

## روش‌های گروهی دنباله‌ای و کاربرد آن در کارآزمایی‌های بالینی

دکتر انوشیروان کاظم نژاد\*؛ دکتر فرناز سهراب‌وند\*\*؛ محمد حسن رخشانی\*\*\*

### چکیده:

یکی از مباحثی که در کلیه علوم مورد دغدغه محققان است، برآورد تعداد نمونه (آزمودنی) کافی برای آزمودن شاخص‌های مورد بررسی مانند میانگین، نسبت و ... است. معمولاً پژوهشگران با یک سطح اطمینان ثابت مانند ۹۵٪ و توان آزمون مانند ۹۰٪ و با در نظر گرفتن خطای معین، و با فرمول‌های تک نمونه‌ای یا دو نمونه‌ای، حجم ثابتی را برای تعداد نمونه یا آزمودنی‌ها برآورد می‌نمایند. در کارآزمایی‌های بالینی به سه دلیل اخلاقی، مالی و اجرایی انجام دادن کارآزمایی با حجم نمونه ثابت از پیش تعیین شده امکانپذیر نیست. یکی از روش‌هایی که می‌تواند در این زمینه مفید باشد، به کارگیری نمونه‌گیری گروهی دنباله‌ای است. طرح‌های گروهی دنباله‌ای، بخشی از طرح آزمایش‌ها در آمار است که داده‌ها را در زمان‌های متوالی مورد تجزیه و تحلیل قرار می‌دهد و تصمیم مناسب در خصوص موضوع مورد بررسی در هر یک از مراحل اتخاذ می‌گردد. در روش‌های گروهی دنباله‌ای سعی بر این است تا با انتخاب آزمون‌های مناسب نظیر پوکاک، ابراین-فلمینگ و ... و با کاهش حجم نمونه نسبت به حالت معمولی، راه‌حلی به منظور مقایسه گروه‌های تحت مطالعه ارائه گردد.

ما این روش‌ها را در بررسی تأثیر دوز چندگانه و تک دوز داروی کورتیکواستروئید در زنان مبتلا به پارگی کیسه آب پیش از موعد و پیش از ترم به کار گرفته ایم و نشان داده ایم که طرح در حدود ۳ سال زودتر از حالت استفاده از روش‌های با حجم نمونه ثابت به انتها می‌رسد.

**کلیدواژه‌ها:** کارآزمایی بالینی، روش‌های گروهی دنباله‌ای، پارگی کیسه آب پیش از موعد و پیش از ترم.

\* دانشیار گروه آمار زیستی دانشکده علوم پزشکی دانشگاه تربیت مدرس.

\*\* استادیار گروه زنان و زایمان دانشگاه علوم پزشکی تهران.

\*\*\* کارشناس ارشد آمار زیستی، دانشکده علوم پزشکی دانشگاه تربیت مدرس.

✉ عهده دار مکاتبات: تهران، دانشگاه تربیت مدرس، دانشکده علوم پزشکی، معاونت پژوهشی، صندوق پستی: ۱۱۱-۱۴۱۱۵

## مقدمه:

افزایش در طول مدت کارآزمایی، باعث افزایش در میزان هزینه‌های مالی طرح می‌شود. مواقعی نیز پیش‌می‌آید که ماهیت طرح به‌گونه‌ای است که برای آشکارسازی یک اختلاف، به عنوان مثال میزان اختلاف در کارایی یک داروی جدید در برابر یک داروی استاندارد، مجبور به انتخاب بیماران (نمونه) زیادی هستیم. این امر در کنار ملاحظات که باید در قبال افراد از دست‌رفته یا گمشده در طول کارآزمایی بالینی و ... بنماییم، باعث افزایش طول مدت طرح و در نتیجه افزایش هزینه‌ها می‌گردد و زمانی که برای هر فرد بیمار به به‌کارگیری روش درمانی گران‌قیمتی نیاز باشد یا آنکه برخی از روش‌ها به شکل توأم در روی بیمار اجرا شود، مانند جراحی و پرتودرمانی به شکل توأم در مورد یک سرطان خاص، شکل حادثی به خود می‌گیرد.

