



دانشکده علوم پزشکی و خدمات بهداشتی درمانی تربت جام
مجله تحقیق و توسعه سلامت
دوره ۴، شماره ۱، بهار ۱۴۰۵



چالش‌های آماری رایج در مطالعات سایتوکائینی

محمد بهرامی (MSc)^۱، فیروز امانی (PhD)^۲، مسعود حسن زاده ماکوئی (PhD)^{۳*}

نامه به سردبیر

چکیده

سایتوکائین‌ها از مهم‌ترین مولکول‌های تنظیم‌کننده پاسخ ایمنی هستند و به‌طور گسترده به‌عنوان بیومارکر در مطالعات ایمونولوژی بالینی و آزمایشگاهی مورد استفاده قرار می‌گیرند. پیشرفت روش‌های اندازه‌گیری چندپانلی امکان بررسی همزمان چندین سایتوکائین را فراهم کرده است، اما تحلیل و تفسیر آماری این داده‌ها همچنان با چالش‌هایی همراه است. هدف این نامه به سردبیر، مرور کوتاه برخی از چالش‌های آماری رایج در مطالعات سایتوکائینی است که می‌تواند بر اعتبار و تفسیر نتایج تأثیر بگذارد. از جمله این چالش‌ها می‌توان به ماهیت غیرنرمال داده‌های سایتوکائینی، استفاده گسترده از شاخص‌های توصیفی نامتناسب مانند $Mean \pm SD$ به جای Median و IQR و انجام چندین مقایسه آماری بدون اعمال اصلاح مناسب اشاره کرد. همچنین، تأکید بیش‌ازحد بر معنی‌داری آماری بدون توجه به اندازه اثر یا اهمیت بیولوژیکال یافته‌ها می‌تواند منجر به تفاسیر اغراق‌آمیز شود. در این نامه تأکید می‌شود که بسیاری از این مسائل با طراحی مناسب مطالعه، گزارش شفاف‌تر نتایج و همکاری نزدیک میان ایمونولوژیست‌ها و متخصصین آمار قابل پیشگیری هستند. توجه به این ملاحظات می‌تواند به بهبود کیفیت تحلیل‌ها، افزایش قابلیت بازتولید نتایج و ارتقای استانداردهای پژوهشی در مطالعات سایتوکائینی کمک کند.

واژه‌های کلیدی: سایتوکائین‌ها، آمار، تحلیل‌های چندمتغیره، ایمونولوژی

نویسنده مسئول: دکتر مسعود حسن‌زاده ماکوئی – آدرس E-mail: dr.m.makoui@zums.ac.ir تلفن: ۰۹۸۲۴۳۳۱۴۰۳۰۰

۱. کمیته تحقیقات دانشجویی، گروه ایمونولوژی، دانشکده پزشکی، دانشگاه علوم پزشکی زنجان، زنجان، ایران.
۲. استاد آمار زیستی، گروه پزشکی اجتماعی و آمار زیستی، دانشکده پزشکی، دانشگاه علوم پزشکی اردبیل، اردبیل، ایران.
۳. استادیار ایمنی شناسی پزشکی، گروه ایمونولوژی، دانشکده پزشکی، دانشگاه علوم پزشکی زنجان، زنجان، ایران.

پذیرش: ۱۴۰۴/۱۱/۱۰

اصلاح: ۱۴۰۴/۱۱/۰۵

دریافت: ۱۴۰۴/۱۰/۰۶

مقدمه

سایتوکاین‌ها مولکول‌های پروتئینی مهمی هستند که از طریق پیام‌رسانی بین سلولی، نقش کلیدی در تنظیم پاسخ ایمنی ایفا می‌کنند. این مولکول‌ها به‌عنوان بیومارکرهای حساس در بسیاری از زمینه‌های ایمونولوژیک، از جمله بیماری‌های التهابی، خودایمنی، عفونت‌های مزمن و سرطان، به‌طور گسترده مورد استفاده قرار گرفته‌اند. با پیشرفت روش‌های آزمایشگاهی، اندازه‌گیری همزمان سطح چندین سایتوکاین در نمونه‌های بالینی و پژوهشی رایج شده است و این امر توانایی پژوهشگران را در بررسی الگوهای پاسخ ایمنی افزایش داده است. با این وجود، تفسیر نتایج سایتوکایینی هنوز با چالش‌های آماری ملموسی همراه است، به‌ویژه از نظر انتخاب روش‌های مناسب تحلیل و گزارش صحیح داده‌ها. از آنجا که این نامه به‌سردبیر به بررسی چالش‌های آماری رایج در مطالعات سایتوکایینی می‌پردازد، تمرکز آن بر روی ویژگی‌های داده‌ها و نحوه تفسیر نتایج آماری در این مطالعات خواهد بود (به‌ویژه در شرایطی که فرضیات کلاسیک آماری ممکن است برقرار نباشند) (۱،۲).

