

تجزیه و تحلیل مطالعات پیش‌آزمون - پس‌آزمون با گروه کنترل: مقایسه دو اندازه اثر

سید حمید شریف نیا* آمنه یعقوب‌زاده**

نوع مقاله:

چکیده

مقاله مروری

طرح‌های پیش‌آزمون - پس‌آزمون با گروه کنترل یکی از طرح‌های متداول در تحقیقات بالینی و آموزشی است. با این که تحلیل داده‌های این طراحی به صورت‌های مختلفی انجام می‌شود، اما با توجه به تعامل بین نمره پیش‌آزمون و مداخله به عنوان عامل تهدیدکننده روایی داخلی مطالعه، کنترل آن با استفاده از آنالیز کوواریانس (ANCOVA) توصیه می‌شود. همچنین با توجه به اهمیت برآورد صحیح اندازه اثر ناشی از مداخله لازم است، تفاوت روش‌های متداول از جهت صحت برآورد آن بررسی شود. در این خصوص ANCOVA یک مدل خطی رایج که ترکیبی از آنالیز واریانس و رگرسیون است، اندازه اثر بهتری نسبت به تحلیل‌های رایج دیگر تخمین می‌زند. در این مقاله برای روشن‌تر شدن اختلاف نتایج آزمون‌های مستقل با ANCOVA، داده‌های یک مطالعه فرضی با این دو آزمون تحلیل و مقایسه شد. نتیجه نشان داد که اندازه اثر به دست آمده از ANCOVA بزرگ‌تر از آزمون‌های مستقل است. از این رو پیشنهاد می‌شود در تحلیل داده‌ها و برآورد مقدار اثر در طرح‌های پیش‌آزمون - پس‌آزمون با گروه کنترل از آزمون ANCOVA استفاده شود.

نویسنده مسئول: سید حمید شریف‌نیا؛ دانشکده پرستاری و مامایی آمل، دانشگاه علوم پزشکی مازندران

e-mail: h.sharifnia@mazums.ac.ir

واژه‌های کلیدی: مطالعات پیش‌آزمون - پس‌آزمون، آنالیز کوواریانس، پرستاری، اندازه اثر

- دریافت مقاله: اردیبهشت ماه ۱۳۹۷ - پذیرش مقاله: مرداد ماه ۱۳۹۷ - انتشار الکترونیک مقاله: ۹۷/۸/۲۹

مقدمه

مطالعات پیش‌آزمون - پس‌آزمون با گروه کنترل که به عنوان طرح تجربی کلاسیک نامیده می‌شود، به طور گسترده‌ای توسط پژوهشگران مورد استفاده قرار می‌گیرد (۱)؛ به خصوص در مطالعاتی با رویکرد آموزشی که تغییر در نمرات پیامد به عنوان نتیجه مداخله پژوهشی و متأثر از یادگیری تلقی می‌شود (۲). به طوری که هدف اکثر طرح‌های تجربی کلاسیک ارزیابی اختلاف نمرات پایه و نمرات کسب شده بعد از مداخله بین گروه‌ها است (۳). بدین منظور در قدم اول، تخصیص تصادفی افراد به گروه مداخله و کنترل صورت می‌پذیرد، تا بدون وجود تورش انتخاب،

افراد در گروه‌ها قرار بگیرند. اگرچه تصادفی‌سازی لزوماً بدین معنا نیست که گروه مداخله و کنترل از نظر ویژگی‌های جمعیت‌شناختی یکسان می‌شوند (۴)، اما با توجه به این که یکسان‌سازی گروه‌ها کار ساده‌ای نیست، باید از تخصیص تصادفی به عنوان روشی برای متعادل‌سازی گروه‌ها نسبت به متغیرهای مخدوش کننده استفاده کرد، تا نتایج تحلیل‌های آماری با اعتبار بیشتری تخمین زده شود (۵). با یا بدون تخصیص تصادفی در صورتی که گروه‌ها از نظر متغیرهای زمینه‌ای همسان نشده باشند، لازم است از روش‌های آماری موجود برای حذف اثر داده‌هایی که همگنی پایه‌ای ندارند، استفاده کنیم (۶). روش‌های آماری که جهت رفع

* گروه آموزشی پرستاری دانشکده پرستاری و مامایی آمل، دانشگاه علوم پزشکی مازندران، ساری، ایران
** گروه آموزشی پرستاری سالمندی دانشکده پرستاری و مامایی، دانشگاه علوم پزشکی تهران، تهران، ایران

ناهمگونی بین گروه‌ها از نظر نمرات پایه‌ای داده‌ها در مطالعات پیش‌آزمون- پس‌آزمون استفاده می‌شوند، بسیار مفید هستند (۷).