می‌دانیم که پس از طی مراحل اولیه یک طرح تحقیقاتی مانند تعریف مسأله، اهداف طرح و ... یکی از ابتدایی‌ترین کارها، تعیین حجم نمونه موردانتظار بر مبنای برخی پارامترهای از پیش تعیین‌شده مثل واریانس و میزان اختلاف میان دو تیمار است. آنچه هم‌اکنون در کشور ما در باب کارآزمایی‌های بالینی مرسوم است، تعیین حجم نمونه ثابت برای انجام دادن کارآزمایی است، اما به سه دلیل اخلاقی، مالی و اجرایی این روش مناسب نیست، مثلاً: هدف یک کارآزمایی بالینی انجام‌شده در کشورهای اسکانندیناوی (۲) مقایسه یک داروی استاندارد (در آن زمان) با پروپانولول در مورد بیماری‌های حاد قلبی بوده است. متخصصان حجم نمونه مورد نیاز را بر مبنای روش‌های ثابت در حدود ۹۰۰۰ نفر تعیین

در مطالعات پزشکی، کارآزمایی‌های بالینی نقش مهمی را ایفاء می‌کنند. پوکاک کارآزمایی‌های بالینی را بخشی از تحقیقات آماری می‌داند که روی جامعه انسانی صورت می‌پذیرد (۱ و ۲). هر پزشک و محقق امور بهداشتی می‌داند که در انجام مطالعات تحقیقاتی در روی جامعه انسانی، باید از هر امری که سلامت روحی و جسمانی افراد تحت مطالعه را به خطر اندازد، پرهیز نماید. رعایت مسائل اخلاق پزشکی امری است که نه تنها فرد محقق به آن اهتمام می‌ورزد، بلکه کمیته اخلاق پزشکی نیز که متشکل از گروهی از متخصصان است، در تمام مراحل تحقیق به آن توجه خاص دارد. از سوی دیگر، کلیه پژوهشگرانی که در زمینه کارآزمایی‌های بالینی فعالیت دارند، اذعان می‌نمایند که اکثر این تحقیقات دراز مدت می‌باشند. این امر ممکن است به دلایل مختلفی به وقوع پیوندد. گاهی اوقات به دلیل اینکه بیماری جزو بیماری‌های نادر است، اخذ بیمار با کندی صورت می‌پذیرد. در برخی مواقع گستره جغرافیایی طرح، به این امر دامن می‌زند. نوع نگرش مردم به تحقیقات پزشکی و میزان اعتماد آن‌ها به مسئولین یک کارآزمایی بالینی، در پاره‌ای از مواقع باعث می‌شود که حتی اگر بیماری در حیطة بیماری‌های مرسوم و رایج نیز باشد، به‌دست آمدن حجم نمونه مورد نیاز کاری بس طاقت‌فرسا گردد. تمام این عوامل همراه با عوامل آشکار و نهان دیگر باعث می‌گردد که طول مدت یک کارآزمایی بالینی افزایش یابد. از آنجا که در عرصه یک کارآزمایی بالینی از افراد متخصص زیادی استفاده می‌شود، این

سیمواستاتین<sup>۲</sup> در کاهش چربی خون در افراد با کلسترول بالا و اثربخشی آن در کاهش مرگ و میر ناشی از بالا بودن کلسترول و حملات قلبی به علت تنگی عروق قلب بود. همچنین تصور می‌شد که سیمواستاتین در طول این طرح در حدود ۳۰ درصد از مرگ‌ها را کاهش می‌دهد. حجم نمونه به روش ثابت در حدود ۴۰۰۰ نفر با  $\alpha = \beta = 0.05$  تعیین گردید. اگر این کارآزمایی با فرض درست بودن کاهش ۳۰ درصدی مرگ و میر ناشی از سیمواستاتین، به روش معمول برپا شود، ۴۰۸ مرگ رخ خواهد داد. این محققان یک طرح گروهی دنباله‌ای را با توقف زودهنگام به نفع فرض صفر به‌کاربردند و تصمیم گرفتند (به دلیل شرایط حاکم بر آزمون گروهی دنباله‌ای) اطلاعات را زمانی که ۱۹۰ مرگ، ۳۰۰ مرگ و ۳۸۰ مرگ رخ دهد، جمع‌آوری کنند. بدین ترتیب، اگر آزمون در مرحله نخست متوقف شود، ۲۱۸ مورد از تعداد مرگ‌ها کاسته می‌شود. با یک تناسب ساده به این نتیجه می‌رسیم که این تعداد کاهش در مرگ‌ها معادل ۲۱۳۸ بیمار کمتر نسبت به حالت حجم نمونه ثابت است. این کاهش برای مراحل بعدی به ۱۰۶۰ و ۲۷۵ بیمار کمتر می‌رسد. در جایی دیگر، لن و روزنبرگ کارآزمایی را در ۴۸ ماه انجام دادند و از ژوئن ۱۹۷۸ شروع کردند. آن‌ها ۳۸۳۷ نفر را به دو گروه دارونما و پروپانولول تقسیم نمودند و داده‌ها را به طور تقریبی هر ۶ ماه یک‌بار جمع‌آوری کردند و نشان دادند که انجام این کارآزمایی به روش گروهی

نمودند. این درحالی بود که اگر این حجم بیمار تحت دو تیمار قرار می‌گرفتند، در حدود ۱۰۰۰ نفر از آن‌ها به علت عوارض مختلف ناشی از بیماری فوت می‌کردند. مسلم است که کمیته اخلاق پزشکی اجازه برپایی این کارآزمایی را به آن‌ها نمی‌دهد.

