

## کاربرد مدل متغیر پنهان بیزی در پیش‌بینی زودهنگام دیابت بارداری در شرایط بدون تست استاندارد مرجع کامل با استفاده از بتا گنادوتروپین جفتی انسانی

مأده امینی<sup>۱</sup>، انوشیروان کاظم‌نژاد<sup>۲</sup>، فرید زایری<sup>۳</sup>، اعظم امیریان<sup>۴</sup>، نورالسادات کریمان<sup>۵</sup>

<sup>۱</sup> دانشجوی دکتری تخصصی آمار زیستی، گروه آمار زیستی، دانشکده علوم پزشکی، دانشگاه تربیت مدرس، تهران، ایران

<sup>۲</sup> استاد، گروه آمار زیستی، دانشکده علوم پزشکی، دانشگاه تربیت مدرس، تهران، ایران

<sup>۳</sup> استاد، مرکز تحقیقات پروتئومیکس و گروه آمار زیستی، دانشکده پیراپزشکی، دانشگاه علوم پزشکی شهید بهشتی، تهران، ایران

<sup>۴</sup> مربی، گروه مامایی، دانشکده پرستاری و مامایی، دانشکده علوم پزشکی جیرفت، جیرفت، ایران

<sup>۵</sup> دانشیار، گروه مامایی و بهداشت باروری، دانشکده پرستاری و مامایی، دانشکده علوم پزشکی شهید بهشتی، تهران، ایران

نویسنده رابط: انوشیروان کاظم‌نژاد، نشانی: تهران، بزرگراه جلال آل‌احمد، پل نصر، دانشگاه تربیت مدرس، دانشکده علوم پزشکی، گروه آمار زیستی، تلفن: ۸۲۸۸۳۸۷۵

پست الکترونیک: kazem\_an@modares.ac.ir

تاریخ دریافت: ۹۸/۰۳/۰۴؛ پذیرش: ۹۸/۸/۱۱

**مقدمه و اهداف:** دیابت بارداری یک عارضه طبی در بارداری بوده که تشخیص دیرهنگام آن می‌تواند اثرات نامطلوبی روی مادر و جنین بگذارد. هدف از انجام این پژوهش، برآورد پارامترهای دقت یک نشانگر زیستی برای پیش‌بینی زودهنگام دیابت بارداری در غیاب یک تست استاندارد مرجع کامل بود.

**روش کار:** این مطالعه روی ۵۲۳ خانم باردار که در سال‌های ۹۶ و ۹۷ به بیمارستان‌های مهدیه و طالقانی تهران مراجعه کرده بودند، انجام شد. مقادیر اندازه‌گیری شده بتا گنادوتروپین جفتی انسانی به عنوان پیشگوی دیابت بارداری برای همه زنان شرکت‌کننده در هفته ۱۷-۱۴ بارداری در چک لیست مربوط ثبت شد. برای برآورد حساسیت، ویژگی و مساحت زیر منحنی راک، از مدل متغیر پنهان بیزی استفاده شد. برآورد بیزی پارامترها در بسته نرم‌افزاری R2OpenBUGS در نرم‌افزار R نسخه ۳،۵،۳ محاسبه شد.

**یافته‌ها:** میان سن بارداری زنان شرکت‌کننده در این مطالعه، ۳۳ سال بود. در غیاب تست مرجع کامل، نتایج مدل به‌کار رفته برای مقدارهای حساسیت، ویژگی و مساحت زیر منحنی راک مربوط به بتا گنادوتروپین جفتی انسانی به ترتیب برابر با ۷۸ درصد (فاصله باورمند ۹۵ درصد: ۰/۱۸۳-۰/۱۶۶)، ۸۳ درصد (فاصله باورمند ۹۵ درصد: ۰/۷۴-۰/۱۸۹) و ۰/۷۲ (فاصله باورمند ۹۵ درصد: ۰/۱۸۸-۰/۱۶۴) بود.

**نتیجه‌گیری:** بر اساس یافته‌های به‌دست آمده در این مطالعه، بتا گنادوتروپین جفتی انسانی در غیاب تست مرجع کامل می‌تواند برای پیش‌بینی زودهنگام دیابت در خانم‌های باردار مناسب باشد.