با توجه به ماهیت مطالعات با طرح‌های کلاسیک تجربی، ضرورت دارد، روایی داخلی که در واقع اعتماد به صحت نتایج در مورد رابطه علیتی است، مدنظر قرار گیرد تا اعتباری برای روایی خارجی که به‌عنوان درجه قابلیت تعمیم‌پذیری اثرات درمان به جمعیت تحت مطالعه، درمان و ابزارهای اندازه‌گیری است، باشد (۸). به عبارت دیگر اگرچه روایی داخلی ضمانتی قطعی برای روایی خارجی محسوب نمی‌شود، اما به‌عنوان یک ضرورت جهت روایی خارجی و تعمیم‌پذیری نتایج به شمار می‌رود (۹). یکی از عوامل تهدیدکننده روایی داخلی و خارجی، نمره پیش‌آزمون و اثر تعاملی (interaction effect) بین پیش‌آزمون و گروه است (۱۰). از طرفی دیگر اثر انتقال (carryover effect) زمانی روی می‌دهد که درمان اولیه، پاسخ مشارکت‌کنندگان را در درمان ثانویه تحت تأثیر قرار می‌دهد (۱۱). به عبارت دیگر اثر انتقال، زمانی رخ می‌دهد که پیش‌آزمون، شرکت‌کنندگان را نسبت به مداخله و درمان حساس می‌کند و پاسخ‌های آن‌ها را (نمرات پس‌آزمون) متفاوت از زمانی می‌کند که با نمرات پیش‌آزمون روبه‌رو نشده بودند (۱۲). شرکت‌کنندگان به واسطه پیش‌آزمون با نمره پایین‌تر در مقایسه با همتایان خود با نمرات بالاتر، احتمالاً نمرات کم‌تری نیز در نمره پس‌آزمون خواهند داشت، حتی اگر درمان تأثیر قابل توجهی بر نمرات پس‌آزمون داشته باشد. در مقابل، ایجاد تغییرات زیاد در نتیجه آزمون

افرادی که نمرات بالاتری در پیش‌آزمون کسب کرده‌اند، نیز مورد تردید خواهد بود (۱۳). اگرچه تهدید بلوغ و تاریخ دو عامل بزرگ تهدیدکننده روایی داخلی در این نوع طرح محسوب می‌شوند، اما تعامل بین نمره پیش‌آزمون و نتیجه درمان (مداخله) به‌عنوان یک عامل تهدیدکننده بزرگ و جدی روایی خارجی محسوب می‌شود (۱۲).

در مجموع بحث قابل توجهی نسبت به تجزیه و تحلیل داده‌های پایه‌ای این نوع مطالعه وجود دارد. به طوری که براساس بررسی متون چهار روش استفاده می‌شود: ۱- تحلیل صرفاً روی نمره‌های پس‌آزمون ۲- تحلیل روی اختلاف نمره ۳- تحلیل آنکوا روی پس‌آزمون با کنترل پیش‌آزمون ۴- تحلیل آنکوا روی اختلاف نمره با کنترل روی پیش‌آزمون (۶). اگرچه در شرایطی که اندازه‌گیری نمرات بعد از مداخله بیش از دو بار صورت بگیرد، تحلیل اندازه‌های تکراری (Repeated Measures ANOVA) نیز استفاده می‌شود (۱۲)، یکی از راه‌های کنترل اثرات نمره پیش‌آزمون به‌عنوان اثر انتقال، استفاده از آزمون آنالیز کوواریانس (ANCOVA) است. به طوری‌که شرایطی برای مطالعه فراهم می‌شود تا اثرات درمان را جدا از اثر بالقوه نمره پیش‌آزمون بررسی کند (۱۴) و به این سوال پاسخ می‌دهد که آیا میانگین گروهی تعدیل شده به‌صورت معناداری با یکدیگر تفاوت دارد یا ندارد؟ (۱۵). آنکوا با ترکیب رگرسیون و آنالیز واریانس به‌عنوان یک تکنیک آماری و با کنترل اثر نمرات پیش‌آزمون در طرح‌های پیش‌آزمون-پس‌آزمون تهدید روایی داخلی و خارجی را کنترل می‌کند و در

