

# یک روش جدید برای تصحیح سوگرایی تاییدی در مطالعات بررسی صحت تست‌های تشخیصی با استفاده از رویکرد بیزین

محمد چهرازی<sup>۱</sup>، منصور شمسی پور<sup>۲</sup>، مهدی نوروزی<sup>۳</sup>، فرهاد جعفری<sup>۴</sup>، فریبا رضانعلی<sup>۵</sup>

<sup>۱</sup> دانشجوی دکتری آمار زیستی، پژوهشگاه رویان جهاد دانشگاهی، مرکز تحقیقات اپیدمیولوژی باروری، گروه اپیدمیولوژی و سلامت باروری، تهران، ایران

<sup>۲</sup> دانشجوی دکتری اپیدمیولوژی، پژوهشگاه رویان جهاد دانشگاهی، مرکز تحقیقات اپیدمیولوژی باروری، گروه اپیدمیولوژی و سلامت باروری تهران، ایران

<sup>۳</sup> کارشناس ارشد اپیدمیولوژی، عضو هیئت علمی دانشکده بهداشت، دانشگاه علوم پزشکی قم، ایران

<sup>۴</sup> استادیار پزشکی اجتماعی، دانشکده پزشکی، دانشگاه شاهد، تهران، ایران

<sup>۵</sup> متخصص زنان و زایمان، پژوهشگاه رویان، پژوهشکده زیست‌شناسی و علوم پزشکی تولید مثل جهاد دانشگاهی، مرکز تحقیقات پزشکی تولید مثل، گروه اندوکرینولوژی و ناباروری زنان، تهران، ایران

نویسنده رابط: منصور شمسی پور، نشانی: تهران، خیابان رسالت، خیابان بنی‌هاشم، میدان بنی‌هاشم، پژوهشگاه رویان، گروه اپیدمیولوژی و سلامت باروری تلفن: ۲۲۳۳۹۹۲۹،

پست الکترونیک: shamsman87@yahoo.com

تاریخ دریافت: ۱۳۹۰/۵/۹؛ پذیرش: ۱۳۹۰/۱۱/۸

**مقدمه و اهداف:** یکی از سوگرایی‌هایی که در مطالعات بررسی صحت آزمون‌های تشخیصی رخ می‌دهد، سوگرایی تاییدی می‌باشد. این

سوگرایی زمانی رخ می‌دهد که تست استاندارد طلایی از افرادی که تست تشخیصی برای آن‌ها صورت گرفته است، انجام می‌شود.

**روش کار:** بیمارانی که حداقل دوبرار متوالی سیکل IVM ICSI را از دست داده بودند در این مدل وارد شدند. بیماران با روش اولترا سونوگرافی غربالگری شدند و بیمارانی که پولیپ داشتند برای زهدان برداری توصیه شدند روش رگرسیون لجستیک با پیامد دو گانه مناسب برای مقادیر گمشده (منفی واقعی و منفی کاذب) برای تعیین حساسیت و ویژگی انتخاب شد. روش‌های بیزین قبل از اطلاع به وجود پولیپ به کار گرفته شد و منفی‌های واقعی و کاذب در چارچوب بیزی تخمین زده شد.

**نتایج:** جمعا ۲۳۸ بیمار غربالگری شدند که ۴۷ نفر آن‌ها دارای پولیپ بودند. به دارندگان پولیپ توصیه شد که زهدان برداری کنند که بر ۴۷ نفر این کار را انجام دادند و ۳۷ نفر آن‌ها از نظر داشتن پولیپ تایید شدند. هیچیک از ۱۹۱ بیمار که پولیپ نداشتند برداری نشدند. نتیجه بررسی نشان داد که ۱۴ مورد منفی کاذب و ۱۷۷ منفی واقعی بودند، در نتیجه حساسیت و ویژگی به آسانی پس از تخمین موارد یافت نشده محاسبه شد. میزان حساسی و ویژگی به ترتیب ۷۴٪ و ۹۴٪ بود.

**نتیجه‌گیری:** تحلیل بیزی به همراه اطلاعات کافی ابزار بسیار قوی در شبیه‌سازی فضای آزمایشی می‌باشد.

**واژگان کلیدی:** سوگرایی تاییدی، روش‌های بیزین، پولیپ

## مقدمه

تشخیص بیماری مورد نظر را معمولا به عنوان استاندارد طلایی در نظر می‌گیرند. در بهترین حالت برای ارزیابی یک تست تشخیصی هر دو تست برای همه بیماران انجام می‌پذیرند. یکی از مشکلاتی که در این مطالعات وجود دارد این است که به دلایلی مانند بالا بودن هزینه، تهاجمی و غیر اخلاقی بودن تست استاندارد و یا استفاده از داده‌هایی که به صورت گذشته نگر گردآوری شده‌اند، تست استاندارد طلایی برای همه افراد صورت نمی‌گیرد. این مساله در کلینیک منطقی به نظر می‌رسد اما در ارزیابی تست‌های تشخیصی باعث ایجاد سوگرایی بنام سوگرایی تاییدی می‌شود. این سوگرایی یک نوع سوگرایی انتخاب است و زمانی اتفاق می‌افتد که