واژگان کلیدی: مدل متغیر پنهان، منحنی راک، روش بیزی، دیابت بارداری، گنادوتروپین جفتی انسانی

### مقدمه

در آزمایش‌های تشخیصی بالینی، پزشکان درباره نتیجه آزمایش می‌توانند بر اساس سابقه بیماری یا بروز علائم توسط خود بیمار یا از طریق تصویربرداری پزشکی و غیره تصمیم‌گیری کنند (۱). اغلب آزمایش‌های تشخیصی بر پایه نشانگرهای زیستی هستند. طی سال‌های متمادی، استفاده از نشانگرهای زیستی که قادر به تشخیص درست وجود و شدت بیماری‌هایی مانند عفونی، مغزی، دیابت و سرطان باشند؛ رو به افزایش است. مرحله کلیدی در ارزیابی قدرت تشخیص یک نشانگر زیستی یا تست تشخیصی تعیین دقت آن است که توانایی تست را در تمایز بین افراد با و

بدون وضعیت واقعی بیماری نشان می‌دهد (۲). در مطالعه‌های تشخیصی، وضعیت واقعی بیماری توسط یک تست مرجع کامل (یا استاندارد طلایی) تعیین می‌شود که یک روش بسیار دقیق و تقریباً بدون خطا در رده‌بندی است. برای ارزیابی دقت تشخیص از معیارهای مختلف مانند حساسیت<sup>۱</sup>، ویژگی<sup>۲</sup> و سطح زیر منحنی راک<sup>۳</sup> می‌توان استفاده کرد. این

<sup>۱</sup> Sensitivity

<sup>۲</sup> Specificity

<sup>۳</sup> Area Under Receiver Operating Characteristic (ROC) curve; AUC

گنادوتروپین جفتی انسانی<sup>۶</sup> است. در این مطالعه با استفاده از مدل متغیر پنهان بیزی، معیارهای دقت تشخیص  $\beta$ -hCG برای ارزیابی توان پیش‌بینی زودرس دیابت بارداری در شرایط نبود OGTT در اوایل سه ماهه دوم بارداری مورد پژوهش قرار گرفت.

## روش کار

### جامعه مورد پژوهش

داده‌های این مطالعه مقطعی-تحلیلی را ۵۲۳ خانم باردار تشکیل می‌دهد که در سال‌های ۹۶ و ۹۷ به بیمارستان‌های مهدیه و طالقانی تهران برای انجام تست‌های حاملگی مراجعه کرده بودند. تعداد نمونه‌های این مطالعه با در نظر گرفتن اطمینان ۹۵ درصد، احتمال خطای نوع اول ۵ درصد و احتمال خطای نوع دوم ۱۰ درصد در گروه شاهد ۴۶۰ نفر و در گروه مورد ۶۳ نفر برآورد شد. روش نمونه‌گیری به صورت چند مرحله‌ای بود. به این ترتیب که ابتدا دو مرکز به تصادف از بین مراکز آموزشی و درمانی وابسته به دانشگاه علوم پزشکی شهید بهشتی انتخاب شدند و تا رسیدن به تعداد نمونه مورد نیاز، نمونه‌گیری ادامه داشت. معیار ورود به مطالعه شامل تک‌قلوبی، عدم وجود سابقه دیابت، عدم وجود بیماری‌های قلبی-عروقی، عدم وجود سابقه پره‌اکلامپسی و فشار خون بالا، عدم مصرف سیگار و مواد مخدر و عدم وجود سابقه مرده‌زایی و سقط مکرر بود. همچنین معیار خروج شامل پایین یا بالا بودن غیر عادی سطوح آلفا فیتوپروتئین و استریول غیرکونژوگه بود. از همه شرکت‌کنندگان پیش از ورود به مطالعه، رضایت‌نامه آگاهانه کتبی اخذ شد. سپس ویژگی‌های بالینی و جمعیت‌شناختی در پرسشنامه مربوطه ثبت شدند. سطح قند خون ناشتا و تست تحمل گلوکز خوراکی با ۷۵ گرم گلوکز با استفاده از دستگاه آنالیزور ۹۱۱ هیتاچی اندازه‌گیری و توسط کارشناس آزمایشگاه برای افراد شرکت‌کننده در هفته ۲۸-۲۴ بارداری ثبت شد. همچنین میزان  $\beta$ -hCG به روش Radio Immuno Assay در یک آزمایشگاه با استفاده از دستگاه گاما کانترا Genesis و توسط پژوهشگر و کارشناس آزمایشگاه در هفته ۱۷-۱۴ بارداری در مقیاس کمی پیوسته و بر حسب میلی‌واحد بین‌المللی در میلی‌لیتر (mIU/ml)<sup>۷</sup> اندازه‌گیری شد و در چک لیستی برای همه افراد شرکت‌کننده به ثبت رسید. در ادامه چگونگی مدل‌بندی پاسخ یک نشانگر زیستی زمانی که وضعیت واقعی بیماری نامعلوم