نتیجه نتایج قابل اعتمادتری ارائه می‌کند (۶). به‌عبارت دیگر با افزایش اعتبار نتایج، تعمیم‌پذیری مطالعه بیشتر می‌شود (۹). آنکوا با کاهش خطای واریانس بین‌گروهی موجب افزایش و تقویت توان آماری نتایج می‌شود (۱۵). برای درک این موضوع ابتدا باید آماره F را که به‌عنوان ارزیابی اختلاف بین گروه‌ها تلقی می‌شود، درک کرد. آماره F با تقسیم واریانس توضیح داده شده بین گروهی (به‌عنوان مثال بین گروه مداخله و کنترل) با واریانس غیرقابل توضیح درون‌گروهی محاسبه می‌شود. بنابراین اگر این میزان از مقدار بحرانی بیشتر باشد، می‌توان نتیجه گرفت که اختلاف بین دو گروه معنادار است (۱۶). همچنین واریانس غیرقابل توضیح شامل واریانس خطا (به‌طور مثال اختلاف بین افراد) همراه با تأثیر عوامل دیگر است (۱۷). بنابراین تأثیر مخدوش‌کننده‌ها مثل نمره پیش‌آزمون (کوواریانت‌ها) در مخرج کسر تأثیر می‌گذارد؛ به‌همین دلیل زمانی که نمره کوواریانت را بر روی نمره پیامد (مثلاً نمره پس‌آزمون) کنترل می‌کنیم، در حقیقت با کاهش مخرج کسر آماره F بیشتر می‌شود و در نتیجه توان آماری برای یافتن تأثیر معنادار (در صورت وجود) افزایش می‌یابد (۱۸). به‌طور کلی یک مزیت تحلیل به کمک آنکوا این است که توان و قدرت بیشتری برای کشف اثرات درمان و مداخله دارد (۱۹).

همانند دیگر آزمون‌های آماری، برای استفاده از آزمون ANCOVA پیش‌فرض‌هایی باید محرز شود که شامل: ۱- خطی بودن رگرسیون (ارتباط خطی رگرسیونی بین متغیر وابسته و کوواریانت)، ۲- همگنی واریانس خطاها

(خطای یک متغیر تصادفی با میانگین صفر و واریانس برابر برای گروه‌های مختلف درمان و مشاهدات)، ۳- استقلال خطاها (عدم همبستگی خطاها)، ۴- توزیع طبیعی خطاها، ۵- ارزیابی همگونی واریانس‌ها به کمک آزمون Levene و Box و ۶- همگنی شیب رگرسیون‌ها (شیب رگرسیونی خطوط مختلف در بین گروه‌ها باید برابر باشد. به‌عبارت دیگر تعامل نمرات کوواریانت و متغیر مستقل در بین گروه‌ها نباید اختلاف معناداری داشته باشد) (۲۰). یعنی این که شیب‌های خطوط رگرسیونی برای کوواریانت‌ها (در ارتباط به متغیر وابسته) نیاز است که در بین گروه‌ها یکسان باشد که به این پیش‌فرض همگنی شیب رگرسیونی گفته می‌شود که می‌تواند با یک آزمون F بر روی تعامل متغیرهای مستقل با کوواریانت ارزیابی شود. اگر آزمون F معنادار بود، بدین معنی است که این پیش‌فرض نقض شده است (۲۱). از میان فرضیات مطرح شده، فرض ارتباط خطی بین نمرات پیش‌آزمون- پس‌آزمون و همگنی شیب رگرسیون‌ها بسیار مهم است. به‌طوری‌که محققان در ابتدا باید این دو پیش‌فرض را ارزیابی کنند (۲۲).