به عنوان مثالی دیگر، اگر در مرحله اول کارآزمایی بالینی در مورد یک داروی جدید، که هدف آن بررسی میزان سلامت و بی‌خطری دارو است، از روش‌های نمونه‌گیری ثابت استفاده کنیم، در عمل نیمی از داوطلبان را در معرض داروی جدید قرار داده‌ایم. حال اگر این دارو از میزان ایمنی کافی برخوردار نباشد، یا تأثیرات جانبی داشته باشد که در درازمدت آشکار می‌شود، در بهترین حالت، سلامت نیمی از داوطلبان را در معرض خطر قرار داده‌ایم.

حال اگر بتوانیم روشی را پیدا کنیم که نیاز به تعداد نمونه کمتری نسبت به روش‌های نمونه‌گیری ثابت داشته باشد و با ماهیت ورود متوالی بیماران در کارآزمایی‌های بالینی سنخیت داشته باشد و هزینه‌ها را نیز کاهش دهد، بسیار مفید است. نام این روش‌های نوین، روش‌های گروهی دنباله‌ای<sup>۱</sup> است.

روش‌های گروهی دنباله‌ای در مورد بیماری‌های زیادی به کار رفته است که می‌توان به نمونه‌های ذیل اشاره نمود.

هاوانگ و همکاران به کارآزمایی بالینی چند مرکزی، سه سو کور، دارونما و کنترلی اشاره نموده‌اند<sup>(۳)</sup> که هدف از آن بررسی اثر داروی

1. Group Sequential Methods

2. Simvastatin

دنباله‌ای ۸ ماه کاهش را در مدت‌زمان طرح نسبت به حالت حجم نمونه ثابت به همراه دارد (۴).

### مواد و روش‌ها:

فرض کنید که طرح تحقیقاتی ما شامل دو تیمار استاندارد و جدید باشد. آنچه در روش‌های نمونه‌گیری ثابت مرسوم است، این است که ابتدا با معلوم‌بودن میزان خطاها و برخی پارامترهای دیگر حجم نمونه مورد نیاز را در هر تیمار معین می‌کنیم. سپس تمام این حجم نمونه را تحت دو تیمار قرار می‌دهیم و پس از آنکه پاسخ‌های لازم از تمام نمونه (با احتساب میزان ریزش، یعنی تعداد نمونه‌هایی که از دست می‌دهیم) به دست آمد، آزمون موردنظر را انجام می‌دهیم. به عنوان مثال، حجم نمونه برای یک آزمون دوطرفه با پاسخ‌های نرمال از فرمول زیر به دست می‌آید.

$$n_f = \frac{(Z_{1-\alpha/2} + Z_{1-\beta})^2 2\sigma^2}{\delta^2}$$

که در آن  $\sigma^2$  واریانس مشترک دو گروه و  $\delta$  میزان اختلاف بالینی در دو تیمار است. در روش‌های گروهی دنباله‌ای، بر خلاف روش‌های نمونه‌گیری ثابت، لزومی به تعیین کردن حجم نمونه قبل از آزمون نیست.

در برخی از روش‌های گروهی دنباله‌ای، محقق تصمیم می‌گیرد که اطلاعات مورد نیازش را در  $K$  مرحله (به عنوان مثال در ۵ یا ۱۰ مرحله) با رجوع به کارآزمایی جمع‌آوری نماید. حال بر مبنای این  $K$  و توان آزمون و خطای مشخص، تعداد بیماران در هر تیمار و در هر مرحله ( $m$ ) را به دست می‌آورد. این

تعداد (برای آزمون پوکاک) از فرمول  $m = \frac{R_p(K, \alpha, \beta) \times n_f}{K}$  به دست می‌آید که مقدار ثابت  $R_p(K, \alpha, \beta)$  از جدول ۱ استخراج می‌گردد. اکنون با استفاده از یک آزمون خاص، مثلا آزمون تی یا آزمون خی-دو معمولی، نتایج  $2m$  بیمار را مقایسه می‌نماییم (عدد ۲ به دلیل این است که دو تیمار داریم). برای مثال برای آزمون دوطرفه برابری دو میانگین در جامعه نرمال، فرمول مربوط برای دو تیمار  $A$  و  $B$  عبارت است از:

$$Z_k = \frac{1}{\sqrt{2mk\sigma^2}} \left[ \sum_{i=1}^{mk} X_{A_i} - \sum_{i=1}^{mk} X_{B_i} \right], k=1,2,\dots,K$$