شاخص‌ها بیان می‌کنند که پاسخ تست تشخیصی مورد نظر چقدر هم‌سو با پاسخ استاندارد طلایی است. در مواقعی ممکن است استفاده از استاندارد طلایی به دلیل‌هایی مانند سخت بودن در اجرا، تأخیر در اجرا، هزینه‌بر بودن و تهاجمی بودن امکان‌پذیر نباشد و وضعیت واقعی بیماری را نتوان به درستی تشخیص داد. در چنین شرایطی ممکن است خطا در رده‌بندی<sup>۱</sup> رخ دهد. این خطا می‌تواند برآورد پارامترهای دقت تشخیص و حتی شیوع بیماری را تحت تأثیر قرار دهد و آن‌ها را با اریبی روبرو کند. از این‌رو بهتر است از روش‌هایی استفاده شود که در نبود استاندارد طلایی، برآورد پارامترها را با دقت بیش‌تری برآورد کنند (۳-۵).

یکی از مدل‌های پرتعدادار برای تحلیل داده‌ها در شرایط فقدان استاندارد طلایی، مدل متغیر پنهان است. این مدل پاسخ تست‌های تشخیصی را به شرط وضعیت پنهان بیماری مدل‌بندی می‌کند. برای برآورد پارامترهای مدل زمانی که نتایج یک تست تشخیصی دقیقاً مشخص نباشد، استفاده از روش‌های حداکثر درست‌نمایی برای برآورد پارامترها ممکن است نامعتبر بوده و به شناسایی مدل منجر نشود. راه حل مناسب در چنین مواقعی استفاده از روش بیزی است. در چهارچوب این روش، با فرض این‌که پارامترها یک متغیر تصادفی هستند، یک توزیع احتمالی با عنوان توزیع احتمال پیشین<sup>۲</sup> در نظر گرفته می‌شود. توزیع پسین پارامترها نیز می‌توانند از راه روش‌های محاسباتی زنجیر مارکوفی مونت کارلو<sup>۳</sup> مانند نمونه‌گیری گیبز<sup>۴</sup> به دست آیند. در نهایت استنباط بر اساس برآوردهای حاصل از توزیع پسین که متداول‌ترین آن‌ها میانگین و واریانس این توزیع هستند، انجام می‌شود (۹-۶).

دیابت بارداری یکی از شایع‌ترین عوارض حاملگی بوده که دارای عوارض نامطلوب در دراز مدت و کوتاه مدت بر مادر و جنین است. بنابراین از دید بالینی، تشخیص زودهنگام آن برای انجام مداخله‌های به موقع اهمیت به‌سزایی دارد. استاندارد طلایی تشخیص دیابت بارداری یعنی تست تحمل گلوکز خوراکی<sup>۵</sup> دارای مشکلاتی از قبیل وقت‌گیر بودن، نیاز به ناشتا بودن از شب پیش از آزمایش، تکرارپذیری کم و انجام آن تقریباً در اواخر دوره بارداری است (۱۰). یکی از نشانگرهای زیستی که ممکن است برای تشخیص دیابت در اوایل دوره بارداری مناسب باشد، بتا

<sup>۱</sup> Misclassification

<sup>۲</sup> Prior distribution

<sup>۳</sup> Markov Chain Monte Carlo; MCMC

<sup>۴</sup> Gibbs sampling

<sup>۵</sup> Oral Glucose Tolerance Test; OGTT

<sup>۶</sup> Beta-human Chronic Gonadotropin; ( $\beta$ -hCG)

<sup>۷</sup> Milli-international units per milliliter; mIU/ml

باشد، شرح داده می‌شود.

می‌توان تعیین کرد که توزیع پاسخ نشانگر زیستی در گروه افراد سالم ( $S_D$ ) چقدر به توزیع پاسخ نشانگر در گروه افراد بیمار ( $S_B$ ) نزدیک است. چوی و همکاران این معیار را از رابطه زیر به دست آوردند:

$$\Delta = \Phi\left(\frac{\delta - \mu_D}{\sqrt{\sigma_D^2}}\right) \quad (5)$$

به طوری که  $\delta$  بیانگر صدک ۹۵ام برای مقادیر نشانگر زیستی در گروه افراد سالم است. اگر میانگین پاسخ نشانگر زیستی در گروه افراد بیمار برابر با صدک ۹۵ام پاسخ نشانگر زیستی در گروه افراد سالم باشد، آن گاه  $\Delta$  در رابطه ۵ برابر با ۰/۵ خواهد شد (۱۱).