پیشرفت علوم پزشکی و تکنولوژی، روش‌های تشخیصی گوناگونی را برای شناسایی بیماران در اختیار پزشکان قرار داده است. مطالعات بررسی صحت تست‌های تشخیصی از جمله مطالعاتی هستند که با بررسی خصوصیات (حساسیت، ویژگی) این تست‌ها به پزشکان در انتخاب مناسب‌ترین تست در ارزیابی وضعیت بالینی افراد و تصمیم‌گیری در مورد روش درمان آن‌ها کمک می‌کنند. برای ارزیابی تست‌های تشخیصی و محاسبه حساسیت و ویژگی این تست‌ها نتایج بدست آمده از تست مورد ارزیابی را با نتایج حاصل از تست استاندارد طلایی که بر روی همان بیماران صورت گرفته است، مقایسه می‌کنند. بهترین تست قابل دسترسی برای

جدیدی به نام  $v$  تعریف شده است که بیانگر فرایند گمشدگی می‌باشد. ما در این مقاله از دیدگاهی دیگری به مساله پرداخته ایم که برخی از مشکلات روش‌های قبلی را ندارد. ما در این روش با استفاده از تعاریف جدید، متغیر  $v$  را حذف کرده و یک مدل جدید بیزی ارائه دادیم به این صورت که با در نظر گرفتن هیستروسکوپی به عنوان استاندارد طلایی حساسیت و ویژگی سونوگرافی واژینال را در تشخیص پولیپ، به روش فراوانی گرا محاسبه کردیم. و سپس با در نظر گرفتن اینکه فقط بیمارانی هیستروسکوپی شوند که نتیجه سونوگرافی واژینال مثبت داشته باشند با یک مدل بیزی حساسیت و ویژگی را محاسبه کردیم و آنرا با حساسیت و ویژگی داده‌های واقعی مقایسه نمودیم که نتایج آن در جداول مربوطه آمده‌اند. بخش‌های بعدی به ترتیب شامل معرفی مدل سپس معرفی داده‌ها و کاربرد مدل در این داده‌ها شرح داده می‌شود. در پایان نتایج و بحث مقاله ارائه می‌شوند.

### روش کار

فرض می‌کنیم که فرایند گمشدن همزمان با نتیجه آزمون تشخیصی رخ می‌دهد (۷). به عبارت دیگر اگر نتیجه تست فردی مثبت شود برای این فرد آزمون طلایی انجام می‌شود ولی اگر منفی شود آزمون طلایی برای این فرد انجام نمی‌شود. پس نتیجه آزمون دوم برای افرادی که نتیجه تست تشخیصی آن‌ها منفی است از دست می‌رود. این فرایند در زیر ترسیم شده است:

$$\text{Test Result} \begin{cases} \text{positive} \{ \text{Gold standard} \\ \text{negative} \{ \text{Miss} \end{cases}$$

پس احتمال گمشدن فردی که نتیجه تست آن مثبت است صفر و برای فردی که منفی است ۱ می‌باشد. اگر  $T$  نتیجه تست تشخیصی و  $V$  نشان دهنده فرایند گمشدگی باشد آنگاه داریم:

$$P(V|T) = \begin{cases} 1 & V=T \\ 0 & V \neq T \end{cases} \quad (1)$$

$$P(V|T) = \frac{P(V,T)}{P(T)} = 1 \xrightarrow{\text{yields}} P(V,T) = P(T) \quad (2)$$

حال فرض می‌کنیم که  $G$  وضعیت تست طلایی فرد را نشان می‌دهد آنگاه هدف یافتن تابعی است که وضعیت تست طلایی افرادی است که نتیجه تست تشخیصی آن‌ها منفی شده است را پیش بینی کند. بدین ترتیب می‌توان حساسیت و ویژگی را برآورد کرد. برای اینکار احتمال زیر را تعریف می‌کنیم:

$$P(G|T, V) = \frac{P(G, T, V)}{P(T, V)} = \frac{P(T)P(G|T)P(V|G, T)}{P(T, V)} \quad (4)$$