پس از اینکه نیمی از این  $2m$  بیمار را تحت تیمار استاندارد و نیم دیگر را تحت تیمار جدید قرار دادیم و مشاهدات لازم را به دست آوردیم، اعداد حاصل را در فرمول فوق قرار می‌دهیم و یک عدد به دست می‌آوریم. نکته قابل توجه در این مرحله این است که چون طرح را به شکل بررسی متوالی انجام می‌دهیم، میزان سطح معناداری از میزان از پیش تعیین شده  $\alpha$  بیشتر می‌شود و سطح اطمینان آزمون کاهش می‌یابد. پس مقایسه مقدار آماره به دست آمده با اعداد جداول معمول توزیع تی و خی-دو صحیح نمی‌باشد و باید تدبیر جدیدی برای آن در نظر بگیریم.

پوکاک با استفاده از محاسبات عددی، حدود بحرانی جدیدی را به دست آورد که در عین حال که در تمام  $K$  مرحله مقدار یکسانی داشت (۲)، میزان سطح معناداری نیز در تمام  $K$  مرحله از میزان از پیش تعیین شده  $\alpha$  تجاوز نمی‌کرد (جدول ۲). حال

غیرقابل اعمال می‌شوند. به عنوان مثال، اگر بیماری رابطه مستقیم با فصل خاصی داشته باشد و طرح تحقیقاتی در فصلی شروع شود که تعداد بیماران در آن فصل ناچیز است، به‌دست‌آوردن تعداد بیمار یکسان برای دو یا چند تیمار کار مشکلی است. لن و دیمتس با ارائه روش اصولی به حل این مشکل پرداختند (۲). روش آن‌ها یک مزیت دیگر نیز داشت و آن این بود که در هر مرحله از آزمون که متوقف شویم، میزان سطح معناداری دقیقاً برابر  $\alpha$  است. حال آنکه در روش‌های قبل، به عنوان مثال، اگر کارآزمایی با ۵ مرحله تکرار برپا می‌شد و پس از ۳ مرحله جمع‌آوری اطلاعات خاتمه می‌یافت، میزان سطح معناداری از  $\alpha$  کمتر و سطح اطمینان افزایش می‌یافت.

پامپالونا و تسیاتیس دست به ابداع آزمون کارآمد دیگری زدند که در مجامع علمی به روش خانواده توانی مشهور است (۷). این روش قابلیت حیرت‌انگیزی در خاتمه زودهنگام کارآزمایی به نفع فرض صفر دارد. در این روش بر خلاف روش‌های قبل، در هر مرحله، از حد بالا و پایین متفاوتی استفاده می‌شود. به‌دست‌آوردن این حدود نیز مستلزم به‌کارگیری محاسبات عددی است و جدول ۳ و پیوست (ب) به ضرایب لازم برای به‌دست‌آوردن این حدود اختصاص داده شده است. در این آزمون، به عنوان مثال برای بررسی مثبت‌بودن تفاضل میانگین دو جامعه، اگر در هر یک از مراحل ۱- $K, \dots, 2, 1$  مقدار آماره مناسب آزمون کمتر یا مساوی حد تحتانی شد ( $a_k$ ) با قبول فرض صفر حکم به خاتمه آزمون می‌دهیم و اگر بزرگتر یا

اگر در یک آزمون دو طرفه مقدار قدر مطلق آماره آزمونی که محقق به‌دست آورده، از عددی که پوکاک به‌دست آورده بزرگ‌تر یا با آن مساوی بود، آزمون در همین مرحله و با حجم نمونه بسیار اندک  $2m$  و با رد فرض صفر خاتمه می‌یابد و در غیراین صورت به مرحله بعد می‌رود و  $2m$  بیمار دیگر می‌گیرد و این روال را با  $m=2m+2$  بیمار ادامه می‌دهد. اگر فرض صفر رد شد که کار تمام است وگرنه به مرحله بعدی می‌رود. حال اگر فرض صفر تا مرحله  $K$  ام دوام آورد، در این مرحله اگر مقدار قدر مطلق آماره آزمون که بر اساس  $2Km$  بیمار به‌دست آورده‌شده از مقدار عدد بحرانی پوکاک بیشتر یا مساوی آن بود، فرض صفر را رد می‌کند و در غیر این صورت آن‌را می‌پذیرد.