در این مطالعه، برآورد پارامترها با رویکرد بیزی و با استفاده از روش مونت کارلوی زنجیرمارکوفی به دست آمد. در روش بیزی، توزیع پسین بر اساس استنباط در مورد پارامترها بوده که این توزیع با تلفیق داده‌های مشاهده شده و توزیع‌های پیشین محاسبه می‌شود. با این حال محاسبه توزیع پسین به علت ابعاد و پیچیدگی مدل به سادگی امکان پذیر نیست. بنابراین برای تحلیل از روش‌های شبیه‌سازی MCMC استفاده شد. در این روش، با استفاده از نمونه‌گیری‌های مکرر از توزیع‌های پسین پارامترها، زنجیرهای مارکوفی حاصل می‌شود که در صورت همگرا بودن زنجیره‌های تولید شده، توزیع حدی این نمونه‌ها به توزیع پسین میل می‌کند. در نتیجه می‌توان نمونه‌های تولید شده را از توزیع‌های حاشیه‌ای پسین پارامترها دانست. برای انجام استنباط بیزی انتخاب توزیع‌های پیشین مناسب برای هر یک از پارامترهای مدل ضرورت دارد. با توجه به عدم وجود دانش پیشین کافی در مورد پارامترها در این مقاله، از توزیع پیشین ناآگاهی بخش استفاده شد. بر این اساس، برای پارامترهای  $\mu_D$  و  $\mu_B$  توزیع پیشین  $N(0,0.001)$  برای  $1/\sigma_D^2$  و  $1/\sigma_B^2$  پیشین  $(0.01, \gamma)$  gamma (0.01, 1) از پیشین  $\pi$  استفاده شد. برآورد پسین پارامترهای دقت تشخیص (شامل میانگین، انحراف معیار و فاصله باورمند<sup>۲</sup> ۹۵ درصد) بر اساس ۴۰۵۰۰ تکرار و رقیق‌سازی ۱۵۰۰۰ به دست آمد. همان‌طور که اشاره شد، در برآوردیابی به روش MCMC بررسی همگرایی فرآیند گیبز به یک توزیع مانا یا توزیع پسین پارامتر موضوع مهمی است. برای تشخیص همگرایی زنجیره‌های تولید شده روش‌های متعددی از جمله روش‌های گرافیکی و آزمون‌های آماری وجود دارد. در این پژوهش، از روش گرافیکی نمودارهای خود هم‌بستگی استفاده شد.

ارزیابی دقت نشانگر زیستی در حالت بدون استاندارد طلایی فرض کنید  $Y_i$  پاسخ نشانگر زیستی برای  $i$ امین فرد در یک نمونه تصادفی  $n$  تایی ( $i = 1, \dots, n$ ) از افراد باشد.  $S_i$  هم بیانگر وضعیت واقعی بیماری اما به صورت پنهان برای  $i$ امین فرد مورد مطالعه است (کد ۱ بیانگر بیمار بودن و کد صفر بیانگر سالم بودن فرد است). حال مدل زیر را در نظر بگیرید:

$$S_i \sim \text{Bernoulli}(\pi) \quad i = 1, \dots, 523 \quad (1)$$

$$Y_i \sim p_1(\cdot | \mu_D, \sigma_D^2)^{S_i} p_2(\cdot | \mu_B, \sigma_B^2)^{1-S_i} \quad (2)$$

در رابطه ۱،  $\pi$  بیانگر احتمال ابتلا به بیماری بوده به طوری که  $P(S_i = 1) = 1 - P(S_i = 0) = \pi$  است.