برخی از پژوهشگران در مطالعه پیش‌آزمون- پس‌آزمون با گروه کنترل، جهت ارزیابی آزمون فرضیه وجود اختلاف معنادار پس از مداخله بین گروه‌ها، از آزمون تی مستقل به‌جای ANCOVA استفاده می‌کنند. برای روشن‌تر شدن اختلاف نتایج آزمون تی مستقل با ANCOVA، نویسندگان مقاله حاضر داده‌های یک مطالعه فرضی را با این دو آزمون تحلیل و اندازه اثر به دست آمده از تحلیل تی مستقل را با اندازه اثر ANCOVA مقایسه می‌کنند.

مثال کاربردی

پژوهشگری با هدف «تعیین تاثیر کلاس‌های آموزشی معاینه شخصی سینه بر باورهای بهداشتی نسبت به غربالگری سرطان پستان» مطالعه تجربی دو گروهی با پیش‌آزمون- پس‌آزمون طراحی کرده است. بدین منظور ۶۰ خانم مراجعه‌کننده به مرکز بهداشتی را براساس معیارهای ورود و خروج طراحی شده در مطالعه انتخاب و به صورت تصادفی به دو گروه مداخله آموزشی و شاهد تقسیم کرد. سپس به افراد هر دو گروه به عنوان پیش‌آزمون، پرسشنامه ۱۰ سوالی «باورهای بهداشتی نسبت به غربالگری سرطان پستان» را تحویل داده و از آن‌ها درخواست شد، به سوالات براساس طیف لیکرتی پنج حالتی پاسخ دهند. در مرحله بعد برای گروه مداخله به مدت چهار روز کلاس آموزشی و فیلم در جهت آموزش معاینه شخصی سینه و اهمیت ارزیابی مرتب آن برگزار شد. در این مدت گروه کنترل نه تنها مداخله آموزشی دریافت نکردند، بلکه هیچ ارتباطی با گروه مداخله نیز نداشتند. در روز پنجم مجدداً به دو گروه پرسشنامه یاد شده داده شد و از آن‌ها درخواست شد به سوالات پاسخ دهند. در نهایت داده‌ها در نرم‌افزار آماری SPSS نسخه ۲۵ به صورت نمره کل پیش‌آزمون و نمره کل پس‌آزمون در دو گروه محاسبه شد (جدول شماره ۲). جهت تجزیه و تحلیل، ابتدا توزیع طبیعی داده‌های هر گروه به کمک آزمون شاپیرو ویلکس ارزیابی شد. سپس با دو رویکرد تی مستقل (نمرات پس‌آزمون گروه مداخله و کنترل) و آزمون آنکوا (کواریانت قرار دادن نمره پیش‌آزمون)،

داده‌ها تحلیل شد. اندازه اثر حاصل از دو رویکرد به کمک نرم‌افزار 3.1 G*Power اندازه گرفته شد.

رویکرد اول: تی مستقل

میانگین و انحراف معیار نمرات پس‌آزمون باورهای بهداشتی... در گروه مداخله و کنترل بعد از مداخله به ترتیب $29/33 \pm 3/18$ و $25/00 \pm 3/41$ بود ($p < 0/001$). اندازه اثر تخمین زده شده $1/313$ به دست آمد.

رویکرد دوم: آنکوا

در ابتدا پیش‌فرض مهم همگنی شیب رگرسیون‌ها ارزیابی شد. با معنادار نشدن ($p = 0/227$) اثر تعاملی نمرات کواریانت و گروه‌ها این پیش‌فرض مهم محرز شد. همچنین آزمون لون همگنی واریانس‌ها را نیز تأیید کرد ($p = 0/469$). در ادامه، اگرچه میانگین و انحراف معیار نمرات پس‌آزمون باورهای بهداشتی... در گروه مداخله و کنترل بعد از مداخله و بدون تعدیل نمرات پیش‌آزمون به ترتیب $29/33 \pm 3/18$ و $25/00 \pm 3/41$ بود، اما با تعدیل اثر نمرات پیش‌آزمون، میانگین گروه مداخله و کنترل به ترتیب ($CI/95: 28/55-29/83$) و $29/19$ و ($CI/95: 24/49-25/77$) شد. در نهایت نتایج آزمون آنکوا نشان داد این اختلاف از نظر آماری معنادار است ($p < 0/001$). میزان η^2 0/890 تخمین زده شد. اندازه اثر تخمین زده شده به کمک این رویکرد تقریباً $2/850$ بود.