به فاصله دو سال از مقاله پوکاک، ابراین و فلمینگ مقاله‌ای ارائه دادند که تفاوت آن با کار پوکاک در اعداد بحرانی به‌دست آمده بود (۵). محققان نشان داده‌اند که بهتر است از آزمون پوکاک برای آزمون‌های ۵ الی ۶ مرحله‌ای و از آزمون ابراین-فلمینگ برای آزمون‌های ۱۰ مرحله‌ای استفاده نماییم. توجه دارید که هر چه تعداد مراحل زیادتر شود،  $m$  یا تعداد بیماران در هر مرحله و در هر تیمار، کاهش می‌یابد. بعد از کار آن‌ها وانگ و تسیاتیس روشی را ابداع کردند که هم دو روش فوق را پوشش می‌داد و هم حالت‌های بین دو روش فوق را در برمی‌گرفت (۶).

دو شرط اعمال‌شده در آزمون‌های مذکور، یعنی معلوم‌بودن تعداد مراحل آزمون ( $K$ ) و مساوی‌بودن تعداد بیماران در هر مرحله ( $m$ ) در پاره‌ای از اوقات

مساوی حد فوقانی ( $b_k$ ) گردید، حکم به رد فرض صفر می‌دهیم. در مرحله K ام اگر مقدار آماره آزمون از حد تحتانی کمتر شد، حکم به قبول فرض صفر می‌دهیم و در غیر این صورت با رد آن به کار خود خاتمه می‌دهیم. ما از این آزمون برای مسأله مقایسه دو دوز متفاوت از داروی کورتیکواستروئید در بیماری پارگی کیسه آب پیش از موعد و پیش از ترم استفاده نموده‌ایم.

توسعه روزافزون روش‌های گروهی دنباله‌ای، حالت‌های عملی زیادی را پوشش داده است. به عنوان مثال اگر داروی جدیدی کشف شود که از لحاظ تأثیرات جانبی و جنبه اقتصادی بر داروی استاندارد رجحان داشته باشد، معمولاً در آزمون میزان کارایی آن، به جای بررسی برتری داروی جدید به داروی استاندارد، از آزمون همسنگ بودن کارایی آن‌ها استفاده می‌شود و محقق بر این باور است که به دلیل دو مزیت یادشده، اگر کارایی داروی جدید نزدیک به داروی استاندارد باشد، می‌توانیم آن‌را جایگزین داروی استاندارد بنماییم.

### یافته‌ها:

بیماری پارگی کیسه آب پیش از موعد و پیش از ترم در ۱٪ کل زایمان‌ها رخ می‌دهد، اما درصد قابل توجهی از مرگ و میر نوزادان بر اثر این ضایعه و تأثیرات مرتبط با آن است، از جمله انتروکولیت نکروزان، اندومتريت، سندرم زجر تنفسی، عفونت منتشر، خونریزی داخل بطنی و تولد با وزن کم. برای جلوگیری از ایجاد التهاب‌های داخل رحم که

بر اثر عفونت ایجاد می‌شود، داروی کورتیکواستروئید به‌طور هفتگی از زمان پارگی کیسه آب تا زمان تولد نوزاد تجویز می‌شود. در مطالعات اخیر تفاوت چندانی بین تجویز مکرر این دارو با تجویز تک‌نوبتی آن مشاهده نشده است (۸). ما قصد داریم به بررسی این ادعا با روش‌های گروهی دنباله‌ای بپردازیم. از میان متغیرهای طرح، متغیر دوره تأخیر<sup>۱</sup> را که عبارت است از فاصله میان پارگی کیسه آب تا تولد نوزاد (برحسب روز) انتخاب می‌کنیم و کارایی آزمون خانواده توانی را در آن نشان می‌دهیم. این آزمون را از آن جهت برگزیده‌ایم که براساس بررسی‌های قبلی تفاوت چندانی بین دو تیمار دیده نشده است و طبیعی است که به دنبال خاتمه زودهنگام کارآزمایی باشیم.

با توجه به وجود اختلاف ۳ روزه مابین دوره تأخیر در دو تیمار تک دوز و دوز چندگانه، حجم نمونه به روش معمولی (ثابت) با  $\alpha = 0.05$ ,  $\beta = 0.1$  در حدود ۸۰ بیمار برای هر تیمار (۱۶۰ بیمار برای کل طرح) تعیین گردید. با توجه به درصد رخداد این عارضه پیش‌بینی شد که برای به‌دست آوردن این تعداد نمونه مدت زمان ۴ سال نیاز است. به این دلیل از روش‌های گروهی دنباله‌ای استفاده نمودیم و مادران بارداری را که دچار این ضایعه شده بودند، به دو طبقه هفته بارداری ۳۲-۲۸ و ۳۲ به بعد تقسیم کردیم. این تقسیم‌بندی بدین‌منظور صورت پذیرفت که هفته بارداری اثر مستقیم بر طول دوره تأخیر دارد. سپس فرمول تعیین حجم نمونه مربوط به آزمون خانواده توانی را به‌کار گرفتیم و تعداد بیماران لازم در هر مرحله و در هر