در رابطه ۲،  $D$  و  $B$  به ترتیب نشانگرهای مربوط به گروه‌های بیمار و سالم هستند.  $p_1$  و  $p_2$  نیز به ترتیب توابع چگالی احتمال نشانگر زیستی در گروه‌های بیمار و سالم را نشان می‌دهند؛ به طوری که  $p_1(\cdot)$  تابع چگالی احتمالی  $N(\mu_D, \sigma_D^2)$  و  $p_2(\cdot)$  تابع چگالی احتمالی  $N(\mu_B, \sigma_B^2)$  هستند.  $\mu_D$  و  $\mu_B$  به ترتیب میانگین نشانگر زیستی را در گروه افراد بیمار و سالم نشان می‌دهند.  $\sigma_D^2$  و  $\sigma_B^2$  هم به ترتیب بیانگر واریانس نشانگر زیستی در گروه افراد بیمار و سالم هستند. برای شناسایی پذیری مدل فرض بر این است که  $\mu_B < \mu_D$  باشد. منحنی راک به ازای نقاط برش<sup>۱</sup> (C) مختلف در بازه  $-\infty < c < \infty$  با رسم حساسیت در مقابل ویژگی- ۱ که مختصات آن به صورت زیر است، به دست می‌آید؛ به طوری که  $\Phi$  تابع توزیع تجمعی متغیر نرمال استاندارد را نشان می‌دهد.

$$\left[ 1 - \Phi\left(\frac{c - \mu_B}{\sqrt{\sigma_B^2}}\right), 1 - \Phi\left(\frac{c - \mu_D}{\sqrt{\sigma_D^2}}\right) \right] \quad (3)$$

در این مطالعه انتخاب نقطه برش بهینه بر اساس شاخص یودن بود. AUC نیز از رابطه زیر محاسبه شد:

$$AUC = \Phi\left(-\frac{\mu_D - \mu_B}{\sqrt{\sigma_D^2 + \sigma_B^2}}\right) \quad (4)$$

با توجه به این که هدف این مطالعه، ارزیابی دقت نشانگر زیستی در پیش‌بینی یک پیامد دو حالتی است، بنابراین با تعریف معیاری

<sup>۱</sup>Cutoff point

<sup>۲</sup> Credible interval

عدم ابتلا به بیماری تخصیص یافتند. همان‌طور که در این جدول ملاحظه می‌شود، طبق نقطه برش بهینه، مقدار حساسیت ۷۸ درصد (فاصله باورمند ۹۵ درصد: ۰/۸۳-۰/۶۶)، مقدار ویژگی ۸۳ درصد (فاصله باورمند ۹۵ درصد: ۰/۸۹-۰/۷۴) و مقدار AUC نیز ۰/۷۲ (فاصله باورمند ۹۵ درصد: ۰/۸۸-۰/۶۴) به دست آمدند. از طرفی مقدار محاسبه شده برای  $\Delta$  (۰/۳۲) گویای آن است که توزیع  $\beta$ -hCG در دو گروه ابتلا و عدم ابتلا به دیابت دارای حداقل هم‌پوشانی بودند. همچنین آزمون گوک هم‌گرایی به توزیع پسین را تأیید کرد ( $|Z| < 2$ ). نمودار خود هم‌بستگی نیز برای همه پارامترهای مدل در شکل شماره ۱ بیانگر هم‌گرایی زنجیره مارکوفی بود. شکل شماره ۲ نمودار منحنی راک برآورد شده به روش بیزی در غیاب تست مرجع کامل را به صورت شهودی نشان می‌دهد.

هم‌چنین در بسته CODA در R از آزمون آماری گوک<sup>۱</sup> استفاده شد. در این آزمون، اگر قدر مطلق آماره آزمون کم‌تر از ۲ باشد، می‌توان نتیجه گرفت که همه زنجیره‌های تولید شده به توزیع پسین همگرا شده‌اند.

در این مطالعه، توصیف متغیرهای کمی به صورت میانگین  $\pm$  انحراف معیار گزارش شد. آزمون کولموگروف-اسمیرنوف تک‌نمونه‌ای برای بررسی نرمال بودن توزیع پاسخ  $\beta$ -hCG به کار رفت. محاسبه شاخص‌های توصیفی با استفاده از نرم‌افزار SPSS نسخه ۲۲ بود. برآورد بیزی پارامترها در بسته نرم‌افزاری R2OpenBUGS در نرم‌افزار R نسخه ۳،۵،۱ محاسبه شد.