تست استاندارد طلایی فقط برای زیر مجموعه‌ای غیر گویا از افرادی که تست تشخیصی برایشان صورت گرفته است. انجام می‌شود و انتخاب افراد برای انجام استاندارد طلایی به نتیجه تست تشخیصی و/یا به متغیری مرتبط با وضعیت بیماری بستگی دارد. معمولاً این سوگرایی را به دو نوع تقسیم‌بندی می‌کنند سوگرایی تاییدی افتراقی و سوگرایی تاییدی جزئی، سوگرایی تاییدی افتراقی زمانی ایجاد می‌شود که استاندارد طلایی که برای افرادی که نتیجه تست تشخیصی مثبت داشته‌اند با استاندارد طلایی که نتیجه تست تشخیصی منفی دارند، متفاوت باشد. سوگرایی تاییدی جزئی زمانی اتفاق می‌افتد که تست استاندارد طلایی فقط برای افرادی صورت می‌گیرد که نتیجه تست تشخیصی مثبت داشته‌اند (۱،۲). بنابراین این سوگرایی ما را در محاسبه حساسیت و ویژگی برای تست‌های تشخیصی دچار خطا می‌کند. در اکثر مطالعات به سوگیری تاییدی در هر نوع آن، از دیدگاه فرایند داده‌های گمشده پرداخته شده است. به این صورت که برای تعدادی از افراد نتیجه تست طلایی وجود ندارد. مدل‌هایی که بر پایه این دیدگاه ساخته می‌شوند، وابسته به فرایند گمشدگی داده‌ها می‌باشد. بگ و گرینز از نخستین محققانی بودند که به تصحیح سوگرایی تاییدی جزئی پرداختند (۳). آن‌ها از نسبت‌های مشاهده شده بیماری و غیر بیماری در بین افرادی که تست استاندارد طلایی انجام داده‌اند برای محاسبه تعداد مورد انتظار بیمار و غیر بیمار در بین افرادی که تست استاندارد طلایی را انجام نداده‌اند، استفاده کردند. آن‌ها پیش فرضی بنام استقلال شرطی برای مدل در نظر گرفتند بدین صورت که فرایند گمشدن یا انتخاب افراد برای تست طلایی فقط به نتیجه تست بستگی دارد و به وضعیت واقعی بیماری فرد بستگی ندارد. به عبارت دیگر رخداد گمشدن افراد به شرط نتیجه تست از وضعیت واقعی بیماری فرد مستقل است در عمل این پیش فرض غیر واقعی بوده و برآوردهای این دو محقق در بعضی از موارد می‌توانند همراه کننده باشند (۴،۵). اما کاسینسکی (۶) برای غلبه بر این مشکل پیش فرض مدل را اینگونه قرار داد که فرایند گمشدن علاوه بر نتیجه تست، به وضعیت واقعی بیمار و اطلاعات مشاهده نشده در ارتباط با بیماری هم بستگی دارد (۵،۷) بدیهی است که با داشتن داده‌های کامل می‌توان مدل‌های متفاوت را از نظر توانایی در پیش‌بینی نتیجه تست طلایی مقایسه کرد. برآورد مدل‌های ذکر شده بر اساس دیدگاه فراوانی گرا و با استفاده از روش حداکثر درست‌نمایی بوده است مشکلات موجود در این روش‌ها در مقاله مارتینز و بوزویانو شرح داده شده‌اند. به نظر می‌رسد مشکلات روش‌های قبلی ناشی از این باشد که متغیر

$$y_{ij} \sim \text{Binomial}(P_{ij}, n_i), \quad \sum_i n_i = N$$

$$P_{11} + P_{12} = 1$$

$$P_{21} + P_{22} = 1$$

$$\text{logit}(P_{ij}) = \beta_0 + \beta_{ij} T_i$$

که در این مدل  $T_i$  یک متغیر نشانگر است که نشان دهنده وضعیت  $i$ ام تست تشخیصی می‌باشد. مثلا  $T_1$  یعنی افرادی که نتیجه تست تشخیصی آن‌ها مثبت می‌باشد.  $\beta_0$  و  $\beta_{ij}$  به ترتیب نشان دهنده عرض از مبدا مدل و اثر متغیر پیش‌گو بر روی نتیجه تست طلایی می‌باشد. در این مدل رگرسیونی نسبت شانس را می‌توان بر حسب احتمال‌های تعریف شده در بخش قبل نوشت. بدین صورت که:

$$\exp(\beta_{11}) - \exp(\beta_{22}) - \frac{\text{SN} \cdot \text{SP}}{(1 - \text{SN}) \cdot (1 - \text{SP})} = \frac{\text{LR}^+}{\text{LR}^-}$$

$$\exp(\beta_{21}) = \exp(\beta_{12}) = \frac{(1 - \text{SN}) \cdot (1 - \text{SP})}{\text{SN} \cdot \text{SP}} = \frac{\text{LR}^-}{\text{LR}^+}$$

$$\text{SP} = \frac{\text{NPV} \cdot n \cdot P(G = 1) - y_{11}}{(1 - \text{NPV}) \cdot n \cdot (1 - P(G = 1))}$$

$$\text{SN} = \frac{\text{PPV} \cdot y_{12}}{(1 - \text{PPV}) \cdot n \cdot P(G = 1)}$$

$$\text{NPV} = \frac{\text{PPV} \cdot y_{12} \cdot (n \cdot P(G = 0) - y_{12})}{(1 - \text{PPV}) \cdot n \cdot P(G = 1) + \text{PPV} \cdot y_{12} \cdot (n \cdot P(G = 0) - y_{12}) - \text{PPV} \cdot y_{12} \cdot y_{11}}$$

$$P(G = 0) = 1 - P(G = 1)$$

چگونگی بدست آمدن این عبارات در قسمت پیوست ۱ توضیح داده شده است. بدین ترتیب توانستیم که ضرایب رگرسیونی مدل را بر حسب حساسیت و ویژگی بدست بیاوریم.