ماهیت داده‌های سایتوکایینی و توزیع آن‌ها

در مطالعات ایمونولوژی که سطح سایتوکاین‌ها اندازه‌گیری می‌شود، این داده‌ها اغلب دارای ویژگی‌هایی هستند که آنان را از فرضیات ساده توزیع نرمال دور می‌کنند. یکی از خصوصیات بارز این داده‌ها، چولگی (skewness) زیاد است که در بسیاری از موارد حتی پس از تبدیل‌های معمول آماری نیز به‌طور کامل تصحیح نمی‌شود. در تحلیل داده‌های ایمنی‌شناختی، به‌ویژه داده‌هایی که از آزمون‌های ELISA یا روش‌های چندپانلی به‌دست می‌آیند، دیده شده که توزیع‌ها غیرنرمال و دارای outlier هستند؛ به نحوی که تحلیل نرمال ممکن است نادرست یا گمراه‌کننده باشد (۱، ۲). این ویژگی‌ها سبب می‌شود آزمون‌های پارامتریک کلاسیک که فرض نرمال بودن توزیع را دارند (مانند t-test یا ANOVA) نتایج قابل اعتماد ارائه ندهند، مگر اینکه قبل از آن بررسی‌هایی برای نرمال بودن داده انجام شود یا از روش‌های جایگزین استفاده گردد. در بسیاری از مطالعات سایتوکایینی، با وجود این موارد، داده‌ها بدون گزارش نحوه بررسی نرمال بودن توزیع و بدون اشاره به روش‌های جایگزین، با آزمون‌های پارامتریک تحلیل

می‌شوند (۱). این ویژگی آماری خاص داده‌های سایتوکایینی نه تنها انتخاب روش تحلیل را تحت‌تأثیر قرار می‌دهد، بلکه می‌تواند در تفسیر علمی نتایج و مقایسه نتایج بین مطالعات مختلف چالش‌هایی ایجاد کند؛ چنانکه پژوهشگران برای توصیف درست تفاوت‌ها و الگوهای بیولوژیک باید توزیع داده را در نظر بگیرند و در گزارش خود به آن اشاره نمایند (۲).

چالش در گزارش شاخص‌های توصیفی

در بسیاری از مطالعات سایتوکایینی، انتخاب شاخص‌های توصیفی داده‌ها نقش مهمی در درک نتایج دارد. به‌طور سنتی، پژوهش‌ها اغلب نتایج را به صورت $\text{Mean} \pm \text{SD}$ گزارش می‌کنند، زیرا این روش برای داده‌های نرمال متقارن مناسب است و از نظر تاریخی در بسیاری از رشته‌ها رایج شده است. اما داده‌های سایتوکایینی، به‌ویژه در مطالعات بالینی و آزمایشگاهی - اغلب دارای چولگی، دامنه وسیع و مقادیری دور از مرکز (outliers) دارند؛ ویژگی‌هایی که می‌توانند میانگین را به‌طور قابل‌توجهی تحت‌تأثیر قرار دهند و نمایانگر واقعی «مرکز» توزیع نباشند. در چنین مواردی، Median و IQR به عنوان شاخص‌های مقاوم‌تر می‌توانند نمای بهتری از داده ارائه دهند، چرا که نسبت به مقادیر دور از مرکز حساسیت کمتری دارند و توزیع داده را بهتر منعکس می‌کنند. در منابع آماری توصیه شده است که برای داده‌های بیولوژیکی غیرنرمال، median و IQR به جای mean و SD گزارش شوند، زیرا گزارش میانگین در چنین توزیع‌هایی می‌تواند منجر به سوءتفسیر شدت و پراکندگی پاسخ ایمنی شود. این موضوع به‌ویژه در مطالعات با حجم نمونه کم و داده‌های آزمایشگاهی که بیشترین نوسانات را دارند قابل توجه است، چرا که در این شرایط مقیاس و شکل توزیع داده از فرضیات کلاسیک دورتر می‌شود (۳، ۴).