بدین ترتیب آزمون با ۲۸ بیمار و مدت زمان یک سال خاتمه می‌یابد. این در حالی است که با روش نمونه‌گیری ثابت به ۱۶۰ بیمار و مدت زمانی نزدیک به ۴ سال نیاز داشتیم. آزمون مشابهی برای بیماری واژنیت صورت پذیرفته است. آزمون اصلی که در پی آشکارسازی عدم تفاوت کارایی مترونیدازول با کلیندامایسین بوده، به روش نمونه‌گیری ثابت انجام شده است. برای بررسی نسبت سلول‌های کلیدی در دو داروی ذکر شده، آزمونی را انجام و نشان دادیم که در کل آزمون کاهش معادل ۲۶ بیمار با روش‌های گروهی دنباله‌ای رخ داده است (در روش نمونه‌گیری ثابت به ۱۲۵ بیمار نیاز داشتیم).

#### بحث:

به سه علت اخلاقی، مالی و اجرایی روش‌های نمونه‌گیری ثابت در کارآزمایی بالینی نه تنها کاربرد ندارد، بلکه مجاز نیز نیست. روش‌های گروهی دنباله‌ای مشکلات فوق را از کاهش حجم نمونه تا سرحد ممکن حل می‌کند، اما مبانی نظری این روش‌ها از پیچیدگی خاصی برخوردار است. مطالعات متعددی برای آشکارسازی قابلیت‌های این روش‌ها انجام شده است. به عنوان مثال می‌توان به بررسی مقایسه پیوند مغز استخوان و شیمی‌درمانی در کم‌خونی حاد میلوئید حاد اشاره نمود که به‌کارگیری این روش‌ها طول مدت طرح را از ۵ به ۳ سال تقلیل داده است (۹). همچنین از این روش‌ها در بررسی کارایی و سلامت داروی زیدوودین<sup>۱</sup> در بیماری ایدز یا بررسی دورژیم دارویی متفاوت در

بیمار و در هر طبقه را ۷ نفر به دست آوردیم و پس از آن با استفاده از جدول اعداد تصادفی، به تخصیص بیماران به تیمارها پرداختیم. برای بیمار اول، اگر عدد تصادفی بدست آمده فرد بود، وی را به تیمار اول (تک دوز) اختصاص می‌دادیم و بیمار بعدی را به تیمار دوم (دوز چندگانه) تخصیص می‌دادیم. این عمل در صورت به دست آمدن عدد زوج حالت معکوس به خود می‌گرفت. پس از اخذ ۲۸ بیمار (۷ بیمار برای هر تیمار و هر طبقه) و با استفاده از اطلاعات اولیه به دست آمده (۸ اقدام به تحلیل داده‌ها نمودیم. براساس اطلاعات مذکور مقدار آماره را از اطلاعات به دست آمده (این اطلاعات از مراکز درمانی شیراز جمع‌آوری شده‌اند) محاسبه نمودیم. ابتدا برای هفته بارداری ۲۸ الی ۳۲ داریم:

$$Z_1 = \frac{\sum X_i - \sum B_i}{\sqrt{m\sigma^2}} = \frac{132 - 169}{\sqrt{7 \times 36.04}} = -2.3$$

به وضوح این مقدار از حد پایین ( $a_1 = -0.765$ ) کوچک‌تر است، لذا بر مبنای این اطلاعات دلیلی بر رد فرض صفر مبنی بر عدم اختلاف دو دوز دارو نداریم. به همین شکل برای ۳۲ هفته بارداری به بعد مقدار آماره به قرار ذیل است:

$$Z_1' = \frac{121 - 78}{\sqrt{7 \times 36.04}} = 2.707$$

چون این عدد از حد فوقانی ( $b_1 = 2.67$ ) بزرگ‌تر است، حکم به رد فرض صفر می‌نمایم.

برطرف نموده است. هرچند که به کارگیری این روش‌ها نیز مانند سایر روش‌های آماری نیازمند ممارست فراوان است. ما در قالب دو مثال نشان دادیم که این روش‌ها به کاهش چشمگیری در حجم نمونه (۱۳۲ نفر در یک مورد و ۲۶ نفر در مورد دیگر) منجر می‌گردند.

افزایش مدت بقای مبتلایان به سرطان ریه، متشکل از داروهای نظیر سیکلوفسفامید<sup>۱</sup>، آدریامایسین<sup>۲</sup>، سیس‌پلاتین<sup>۳</sup> و امیدازول کربوکسامید<sup>۴</sup> استفاده شده است (۱۰). هم‌اکنون دانشمندان جداول مبسوطی را برای کاربران معمولی این روش‌ها ارائه نموده‌اند که نیاز به دانستن ریاضیات پیشرفته را تا حد زیادی

پیوست الف: جداول مربوط به ضرایب لازم جهت تعیین حجم نمونه و آزمون پوکاک و خانواده توانی (جداول ۱، ۲ و ۳):

جدول ۱- ثابت‌های  $R_p(K, \alpha, \beta)$  برای تعیین حجم نمونه در گروه‌ها با  $K$  گروه مشاهده.