#### توزیع‌های پیشین

برای پارامترهای مورد نظر در مدل بیزی بدلیل نبودن اطلاعات لازم از پیشین‌های ناآگاهی بخش استفاده شد. بدین ترتیب برای عرض از مبدا مدل رگرسیون لوجستیک از توزیع نرمال با میانگین صفر و واریانس زیاد ۱۰۰۰ استفاده شد.

$$\beta_0 \sim N(0, 1000)$$

بنا به تعریف (۲) احتمال (۳) بفرم زیر درمی‌آید:

$$P(G|T, V) = \frac{P(T)P(G|T)P(V|G, T)}{P(T)} = P(G|T)P(V|G, T) \quad (5)$$

همان‌طور که در فرمول (۱) عنوان کردیم  $T$  و  $V$  همزمان رخ می‌دهند پس اطلاع از رخداد  $T, G$  درباره رخداد  $V$  نیز اطلاع کامل می‌دهد. پس مقدار  $P(V|G, T)$  برابر ۱ است یعنی

$$P(G|T, V) = P(G|T)$$

می‌توان نتیجه گرفت که فرآیند گمشدگی بشرط اطلاع از نتیجه تست تشخیصی، از وضعیت واقعی بیماری افراد مستقل است (اگر فرض کنیم که نتیجه تست طلایی وضعیت واقعی بیماری افراد را مشخص می‌کند) ولی از طرفی وضعیت واقعی بیماری افراد تاثیر مستقیم بر فرآیند گمشدگی دارد. مثلا فردی که واقعا بیمار است احتمال مثبت شدن نتیجه تست تشخیصی و در نتیجه احتمال گمشدن آن کمتر است. حال احتمال  $P(G|T)$  را برای حالات  $G, T$  بصورت زیر بدست می‌آید.

$$P(G^+|T^+) = \frac{\text{SN}P(G^+)}{\text{SN}P(G^+) + (1 - \text{SP})P(G^-)} = P_{11}$$

$$P(G^+|T^-) = \frac{(1 - \text{SN})P(G^+)}{\text{SP}P(G^-) + (1 - \text{SN})P(G^+)} = P_{21}$$

$$P(G^-|T^+) = \frac{(1 - \text{SP})P(G^-)}{\text{SN}P(G^+) + (1 - \text{SP})P(G^-)} = P_{12}$$

$$P(G^-|T^-) = \frac{\text{SP}P(G^-)}{(1 - \text{SN})P(G^+) + (\text{SP})P(G^-)} = P_{22}$$

در این فرمول‌ها  $\text{SN}$  یعنی حساسیت و  $\text{SP}$  نشان دهنده ویژگی می‌باشد. حال بر اساس این احتمال‌ها یک مدل رگرسیون لوجستیک با یک متغیر مستقل می‌سازیم. متغیر پیش‌گو را نتیجه تست تشخیصی قرار می‌دهیم. بدین ترتیب از توزیع احتمالی که در تقسیم افراد در دو گروه نتیجه تست طلایی استفاده کرده و کلیه افراد را به دو گروه تقسیم می‌کنیم.

#### مدل بیزی

در این قسمت با استفاده از یک مدل رگرسیونی لوجستیک با پیامد دو جمله‌ای برای پیش‌بینی مقادیر گمشده می‌پردازیم. اگر  $y_{ij}$  نشان دهنده تعداد افراد  $i$ امین موقعیت در یک جدول  $2 \times 2$  باشد و  $n_i$  تعداد افراد با نتیجه تست مثبت یا منفی (بسته به اندیس  $i$ ) باشد، آنگاه مدل زیر را می‌توان بنا به تعاریف ذکر شده بیان کرد:

که تمامی زنان قبل از درمان IVF تحت بررسی تشخیصی با هیستروسکوپی قرار بگیرند (۱۳) اما با توجه به اینکه هیستروسکوپی یک روش تهاجمی می‌باشد ممکن است همه بیمارانی که به روش سونوگرافی واژینال ارزیابی می‌شوند تحت هیستروسکوپی قرار نگرفته و فقط بیمارانی که نتیجه سونوگرافی واژینال مثبت داشته‌اند، هیستروسکوپی شوند در نتیجه سوگیری تاییدی رخ می‌دهد. نخست ما اطلاعات مربوط به ۲۳۸ بیمار مراجعه کننده به پژوهشگاه رویان را که حداقل دو بار شکست مکرر در سیکل‌های IVF/ICSI داشته‌اند و تحت هیستروسکوپی و سونوگرافی قرار گرفته‌اند، آنالیز می‌کنیم. نتایج این دو تست در جدول شماره ۱ نشان داده شده است.

جدول شماره ۱- نتایج تست تشخیصی و تست استاندارد

	<b>G<sup>+</sup></b>	<b>G<sup>-</sup></b>
<b>T<sup>+</sup></b>	<b>37</b>	<b>10</b>
<b>T<sup>-</sup></b>	<b>13</b>	<b>178</b>

با در نظر گرفتن هیستروسکوپی به عنوان استاندارد طلایی حساسیت و ویژگی سونوگرافی واژینال را در تشخیص پولیپ، به روش فراوانی گرا محاسبه کردیم.

$$SN = \frac{37}{50} = 0.74$$

$$SP = \frac{178}{188} = 0.946$$

$$PPV = \frac{37}{47} = 0.787$$

$$NPV = \frac{178}{191} = 0.932$$

و سپس با در نظر گرفتن اینکه فقط بیمارانی هیستروسکوپی شوند که نتیجه سونوگرافی واژینال مثبت داشته باشند با یک مدل بیزی حساسیت و ویژگی را محاسبه کردیم و آنرا با حساسیت و ویژگی داده‌های واقعی مقایسه نمودیم. مشکل این وضعیت این است که برای افرادی که نتیجه تست تشخیصی آن‌ها منفی است نتیجه تست طلایی به دلایل متعددی از جمله دلایل اخلاقی وجود ندارد. پس در جدول شماره ۲ به فرم زیر می‌شود.