انجام چندین مقایسه آماری در پنل‌های سایتوکایینی

یکی از ویژگی‌های رایج مطالعات سایتوکایینی این است که پژوهشگران به‌طور همزمان چندین مارکر ایمنی را اندازه‌گیری و مقایسه می‌کنند، مثلاً در پنل‌های multiplex که ده‌ها تا چندین مارکر را بررسی می‌کنند. چنین رویکردی اگرچه اطلاعات غنی‌تری از وضعیت ایمنی بدن فراهم می‌کند، اما همراه با چالشی آماری مهم است: انجام چندین آزمون همزمان احتمال خطای نوع اول (false positive) را افزایش

معنی‌دار لزوماً به معنا و پیامدهای بیولوژیکال معنی‌دار نیست (۷،۸). برای روشن‌تر شدن این چالش‌ها، می‌توان به یک سناریوی رایج در مطالعات ساینوکائینی اشاره کرد. فرض شود در یک مطالعه، سطح چند ساینوکائین التهابی از جمله IL-6 در دو گروه بیماران مبتلا به یک بیماری التهابی و افراد سالم اندازه‌گیری شده است. در چنین شرایطی، داده‌های IL-6 در گروه بیماران اغلب دارای توزیع چول و شامل مقادیر دور از مرکز هستند، به‌طوری که تعداد محدودی از نمونه‌ها مقادیر بسیار بالاتری نسبت به سایر افراد نشان می‌دهند. در صورت گزارش این داده‌ها به صورت $\text{Mean} \pm \text{SD}$ ، میانگین به‌طور قابل توجهی تحت‌تأثیر این مقادیر قرار می‌گیرد و ممکن است تصویری اغراق‌آمیز از شدت پاسخ ایمنی در کل گروه ارائه شود. در مقابل، گزارش داده‌ها به صورت Median و IQR می‌تواند توصیف دقیق‌تری از مرکز و پراکندگی واقعی توزیع فراهم کند. علاوه بر این، در همین مطالعه، انجام مقایسه‌های آماری جداگانه برای چندین ساینوکائین می‌تواند احتمال مشاهده تفاوت‌های آماری تصادفی را افزایش دهد، به‌ویژه اگر اصلاحی برای آزمون‌های متعدد در نظر گرفته نشود. حتی در مواردی که تفاوت آماری برای یک ساینوکائین مشاهده می‌شود، این اختلاف لزوماً به معنای اهمیت بیولوژیکال قابل توجه نیست و تفسیر نتایج نیازمند توجه هم‌زمان به بزرگی اثر و زمینه ایمنی‌شناختی مطالعه است. این مثال نشان می‌دهد که چگونه ویژگی‌های آماری داده‌های ساینوکائینی می‌توانند بر گزارش و تفسیر نتایج تأثیر بگذارند.

نقش همکاری ایمونولوژیست و متخصصین آمار

یکی از عوامل کلیدی در ارتقای کیفیت تحلیل داده‌ها در مطالعات ایمونولوژیک، همکاری بین‌رشته‌ای میان متخصصان ایمونولوژی و متخصصین آمار زیستی است. طراحی مناسب مطالعه، انتخاب روش‌های آماری متناسب با ویژگی‌های داده و گزارش دقیق نتایج از ابتدا تا انتها نیازمند فهم هم‌زمان مبانی بیولوژیک و روش‌های آماری است. همچنین پژوهش‌ها تأکید دارند که در گزارش نتایج، ارتباط شفاف میان تحلیل‌های آماری و مفاهیم علوم زیستی، مانند اهمیت بیولوژیکال یافته‌ها، موجب درک بهتر نتایج و پذیرش گسترده‌تر آن‌ها در جامعه علمی می‌شود. از این رو، تعامل فعال میان ایمونولوژیست‌ها و متخصصین آمار می‌تواند کیفیت