$1-\beta = 0.8$			$1-\beta = 0.9$			$1-\beta$ K
$\alpha = 0.1$	$\alpha = 0.05$	$\alpha = 0.01$	$\alpha = 0.1$	$\alpha = 0.05$	$\alpha = 0.01$	
۱	۱	۱	۱	۱	۱	۱
۱/۱۱	۱/۱	۱/۰۸۴	۱/۱۲۱	۱/۱۱۰	۱/۰۹۲	۲
۱/۱۶۶	۱/۱۵۱	۱/۱۲۵	۱/۱۸۴	۱/۱۶۶	۱/۱۳۷	۳
۱/۲۰۲	۱/۱۸۳	۱/۱۵۲	۱/۲۲۴	۱/۲۰۲	۱/۱۶۶	۴
۱/۲۲۸	۱/۲۰۷	۱/۱۷۰	۱/۲۵۴	۱/۲۲۹	۱/۱۸۷	۵
۱/۲۴۹	۱/۲۲۵	۱/۱۸۵	۱/۲۲۷	۱/۲۴۹	۱/۲۰۳	۶
۱/۲۶۶	۱/۲۳۹	۱/۱۹۷	۱/۲۹۶	۱/۲۶۵	۱/۲۱۶	۷
۱/۲۸۰	۱/۲۵۲	۱/۲۰۶	۱/۳۱۱	۱/۲۷۹	۱/۲۲۶	۸
۱/۲۹۲	۱/۲۶۲	۱/۲۱۵	۱/۳۲۵	۱/۲۹۱	۱/۲۳۶	۹
۱/۳۰۲	۱/۲۷۱	۱/۲۲۲	۱/۳۳۷	۱/۳۰۱	۱/۲۴۳	۱۰
۱/۳۱۲	۱/۲۷۹	۱/۲۲۸	۱/۳۴۸	۱/۳۱	۱/۲۵۰	۱۱
۱/۳۲۰	۱/۲۸۷	۱/۲۳۴	۱/۳۵۷	۱/۳۱۸	۱/۲۵۷	۱۲
۱/۳۴۱	۱/۳۰۵	۱/۲۴۸	۱/۳۸۱	۱/۳۳۸	۱/۲۷۲	۱۵
۱/۳۶۷	۱/۳۲۷	۱/۲۶۴	۱/۴۱۱	۱/۳۶۳	۱/۲۹۱	۲۰

1. Cyclophosphamide

2. Adriamycin

3. Cisplatin

4. Imidazole Carboxamide



جدول ۲- آزمون پوکاک:  $C_p(K, \alpha)$  برای آزمون‌های دوطرفه با  $K$  مرحله تکرار.

$\alpha = 0.1$	$\alpha = 0.05$	$\alpha = 0.01$	$K$
۱/۶۴۵	۱/۹۶	۲/۵۷۶	۱
۱/۶۴۵	۱/۱۷۸	۱/۷۷۲	۲
۱/۹۹۲	۱/۲۸۹	۲/۸۷۳	۳
۲/۰۶۷	۲/۳۶۱	۲/۹۳۹	۴
۲/۱۲۲	۲/۴۱۳	۲/۹۸۶	۵
۲/۱۶۴	۲/۴۵۳	۳/۰۲۳	۶
۲/۱۹۷	۲/۴۵۸	۳/۰۵۳	۷
۲/۲۲۵	۲/۵۱۲	۳/۰۷۸	۸
۲/۲۴۹	۲/۵۳۵	۳/۰۹۹	۹
۲/۲۷	۲/۵۵۵	۳/۱۱۷	۱۰
۲/۲۸۸	۲/۵۷۲	۳/۱۳۳	۱۱
۲/۳۰۴	۲/۵۸۸	۳/۱۴۷	۱۲
۲/۳۳۴	۲/۶۲۶	۳/۱۸۲	۱۵
۲/۳۹۲	۲/۶۷۲	۳/۲۲۵	۲۰

جدول ۳- ضرایب لازم برای آزمون خانواده توانی.