اگرچه اندازه اثر توسط فرمول‌های متعددی تخمین زده می‌شود، اما به‌طور کلی به دو صورت به دست می‌آید: ۱- تفاوت استاندارد شده بین دو میانگین (Cohen's d) و ۲- همبستگی (Pearson's r) بین متغیر مستقل و نمرات در

متغیر وابسته (۲۳). ارزش‌های عددی (جدول شماره ۱) به کمک $Pearson's r$ ، بین صفر تا یک است در صورتی که ارزش‌های $Cohen's d$ بیش از یک می‌تواند باشد (۲۴). از طرفی یکی دیگر از مفاهیم کلیدی در این زمینه توان آماری

(Statistical power) است که صحت احتمال رد فرض صفر را اندازه می‌گیرد (۲۵) و تحت تأثیر چهار عامل قرار دارد: ۱- اندازه اثر، ۲- سطح معناداری ()، ۳- توان آزمون و ۴- تعداد نمونه‌های شرکت‌کننده در مطالعه (۲۶).

جدول ۱- راهنمای اندازه اثر

اندازه	Pearson's r	Cohen's d
کوچک	۰/۱-۰/۳	<۰/۲۵
متوسط	۰/۳-۰/۵	۰/۲۵-۰/۴
بزرگ	۰/۵-۱	۰/۴-

جدول ۲- نمرات به دست آمده از پرسشنامه باورهای بهداشتی نسبت به غربالگری سرطان سینه

گروه	پیش‌آزمون	پس‌آزمون	گروه	پیش‌آزمون	پس‌آزمون
۱	۲۵	۳۲	۲	۱۸	۲۰
۱	۲۳	۲۷	۲	۲۵	۲۵
۱	۲۱	۲۶	۲	۲۴	۲۵
۱	۲۴	۲۹	۲	۲۳	۲۳
۱	۲۶	۳۰	۲	۲۱	۲۲
۱	۲۵	۳۱	۲	۲۵	۲۴
۱	۲۶	۳۳	۲	۲۳	۲۳
۱	۲۱	۲۶	۲	۲۱	۲۲
۱	۲۰	۲۴	۲	۲۶	۲۷
۱	۲۰	۲۵	۲	۲۵	۲۷
۱	۲۶	۳۳	۲	۱۷	۱۸
۱	۲۴	۲۸	۲	۲۶	۲۶
۱	۲۲	۲۷	۲	۲۵	۲۶
۱	۲۵	۳۰	۲	۲۴	۲۴
۱	۲۷	۳۱	۲	۲۲	۲۳
۱	۲۶	۳۲	۲	۲۶	۲۵
۱	۲۹	۳۵	۲	۲۴	۳۶
۱	۲۲	۲۷	۲	۲۲	۲۳
۱	۲۱	۲۵	۲	۲۷	۲۸
۱	۲۱	۲۶	۲	۲۶	۲۸
۱	۲۷	۳۴	۲	۱۸	۱۹
۱	۲۵	۲۹	۲	۲۷	۲۷
۱	۲۳	۲۸	۲	۲۶	۲۷
۱	۲۶	۳۱	۲	۲۵	۲۵
۱	۲۸	۳۲	۲	۲۳	۲۴
۱	۲۷	۳۳	۲	۲۷	۲۶
۱	۳۰	۳۵	۲	۲۵	۲۵
۱	۲۳	۲۸	۲	۲۳	۲۴
۱	۲۲	۲۶	۲	۲۶	۲۵
۱	۲۲	۲۷	۲	۲۶	۲۶