می‌دهد. در آزمون تک‌تک فرضیه‌ها با سطح خطای α (مثلاً ۰/۰۵)، اگر تعداد زیادی آزمون مستقل انجام شود، احتمال این که حداقل یکی از نتایج به‌طور تصادفی معنی‌دار به‌نظر برسد افزایش می‌یابد. در واقع با افزایش تعداد آزمون‌ها، احتمال رخداد خطای نوع اول در کل مجموعه داده‌ها بالا می‌رود که می‌تواند منجر به نتایج مثبت کاذب شود. برای حل این مسئله، روش‌های اصلاح چندگانه مانند Bonferroni یا کنترل نرخ کشف غلط (FDR) پیشنهاد شده‌اند، که با تنظیم آستانه معنی‌داری به‌گونه‌ای تلاش می‌کنند که احتمال اشتباهات تصادفی کاهش یابد. در مطالعات بیولوژیک و ایمونولوژیک، توجه به این اصلاحات اهمیت دارد، زیرا عدم استفاده از آن‌ها می‌تواند باعث افزایش گزارش تفاسیر نادرست درباره تفاوت‌های ساینوکائینی شود و نتایج را به‌صورت غیرقابل‌اطمینان نشان دهد. بر اساس بررسی‌های آماری در علوم زیستی، استفاده از استراتژی‌های اصلاح چندگانه هنگام انجام تعداد زیادی آزمون برای کنترل خطر کشف‌های غلط توصیه می‌شود و نقش مهمی در بهبود اعتبار نتایج دارد (۵،۶).

تفسیر بیش‌ازحد معنی‌داری آماری

در پژوهش‌های علمی، معنی‌داری آماری (statistical significance) که معمولاً با P-value گزارش می‌شود، تنها نشان‌دهنده‌ی این است که احتمال رخداد تفاوت مشاهده‌شده به‌صورت تصادفی در چارچوب فرض صفر کم بوده است و به‌طور مستقیم به اهمیت یا بزرگی اثر اشاره نمی‌کند. به عبارت دیگر، یک تفاوت در سطوح ساینوکائین‌ها می‌تواند از نظر آماری معنی‌دار باشد اما از نظر زیستی یا بیولوژیکال اهمیت کمی داشته باشد؛ به‌ویژه زمانی که حجم نمونه بزرگ باشد، حتی اختلاف‌های کوچک می‌توانند P معنی‌دار بدهند بدون اینکه اثر واقعی زیادی داشته باشند. از سوی دیگر، در مطالعات با حجم نمونه کم ممکن است اختلاف‌های بیولوژیکال مهم از نظر آماری معنی‌دار نشوند. بنابراین، تکیه صرف بر معنی‌داری آماری می‌تواند منجر به تفاسیر اغراق‌آمیز از نتایج شود، مگر اینکه همراه با معیارهایی مانند اثر اندازه (effect size) و بحث درباره‌ی اهمیت بالینی/بیولوژیکال گزارش و تفسیر شود. در این زمینه مقالات برجسته تأکید کرده‌اند که تعبیر نتایج باید فراتر از مقدار P باشد و شامل ارزیابی مقدار و اهمیت واقعی تغییر نیز باشد، زیرا آماری

پاسخ ایمنی شود. توجه به این مسائل نه به عنوان نقد فردی، بلکه در راستای بهبود کیفیت پژوهش‌های ایمونولوژیک ضروری است. در این مسیر، طراحی دقیق مطالعه و همکاری مؤثر میان ایمونولوژیست‌ها و متخصصین آمار زیستی می‌تواند به افزایش شفافیت گزارش، اعتبار نتایج و قابلیت بازتولید مطالعات کمک کند. پرداختن آگاهانه به این چالش‌ها گامی مهم در جهت ارتقای استانداردهای تحلیلی در مطالعات سائتوکایینی خواهد بود.

سپاسگزاری

موردی ندارد.

تضاد منافع

وجود ندارد.

تحلیلی، شفافیت گزارش و قابلیت بازتولید نتایج را به طور قابل توجهی افزایش دهد و فضای علمی را به سمت استانداردهای بالاتر سوق دهد (۹، ۱۰).