جدول شماره ۲- مکانیسم گم شدگی داده‌ها

	<b>G<sup>+</sup></b>	<b>G<sup>-</sup></b>
<b>T<sup>+</sup></b>	<b>37</b>	<b>10</b>
<b>T<sup>-</sup></b>	<b>NA</b>	<b>NA</b>

NA: Not Available

اما برای شیوع بیماری در جامعه از یک توزیع آگاهی بخش بتا استفاده می‌کنیم. برای بدست آوردن فوق پارامترها از روش ذهنی صدکی استفاده شد (۸).

$$p \sim \text{beta}(a, b)$$

بدین صورت که از دو متخصص زنان به طور مستقل خواسته شد تا صدک‌های ۵ و ۹۵ را درباره میزان شیوع بیماری ارائه دهند. سپس میانگین صدک‌های آن‌ها را در نظر می‌گیریم. از هر متخصص خواسته شد تا بهترین نظر خود را بدون در نظر گرفتن متون، درباره حد بالا و پایین شیوع پولیپ در جامعه بیماران مورد نظر بیان کند. نتایج این بررسی پس از میانگین گرفتن از صدک‌ها به صورت زیر بدست آمد:

$$P(p < 0.33) = 0.95$$

$$P(p < 0.001) = 0.05$$

بدین ترتیب با استفاده از این دو معادله مقادیر فوق پارامترها به راحتی بدست می‌آیند. پس توزیع پیشین شیوع به صورت زیر می‌شود:

$$p \sim \text{beta}(0.59, 5.61)$$

بدین ترتیب با استفاده از قضیه بیز می‌توان تابع چگالی پسین را بدست آورد.

$$P(SN, SP, \beta_0 | y) \propto P(y | SN, SP, \beta_0) P(SN, SP, \beta_0)$$

#### مثال با داده‌های واقعی و کاربرد مدل

شکست مکرر IVF یکی از معضلات مراکز نازایی می‌باشد. ناهنجاری‌های ساختمانی در حفره رحم نظیر فیروئیدها و ناهنجاری‌های مولرین می‌توانند نقش مهمی در شکست لانه‌گزینی جنین در طی سیکل‌های IVF ایفا نمایند (۹) اصلاح پاتولوژی‌های حفره رحمی به عنوان یک اقدام درمانی جهت بهبود نتایج سیکل‌های ART در این افراد پیشنهاد شده است. هیستروسکوپی یک روش ارزشمند و به عنوان استاندارد طلایی جهت ارزیابی حفره رحم در بیماران نازا می‌باشد. این روش امکان دید مستقیم حفره رحم و سرویکس را فراهم نموده و دقت و صحت تشخیص را افزایش می‌دهد (۱۰). ناهنجاری‌های رحمی به خوبی با هیستروسکوپی قابل تشخیص می‌باشند (۱۱) نتایج تعدادی از مطالعات حساسیت، ویژگی و ارزش پیشگویی منفی و مثبت سونوگرافی واژینال را مشابه هیستروسکوپی گزارش کرده‌اند (۹) برخی از مطالعات نیز عدم ارتباط قوی بین سونوگرافی واژینال استاندارد و یافته‌های هیستروسکوپی را گزارش کرده‌اند (۱۲). بنابراین پیشنهاد می‌گردد

پسین قابل برآورد هستند. این شبیه سازی با استفاده از زنجیر مارکوف مونت کارلویی و با الگوریتم متروپولیس بلوکی تطبیقی انجام شده است (۱۴). این الگوریتم در نرم افزار OpenBUGS 3.1.2 اجرا شده است (۱۵). بعد از ۱۰۰۰۰ بار تکرار الگوریتم، زنجیر به توزیع پسین مورد نظر همگرا شد و ۵۰۰۰ نمونه اول به عنوان دوره سوخته کنار گذاشته شد. برای بررسی همگرایی زنجیر از روش‌های گرافیکی مثل نمودار اثر، نمودار خود همبستگی و نمودار تابع چگالی پسین استفاده شد (۱۶). دستور نوشته شده در محیط Openbugs در ضمیمه ۲ آورده شده است. نتایج روش بیزی بعد از نمونه‌گیری (شبیه‌سازی) از تابع پسین در جدول شماره ۳ آورده شده است.

برای حل این مشکل چندین راه وجود دارد از جمله اینکه فرض می‌کنیم که نتیجه تست طلایی تمامی این افراد نیز منفی است که این کار از نظر منطقی صحیح نیست. چون می‌دانیم که احتمال خطا در نتیجه تست وجود دارد و ممکن است ارزش اخباری منفی ۱ نباشد. با اینکار مقدار حساسیت و ویژگی به شدت افزایش یافته و در برآورد آن‌ها دچار بیش برآورد می‌شویم. روش دیگر اینکه نتیجه تست طلایی افرادی را که در آزمایش‌های قبلی بررسی شده‌اند جایگزین شوند که در این روش نیز برآوردها می‌توانند دچار اربیبی شوند زیرا شرایط حاکم بر آزمایشات قبلی ممکن است با شرایط فعلی متفاوت باشد. در روش بیزی، مدل (۱) تابع پسین حساسیت، ویژگی و عرض از مبدا مدل رگرسیون لوژستیک می‌باشد. پارامترهای مدل (۱) با استفاده از شبیه‌سازی توزیع