نتیجه‌گیری

همان‌طور که در مثال بالا نشان داده شد از دو رویکرد مختلف برای مقایسه و تحلیل نمرات به دست آمده در طرح تجربی دو گروهی با پیش‌آزمون- پس‌آزمون استفاده شد. اندازه اثر به دست آمده از آنکوا بیش از رویکرد آزمون‌تی‌مستقل بود. به نظر می‌رسد آزمون آنکوا که در حقیقت یک رگرسیون خطی چندگانه را اجرا می‌کند، رویکردی مناسب برای مقایسه و تحلیل این دسته طرح‌ها می‌باشد. اما برخی دیگر بیان می‌کنند در صورتی که پیش شرط مهم همگنی شیب رگرسیون‌ها برقرار نباشد، باید از رویکردهای دیگری مثل آنکوا ناپارامتری Quade's، آنکوا ناپارامتری Sen's و Puri و یا آنکوا ناپارامتری برای رتبه‌ها استفاده کرد (۶).

نویسندگان مقاله حاضر تاکید دارند که در مطالعات پیش‌آزمون- پس‌آزمون با گروه کنترل،

جهت تعدیل نمره پیش‌آزمون از آزمون آنکوا استفاده کنند. همچنین بر این دو نکته اصرار دارند؛ اول، در صورتی که نمونه‌ها به صورت تصادفی شده به گروه‌ها تخصیص داده نشده باشند، محققان نتایج تحلیل این نوع طرح‌ها را با احتیاط تفسیر کنند و دوم این‌که اندازه اثر آزمون را تخمین زده و گزارش نمایند.

تشریح و قدردانی

نویسندگان مقاله حاضر بر این باورند که پیشنهادات و توصیه‌های دقیق مشاور آمار مجله حیات در اصلاح و ارتقا سطح کیفی مقاله حاضر کمک بسیار شایان و ارزنده‌ای در محتوای مقاله داشته است و توانسته موجب روشن‌تر شدن اهمیت موضوع شود. لذا نویسندگان بر خود واجب می‌دانند مراتب قدردانی و قدرشناسی از مشاور محترم آمار مجله حیات را بیان کنند.

منابع

- 1 - Schmidt NA, Brown JM. Evidence-based practice for nurses: appraisal and application of research. 3rd ed. Burlington, Massachusetts: Jones & Bartlett Learning; 2014.
- 2 - Dugard P, Todman J. Analysis of pre-test-post-test control group designs in educational research. *Educational Psychology*. 1995; 15(2): 181-98.
- 3 - Johnson B, Christensen LB. Educational research: quantitative, qualitative, and mixed approaches. 4th ed. Thousand Oaks, California: SAGE Publications; 2010.
- 4 - Altman DG, Dore CJ. Randomisation and baseline comparisons in clinical trials. *Lancet*. 1990 Jan 20; 335(8682): 149-53.
- 5 - Zhao W, Berger V. Imbalance control in clinical trial subject randomization-from philosophy to strategy. *J Clin Epidemiol*. 2018 Sep; 101: 116-118.
- 6 - Bonate PL. Analysis of pretest-posttest designs. 1st ed. Boca Raton: Chapman & Hall/CRC; 2000.
- 7 - Roberts MC, Ilardi SS. Handbook of research methods in clinical psychology. Chichester: Wiley-Blackwell; 2008.