نتیجه گیری

مطالعات سائتوکایینی نقش اساسی در درک سازوکارهای ایمنی و شناسایی بیومارکرهای مرتبط با بیماری‌های التهابی، خودایمنی، عفونی و بدخیمی‌ها دارند. با این حال، ویژگی‌های خاص داده‌های سائتوکایینی، از جمله توزیع‌های غیرنرمال، نوسانات بالا و بررسی همزمان چندین مارکر، تحلیل و تفسیر آماری این داده‌ها را با چالش‌هایی همراه می‌سازد. انتخاب نامناسب شاخص‌های توصیفی، انجام چندین مقایسه آماری بدون اصلاح مناسب و تفسیر بیش از حد معنی‌داری آماری می‌تواند منجر به برداشت‌های نادرست از شدت یا اهمیت

References

- McGuinness D, Bennett S, Riley E. Statistical analysis of highly skewed immune response data. *J Immunol Methods*. 1997;201(1):99-114.
- Ballenberger N, Lluís A, von Mutius E, Illi S, Schaub B. Novel Statistical Approaches for Non-Normal Censored Immunological Data: Analysis of Cytokine and Gene Expression Data. *PLOS ONE*. 2012;7(10):e46423.
- Whitley E, Ball J. Statistics review 1: presenting and summarising data. *Crit Care*. 2002;6(1):66-71.
- Habibzadeh F. How to report the results of public health research. *Journal of Public Health and Emergency*. 2017;1.
- Pfister IB, Zandi S, Gerhardt C, Spindler J, Reichen N, Garweg JG. Risks and Challenges in Interpreting Simultaneous Analyses of Multiple Cytokines. *Transl Vis Sci Technol*. 2020;9(7):27.
- Menyhart O, Weltz B, Györfy B. MultipleTesting.com: A tool for life science researchers for multiple hypothesis testing correction. *PLOS ONE*. 2021;16(6):e0245824.
- Sharma H. Statistical significance or clinical significance? A researcher's dilemma for appropriate interpretation of research results. *Saudi J Anaesth*. 2021;15(4):431-4.
- Lovell DP. Biological Importance and Statistical Significance. *Journal of Agricultural and Food Chemistry*. 2013;61(35):8340-8.
- Pomann GM, Boulware LE, Cayetano SM, Desai M, Enders FT, Gallis JA, et al. Methods for training collaborative biostatisticians. *J Clin Transl Sci*. 2020;5(1):e26.
- Simkus A, Coolen-Maturi T, Coolen FPA, Bendtsen C. Statistical Perspectives on Reproducibility: Definitions and Challenges. *Journal of Statistical Theory and Practice*. 2025;19(3):40



Common Statistical Challenges in Cytokine Studies

Mohammad Bahrami (MSc)¹, Firouz Amani (PhD)², Masoud Hassanzadeh Makoui (PhD)³

Letter to Editor

Abstract

Cytokines are among the most important regulators of immune responses and are widely used as biomarkers in both clinical and experimental immunology studies. Advances in multiplex measurement techniques have enabled the simultaneous assessment of multiple cytokines; however, the statistical analysis and interpretation of these data remain challenging. This letter provides a overview of common statistical challenges encountered in cytokine studies that may affect the validity and interpretation of results. These challenges include the non-normal distribution of cytokine data, the widespread use of inappropriate descriptive measures such as mean $\pm$ standard deviation instead of median and interquartile range, and the performance of multiple statistical comparisons without adequate adjustment. In addition, an overemphasis on statistical significance without consideration of effect size or biological relevance may lead to exaggerated interpretations of findings. This letter highlights that many of these issues can be addressed through appropriate study design, more transparent reporting of results, and close collaboration between immunologists and statisticians. Greater attention to these considerations may improve the quality of statistical analyses, enhance the reproducibility of findings, and raise research standards in cytokine studies.

Keywords: Cytokines, Biostatistics, Multivariate Analysis, Immunology

***Corresponding Author:** Dr. Masoud Hassanzadeh Makoui - E-mail address: dr.m.makoui@zums.ac.ir
Tell: +982433140300

1. Student Research Committee, Department of Immunology, School of Medicine, Zanjan University of Medical Sciences, Zanjan, Iran
2. Department of Social Medicine and Biostatistics, Faculty of Medicine, Ardabil University of Medical Sciences, Ardabil, Iran.
3. Department of Immunology, School of Medicine, Zanjan University of Medical Sciences, Zanjan, Iran.

Received: 2025/12/27

Revised: 2026/01/25

Accepted: 2026/01/25