جدول شماره ۳- برآورد پارامترهای شبیه سازی شده با الگوریتم MCMC در چهارچوب بیزی

پایان	شروع	حد بالا اطمینان	حد پایین اطمینان	انحراف معیار	میانگین پسین	پارامتر
۱۰۰۰۰	۵۰۰۱	۰/۹۲	۰/۵۱	۰/۱۱	۰/۷۴	حساسیت
۱۰۰۰۰	۵۰۰۱	۰/۹۸	۰/۸۳	۰/۰۴	۰/۹۲	ویژگی
۱۰۰۰۰	۵۰۰۱	۰/۸۸	۰/۶۵	۰/۰۶	۰/۷۸	ارزش اخباری مثبت
۱۰۰۰۰	۵۰۰۱	۰/۹۸	۰/۸۳	۰/۰۴	۰/۹۲	ارزش اخباری منفی

همانطور که مشاهده می‌کنید برآوردها تقریباً برابر زمانی است که تمام داده در دسترس می‌باشد.

## بحث

کرد. از آنجایی که فقط توزیع احتمالی در دو گروه تست طلایی برای نتیجه تست تشخیصی مثبت کاملاً مشخص شده است پس مدل رگرسیونی را باید بر پایه این توزیع احتمالی در نظر گرفت یعنی نتیجه تست تشخیصی را باید به عنوان یک متغیر پیشگو برای وضعیت واقعی بیماری در نظر گرفت. از این طریق می‌توان به همین توزیع احتمالی در گروه تست منفی نیز پی برد. به عبارت دیگر روند احتمالی که در گروه تست مثبت وجود دارد الگویی برای پیش‌بینی توزیع احتمالی تست منفی است. در هر دو گروه تست مثبت و منفی از یک تست تشخیصی یکسان و در شرایط آزمایش یکسان استفاده شده است پس انتظار داریم که روند احتمالی در دو گروه مشابه باشد اما از آنجایی که افرادی که گمشده اند با احتمال بیشتری سالم هستند پس روند احتمالی در این گروه برعکس گروه تست مثبت است. برای حل این مشکل فرض کردیم که در جدول شماره ۱ جای  $\gamma_{21}$  و  $\gamma_{22}$  را عوض می‌کنیم که روند احتمالی موجود در گروه تست مثبت حفظ شود. برای بهتر شدن مدل و پیش‌بینی بهتر مقادیر گمشده می‌بایست اطلاعاتی را خارج از مطالعه و بر اساس نظر یک کارشناس وارد

ما در این مقاله یک مدل آماری جدید بر اساس احتمال‌های شرطی ارائه کردیم که پارامترهای این مدل یعنی حساسیت و ویژگی را از دیدگاه بیزی (۱۷) بدست آوردیم. این احتمال‌های شرطی در واقع همان ارزش اخباری مثبت و منفی هستند که اطلاعات مدل را در خود جای داده‌اند. فرم گمشدن داده‌ها در این مقاله با سایر مقالات متفاوت می‌باشد به طوریکه کلیه نتایج تست طلایی برای افراد با تست منفی وجود ندارد. همچنین از پیش فرض استقلال شرطی که توسط Greene عنوان شد استفاده شد که در مورد روند گمشدگی داده‌های این مقاله کاملاً صدق می‌کند. زمانی که نتیجه تست تشخیصی برای فردی مشخص شود آنگاه بطور اتوماتیک وضعیت گمشدگی نیز مشخص می‌شود. پس وضعیت واقعی بیماری فرد به شرط دانستن نتیجه تست از وضعیت گمشدن فرد مستقل است از طرف دیگر وضعیت واقعی بیماری فرد، شانس گمشدن آن را افزایش یا کاهش می‌دهد. پس می‌توان مدل رگرسیونی را بر پایه احتمالات شرطی ( $G|T$ ) تعریف

منفی می‌باشد. این تعریف با معادلات بدست آمده در مدل ما کاملاً مطابقت دارد. در پایان امید است تا در آینده با ارائه مدل‌های قوی‌تر و قدرتمندتر گامی در پیشرفت و هر چه بهتر شدن نتایج حاصل از تست‌های تشخیصی کنیم.

### نتیجه‌گیری

بدین ترتیب به نظر می‌رسد که آنالیز بیزی به همراه اطلاعات آگاهی بخش یک ابزار قوی در شبیه‌سازی فضای آزمایش‌ها می‌باشد و می‌توان در تصحیح این نوع سوگرایی تاییدی از آن استفاده کرد.

### شکر و قدردانی

نویسندگان این مقاله بر خود لازم می‌دانند مراتب قدردانی خود را از خانم دکتر اشرف معینی و سایر پرسنل پژوهشگاه به دلیل فراهم آوردن امکان دسترسی به داده‌های بیماران، به عمل آورند.