- 8 - Adams KA, Lawrence EK. Research methods, statistics, and applications. 1st ed. Thousand Oaks, California: SAGE Publications; 2014.
- 9 - Bonita R, Beaglehole R, Kjellstrom T. Basic epidemiology. 2nd ed. Geneva: World Health Organization; 2006.
- 10 - Lavrakas PJ. Encyclopedia of survey research methods. Thousand Oaks, California: SAGE Publications; 2008.
- 11 - Polit DF, Beck CT. Nursing research: generating and assessing evidence for nursing practice. 10th ed. Philadelphia: Lippincott Williams & Wilkins; 2017.
- 12 - Dimitrov DM, Rumrill PD Jr. Pretest-posttest designs and measurement of change. *Work*. 2003; 20(2): 159-65.
- 13 - Grove SK, Gray JR, Burns N. Understanding nursing research: building an evidence-based practice. 6th ed. St. Louis, Missouri: Elsevier Saunders; 2015.
- 14 - Verma JP. Data analysis in management with SPSS software. New Delhi, India: Springer; 2013.
- 15 - Tabachnick BG, Fidell LS. Using multivariate statistics. 7th ed. London: Pearson; 2018.
- 16 - Pace L. The excel data and statistics cookbook. 3rd ed. United States; 2012.
- 17 - Jackson SL. Research methods: a modular approach. 1st ed. Belmont, CA: Thomson Wadsworth; 2007.
- 18 - Kirk RE. Experimental design: procedures for the behavioral sciences. 4th ed. Thousand Oaks, California: SAGE Publications; 2012.
- 19 - Vickers AJ, Altman DG. Analysing controlled trials with baseline and follow up measurements. *BMJ*. 2001; 323: 1123-1124.
- 20 - Field A. Discovering statistics using IBM SPSS statistics 4e + eBook + Web Assign Single Term + SPSS version 23.0. 5th ed. London: SAGE Publications; 2018.
- 21 - Morgan GA, Leech NL, Gloeckner GW, Barrett KC. IBM SPSS for introductory statistics: use and interpretation. 5th ed. New York: Taylor & Francis; 2013.
- 22 - Gliner JA, Morgan GA, Harmon RJ. Pretest-posttest comparison group designs: analysis and interpretation. *J Am Acad Child Adolesc Psychiatry*. 2003 Apr; 42(4): 500-3.
- 23 - Rosnow RL, Rosenthal R. Definition and interpretation of interaction effects. *Psychological Bulletin*. 1989; 105(1): 143-146.
- 24 - Sullivan GM, Feinn R. Using effect size-or why the p-value is not enough. *J Grad Med Educ*. 2012 Sep; 4(3): 279-82.
- 25 - Haghdoost AA. [Do you want to gain a profound insight into sample size and statistical power]. *Iranian Journal of Epidemiology*. 2009; 5(1): 57-63. (Persian)
- 26 - Mayers A. Introduction to statistics and SPSS in psychology. Harlow: Pearson Education Limited; 2013.

Analysis of pretest-posttest studies with control group: Comparison of two types of effect sizes

Seyed Hamid Sharif Nia* Ameneh Yaghoobzadeh**

Abstract

Article type:
Review Article

Received: May 2018
Accepted: Aug. 2018
e-Published: 20 Nov. 2018

Corresponding author:
Seyed Hamid Sharif Nia
e-mail:
h.sharifnia@mazums.ac.ir

The pre-post test study with control group is one of the most widely used methods to evaluate clinical research and educational designs. The data analysis of this design is done in a variety of ways. However, considering that the interaction between the pre-test score and intervention score is considered a threat for external validity in these types of study designs, the ANCOVA test is recommended for its control. Also, due to the importance of accurately estimating the effect size of intervention, it is necessary to assess the differences in the accuracy of its estimate among conventional methods. In this regard, the ANCOVA test, a common linear model that combines the analysis of variance and regression, estimates the effect size better than other common analyzes. In this article, the data of a hypothetical study were analyzed using these two tests to clarify the different results between the independent *t*-test and the ANCOVA test. The result showed that the obtained effect size from ANCOVA was greater than that from the independent *t*-test. Therefore, it is suggested that the ANCOVA test be used to analyze the data and estimate the effect size in the pre-post test designs with control group.

Key words: pre-post test study, analysis of covariance, nursing, effect size

Please cite this article as:

- Sharif Nia SH, Yaghoobzadeh A. [Analysis of pretest-posttest studies with control group: Comparison of two types of effect sizes]. *Hayat, Journal of School of Nursing and Midwifery, Tehran University of Medical Sciences.* 2018; 24(3): 261-268. (Persian)

* Dept. of Nursing, School of Nursing and Midwifery Amol, Mazandaran University of Medical Sciences, Sari, Iran

** Dept. of Geriatric Nursing, School of Nursing and Midwifery, Tehran University of Medical Sciences, Tehran, Iran