$\theta = \delta$	$\theta = \delta/2$	$\theta = 0$	$\tilde{R}$	$\tilde{C}_2$	$\tilde{C}_1$	$K$
۱۰۰/۰	۱۰۰/۰	۱۰۰/۰	۱/۰۰۰	۱/۶۴۵	۱/۶۴۵	۱
۷۶/۶	۹۰/۵	۷۶/۶	۱/۰۲۸	۱/۶۶۸	۱/۶۶۸	۲
۷۲/۲	۸۶/۰	۷۲/۲	۱/۰۵۲	۱/۶۸۷	۱/۶۸۷	۳
۶۹/۰	۸۳/۴	۶۹/۰	۱/۰۷۱	۱/۷۰۲	۱/۷۰۲	۴
۶۷/۰	۸۱/۸	۶۷/۰	۱/۰۸۵	۱/۷۱۳	۱/۷۱۳	۵
۶۳/۳	۷۸/۷	۶۳/۳	۱/۱۲۵	۱/۷۴۵	۱/۷۴۵	۱۰
۶۲/۲	۷۷/۶	۶۲/۲	۱/۱۴۵	۱/۷۶۰	۱/۷۶۰	۱۵
۶۱/۶	۷۷/۱	۶۱/۶	۱/۱۵۸	۱/۷۷۰	۱/۷۷۰	۲۰

توضیح ۱: مقادیر ثابت‌ها برای  $\theta = \delta$ ،  $1 - \beta = 0.95$ ،  $\alpha = 0.05$ ،  $\Delta = 0$  تنظیم گردیده است.

توضیح ۲: در سه ستون آخر حجم نمونه مورد انتظار به عنوان درصدی از حجم نمونه ثابت در  $\theta$  های مد نظر ارائه شده است.

پیوست ب: محاسبه حدود بحرانی در آزمون خانواده توانی.

فرض کنید که آماره  $Z_k$  عبارت است از  $Z_k = \frac{1}{\sqrt{2n\sigma^2}} \left[ \sum X_A - \sum X_B \right]$  که در آن  $X_A$ ,  $X_B$  پاسخ در دو تیمار است. خواهان آزمون فرض های  $H_1: \theta > 0, H_0: \theta = 0$  با خطای نوع اول  $\alpha$  و توان  $1 - \beta$  در  $\theta = \delta$  هستیم.

حدود تحتانی و فوقانی  $k = 1, \dots, K$  به قرار زیر هستند:

$$b_k = \tilde{C}_1(K, \alpha, \beta, \Delta) \left( \frac{k}{K} \right)^{\Delta-0.5} \quad \text{و} \quad a_k = \delta \sqrt{I_k} - \tilde{C}_2(K, \alpha, \beta, \Delta) \left( \frac{k}{K} \right)^{\Delta-0.5}$$

مقادیر  $\tilde{C}_1(K, \alpha, \beta, \Delta), \tilde{C}_2(K, \alpha, \beta, \Delta)$  به گونه ای انتخاب می شوند که ما را از رسیدن به توان و خطای مد نظر مطمئن سازند. شرط  $a_K = b_K$  ما را از خاتمه آزمون مطمئن می سازد. در این حالت سطح اطمینان نیز به تعداد ذیل می رسد.

$$I_K = \frac{\left\{ \tilde{C}_1(K, \alpha, \beta, \Delta) + \tilde{C}_2(K, \alpha, \beta, \Delta) \right\}^2}{\delta^2} \quad \text{و} \quad I_k = \frac{k}{K} \times I_K$$

#### References:

1. Pocock S J. Group sequential methods in the design and analysis of clinical trials. *Biometrika* 1977; 64:191-199.
2. Pocock SJ. *Clinical trials: a practical approach*. 6, New York: Wiley; 1992.
3. Hawang I.K, Shih WJ, Dwcani JS. Group sequential designs using a family of type one error probability spending function. *Statistics in Med* 1990; 9:1439-1445.
4. Lan KKG, DeMets DL. Discrete sequential boundaries for clinical trials. *Biometrika* 1983; 70:659-663.
5. O'Brein PC, Fleming TR. A multiple testing procedure for clinical trials. *Biometrics* 1979; 35: 549-556.
6. Wang SK, Tsiatis AA. Approximately optimal one parameter boundaries for group sequential trials. *Biometrics* 1987; 43: 193-200.
7. Pampalona S, Tsiatis AA. Group sequential designs for one-sided and two-sided hypothesis testing with provision for early stopping in favor of the null hypothesis. *J of Statistical Planning And Inference* 1994; 42: 19-35.
8. Vermillion ST, Sopher DE. Neonatal sepsis and death after multiple course of antenatal betametasone therapy. *Am J Obstet Gynecol* 2000; 183:810-814.
9. Whitehead J, Straton I. Group sequential clinical trials with triangular continuation regions. *Biometrics* 1983; 39: 227-236.

10. Fleming TR, Harrington DP, O'Brien PC. Designs for group sequential tests. *Control Clinical Trials* 1984; 5: 348-361.