کرد چون توزیع احتمالی موجود اطلاعات کافی درباره مقادیر گمشده اختیار ما قرار نمی‌دهد. برای اینکار شیوع بیماری در جامعه را بصورت عدم قطعیت در نظر گرفتیم و یک توزیع احتمالی به آن نسبت دادیم با این کار می‌توان اطلاعات مفیدی در مورد توزیع کلی استاندارد طلایی به مدل داد. مدل برای تست تشخیصی سونوگرافی با استاندارد طلایی هیستروسکوپی در تشخیص پولیپ بخوبی جواب می‌دهد اما در سایر تست باید بررسی شود که چگونه عمل می‌کند. بدین صورت که بعد از تعیین توزیع احتمالی شیوع بیماری و ترکیب آن با توزیع احتمالی تست تشخیصی برآوردها را می‌توان بدست آورد. در این مدل نسبت درست‌نمایی مثبت به نسبت درست‌نمایی منفی (۱۸) را برابر نسبت شانس مدل رگرسیون لوژستیک در نظر گرفتیم. نسبت شانس در مدل رگرسیونی برابر شانس مثبت شدن استاندارد طلایی در افراد تست مثبت به شانس منفی شدن استاندارد طلایی به افراد تست

پیوست شماره ۱- اثبات نسبت شانس در مدل رگرسیونی

$$\text{odds}_{11} = \frac{P_{11}}{1 - P_{11}} = \frac{\frac{\text{SNP}(G^+)}{\text{SNP}(G^+) + (1 - \text{SP})P(G^-)}}{\frac{(1 - \text{SN})P(G^+)}{\text{SPP}(G^-) + (1 - \text{SN})P(G^+)}} = \frac{\text{SNP}(G^+)}{(1 - \text{SN})P(G^-)}$$

$$\text{odds}_{21} = \frac{P_{21}}{1 - P_{21}} = \frac{\frac{(1 - \text{SN})P(G^+)}{\text{SPP}(G^-) + (1 - \text{SN})P(G^+)}}{\frac{\text{SPP}(G^-)}{\text{SPP}(G^-) + (1 - \text{SN})P(G^+)}} = \frac{(1 - \text{SN})P(G^+)}{\text{SPP}(G^-)}$$

$$\exp(\beta_{11}) = \exp(\beta_{22}) = \frac{\text{odds}_{11}}{\text{odds}_{21}} = \frac{\text{SN} \cdot \text{SP}}{(1 - \text{SN}) \cdot (1 - \text{SP})}$$

$$\text{odds}_{12} = \frac{P_{12}}{1 - P_{12}} = \frac{\frac{(1 - \text{SP})P(G^-)}{\text{SNP}(G^+) + (1 - \text{SP})P(G^-)}}{\frac{\text{SNP}(G^+)}{\text{SNP}(G^+) + (1 - \text{SP})P(G^-)}} = \frac{(1 - \text{SP})P(G^-)}{\text{SNP}(G^+)}$$

$$\text{odds}_{22} = \frac{P_{22}}{1 - P_{22}} = \frac{\frac{(\text{SP})P(G^-)}{(1 - \text{SN})P(G^+) + (\text{SP})P(G^-)}}{\frac{(1 - \text{SN})P(G^+)}{(1 - \text{SN})P(G^+) + (\text{SP})P(G^-)}} = \frac{(\text{SP})P(G^-)}{(1 - \text{SN})P(G^+)}$$

$$\exp(\beta_{21}) = \exp(\beta_{12}) = \frac{\text{odds}_{12}}{\text{odds}_{22}} = \frac{(1 - \text{SN}) \cdot (1 - \text{SP})}{\text{SN} \cdot \text{SP}}$$

$$\text{PPV} = \frac{y_{11}}{y_{11} + y_{12}} \Rightarrow y_{11} = \frac{\text{PPV} \cdot y_{12}}{1 - \text{PPV}}$$

$$y_{11} + y_{21} = n \cdot P(G = 1)$$

$$\text{SN} = \frac{y_{11}}{y_{11} + y_{21}} = \frac{\text{PPV} \cdot y_{12}}{(1 - \text{PPV}) \cdot n \cdot P(G = 1)}$$

$$NPV = \frac{y_{22}}{y_{21} + y_{22}} \Rightarrow y_{22} = \frac{NPV \cdot y_{21}}{1 - NPV}$$

$$y_{12} + y_{22} = n \cdot P(G = 0)$$

$$SP = \frac{y_{22}}{y_{12} + y_{22}} = \frac{NPV \cdot (n \cdot P(G = 1) - y_{11})}{(1 - NPV) \cdot n \cdot (1 - P(G = 1))}$$

### پیوست شماره ۲:- کد های مورد استفاده در برنامه Openbugs

در مدل سلسله مراتبی از توزیع یکنواخت برای حساسیت و ویژگی، توزیع نرمال برای عرض از مبدا مدل و توزیع بتا برای شیوع بیماری استفاده شد.

```

model{
for(i in 1:a){
for(j in 1:b){
y[i,j]~dbin(p[i,j],n[i])
p[i,j]<-exp(beta0+beta[i,j]*x[i,j])/(1+exp(beta0+beta[i,j]*x[i,j]))
}}
beta[1,1]<-log((SN*SP)/((1-SN)*(1-SP)))
beta[2,2]<-log((SN*SP)/((1-SN)*(1-SP)))
beta[2,1]<-log(((1-SN)*(1-SP))/(SN*SP))
beta[1,2]<-log(((1-SN)*(1-SP))/(SN*SP))
SN~dunif(0,1)
SP~dunif(0,1)
beta0~dnorm(0,0.01)
}sensitivity<-y[1,1]/(y[1,1]+y[2,1])
specificity<-y[2,2]/(y[2,1]+y[2,2])
}

