103(18): 1400-7.

## Mendelian randomization, a method for inferring causal relationships in observational studies and an alternative to clinical trial studies: a brief report

### Abstract

Received: 02 Oct. 2022 Revised: 09 Oct. 2022 Accepted: 13 Jun. 2022 Available online: 22 Jun. 2022

Mahdi Akbarzadeh Ph.D.<sup>1</sup>  
 Danial Habibi Ph.D.<sup>1</sup>  
 Goodarz Kolifarhood Ph.D.<sup>1</sup>  
 Mohammad Bidkhori M.Sc.<sup>2</sup>  
 Fereidoun Azizi M.D.<sup>3</sup>  
 Maryam S. Daneshpour Ph.D.<sup>1\*</sup>

1- Cellular and Molecular Endocrine Research Center, University of Medical Sciences, Tehran, Iran.

2- Department of Epidemiology and Biostatistics, School of Public Health, Tehran University of Medical Sciences, Tehran, Iran.

3- Endocrine Research Center, University of Medical Sciences, Tehran, Iran.

\* Corresponding author: Cellular and Molecular Research Center, Research Institute for Endocrine Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran.  
 Tel: +98-21-22416264  
 E-mail: daneshpour@sbmu.ac.ir

**Background:** Mendelian randomization (MR) is a new generation in the statistical method that uses genetic variants as instrumental variables in data from non-experimental studies to evaluate and estimate the causal effects of risk factors.

**Methods:** The weakness of observational studies to detect causality, the difficulties of conducting clinical trials, the dramatic advancement of Genome-Wide Association Studies (GWAS) have led to the emergence of a new type of study called MR. It is increasingly being used to determine causality MR is an approach based on meta-analysis methods. The main idea of the MR is based on using the instrument variable (IV) to find the causality between exposure and outcome. This variable does not need to adjust the confounding effects found in observational studies.

**Results:** Data for this study were collected from the beginning of January 2003 to October 2020 in PubMed. Our results showed that MR has an increasing trend. The data used in MR includes summarized statistical data, individual-level data, and meta-analysis. Choosing the suitable IV is essential to successfully conduct an MR. For an unbiased estimate, three main hypotheses should be considered: 1) The IV has a strong relationship with the desired exposure (i.e., potential risk factor), 2) The IV is not related to the confounding variable, and 3) The IV is not directly related to the outcome and should only relate to the outcome through exposure. If these conditions are not met, one solution is to use robust methods. Besides, this research introduced the study designs, estimation methods, limitations, software packages, and some applications of MR in medical research.

**Conclusion:** When we seek to find a causal relationship, but it is not possible to use a clinical trial as a standard method, the MR design can be used in observational studies. Therefore, it is possible to obtain causal relationships between exposure and outcome using the MR.

**Keywords:** causality, genetics, medical research, mendelian randomization, observational study.