```

## منابع

1. Lu Y, et al. A Bayesian approach to simultaneously adjusting for verification and reference standard bias in diagnostic test studies. *Stat Med* 2010; 29: 2532-43.
2. Whiting P, et al., Sources of variation and bias in studies of diagnostic accuracy. *Annals of Internal Medicine* 2004; 140: 189.
3. Begg CB, Greenes RA. Assessment of diagnostic tests when disease verification is subject to selection bias. *Biometrics* 1983; 207-15.
4. de Groot JAH, et al. Correcting for Partial Verification Bias: A Comparison of Methods. *Annals of Epidemiology*, 2011; 21: 139-48.
5. Martinez EZ, Alberto J, Achcar F. Louzada-Neto, Estimators of sensitivity and specificity in the presence of verification bias: A Bayesian approach. *Computational Statistics & Data Analysis* 2006; 51: 601-11.
6. Kosinski AS, Barnhart HX., Accounting for nonignorable verification bias in assessment of diagnostic tests. *Biometrics* 2003; 59: 163-71.
7. Buzaianu, M, Kadane JB. Adjusting for verification bias in diagnostic test evaluation: a Bayesian approach. *Statistics in medicine* 2008; 27: 2453-73.
8. Liu J, et al. Bayesian analysis of a matched case-control study with expert prior information on both the misclassification of exposure and the exposure-disease association. *Stat Med* 2009; 28: 3411-23.
9. Ayida G, et al. Uterine cavity assessment prior to in vitro fertilization: comparison of transvaginal scanning, saline contrast hysterosonography and hysteroscopy. *Ultrasound Obstet Gynecol* 1997; 10: 59-62.
10. Crosignani PG, and Rubin BL. Optimal use of infertility diagnostic tests and treatments. The ESHRE Capri Workshop Group. *Hum Reprod* 2000; 15: 723-32.
11. Seiner P, et al. Hysteroscopy in an Ivf Er Program: Clinical experience with 360 infertile patients. *Acta obstetricia et gynecologica Scandinavica* 1988; 67: 135-7.
12. Golan A, et al. Hysteroscopy is superior to hysterosalpingography in infertility investigation. *Acta obstetricia et gynecologica Scandinavica* 1996; 75: 654-6.
13. Golan A, et al. Diagnostic hysteroscopy: its value in an in-vitro fertilization/embryo transfer unit. *Human Reproduction* 1992; 7: 1433.
14. Yeung S, Wilkinson D. Adaptive Metropolis-Hastings samplers for the Bayesian analysis of large linear Gaussian systems. *Comp. Sci. Statist* 2002; 33: 128-38.
15. Spiegelhalter D, et al., WinBUGS Version 1.4 User Manual. 2005.
16. Cowles MK, Carlin BP, Markov Chain Monte Carlo convergence diagnostics: a comparative review. *Journal of the American Statistical Association*, 1996.
17. Gelman A, et al. Bayesian data analysis. 1995, New York: Chapman and Hall.
18. Deeks JJ, Altman DG. Statistics Notes: Diagnostic tests 4: likelihood ratios. *BMJ: British Medical Journal* 2004; 329: 168.

Iranian Journal of Epidemiology 2012; 8(2): 20-28.

**Original Article**

# A New Method for Correcting Verification Bias in Diagnostic Accuracy Studies Using A Bayesian Approach

Cheharazi M<sup>1</sup>, Shamsipour M<sup>2</sup>, Norouzi M<sup>3</sup>, Jafari F<sup>4</sup>

1- PhD student of Biostatistics, Department of Epidemiology and Reproductive Health, Reproductive Epidemiology Research Center, Royan Institute, ACECR, Tehran, Iran

2- PhD student of epidemiology, Department Of Epidemiology And Reproductive Health, Reproductive Epidemiology Research Center, Royan Institute, ACECR, Tehran, Iran

3- Msc of Epidemiology, Faculty member of Ghom University of Medical Science, Ghome, Iran

4- Assistant Professor of Community Medicine, shahed university, Tehran, Iran

**Corresponding author:** Shamsipour M., shamsman87@yahoo.com

**Background & Objectives:** One of the problems of diagnostic accuracy studies is verification bias. It occurs when standard test performed only for non-representative subsample of study subjects that diagnostic test done for them.

**Methods:** Patients that have had at least twice repeated failures in cycles IVF ICSI were included in this model. Patients were screened by using an ultrasonography and those with polyps recommended for hysteroscopy. A logistic regression with binomial outcome fit to predict the missing values (false and true negative), sensitivity and specificity. Bayesian methods was applied with informative prior on polyp prevalence. False and true negatives were estimated in Bayesian framework.

**Results:** A total of 238 patients were screened and 47 had polyps. Those with polyps are strongly recommended to undergo hysteroscopy, 47/47 decided to have a hysteroscopy and 37/47 were confirmed to have polyps. None of the 191 patients with no polyps in ultrasonography had hysteroscopy. The false negative was obtained 14 and true negative 177, so sensitivity and specificity was estimated easily after estimating missing data. Sensitivity and specificity were equal to 74% and 94% respectively.

**Conclusion:** Bayesian analyses with informative prior seem to be powerful tools in simulation experimental space.

**Keywords:** Verification bias, Bayesian method, Polyp