## یافته‌ها

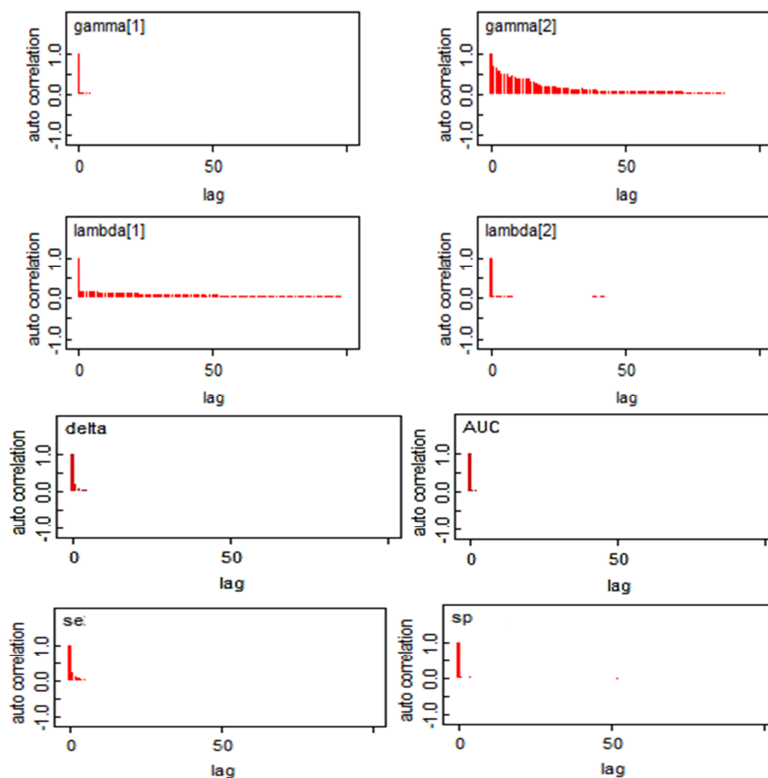
در مجموع ۵۲۳ زن باردار در محدوده سنی ۲۰-۴۰ ساله در این مطالعه شرکت داشتند. میانگین  $\pm$  انحراف معیار سن زنان باردار  $28/76 \pm 5/33$  سال بود. هم‌چنین میانگین  $\pm$  انحراف معیار شاخص توده بدنی  $24/57 \pm 3/22$  کیلوگرم بر متر مربع بود. میانگین  $\pm$  انحراف معیار  $\beta$ -hCG برای زنان شرکت‌کننده نیز  $1/17 \pm 0/77$  mIU/ml بود. در ادامه به برآورد معیارهای دقت تشخیص در غیاب یک تست استاندارد طلایی با روش بیزی پرداخته می‌شود.

ابتدا تعیین شد که توزیع پاسخ  $\beta$ -hCG در غیاب تست استاندارد مرجع کامل نرمال نبود ( $P < 0/001$ ). از این رو با انجام تبدیل‌های مختلف (مانند لگاریتمی و باکس-کاکس) نرمال بودن پاسخ این نشانگر بررسی شد. سرانجام با تبدیل لگاریتمی نرمال بودن توزیع مورد تأیید قرار گرفت ( $P < 0/05$ ). سپس مدل متغیر پنهان به داده‌ها برازش داده شد و برآورد بیزی میانگین، واریانس، حساسیت، ویژگی و سطح زیر منحنی راک به دست آمدند. در جدول شماره ۱، برآورد پسین پارامترهای مدل به همراه معیارهای دقت تشخیص بر اساس میانگین، انحراف معیار و فاصله باورمند ۹۵ درصد گزارش شده است. با استفاده از شاخص یودن، مقدار نقطه برش بهینه ۱۱۶ mIU/ml محاسبه شد و این یعنی فردی که میزان  $\beta$ -hCG آن بیشتر از ۱۱۶ mIU/ml است، به عنوان شخص در معرض خطر ابتلا به دیابت بارداری در هفته ۱۷-۱۴ بارداری شناخته می‌شود. با برآورد متغیر پنهان و تشکیل رده‌ها، از بین ۵۲۳ نفر، حدود ۷۲ نفر (۱۴ درصد) در گروه در معرض خطر ابتلا به دیابت بارداری قرار گرفتند. ۴۵۱ نفر (۸۶ درصد) نیز به گروه

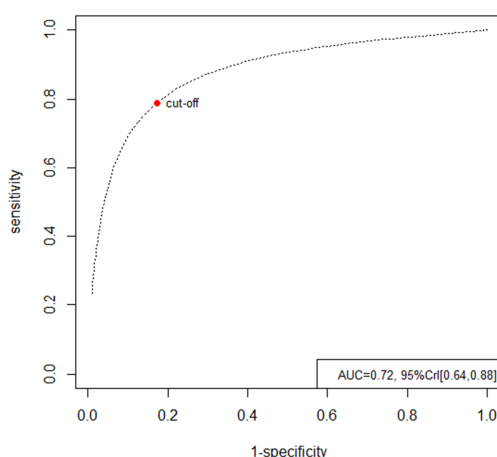
<sup>۱</sup>Geweke test

جدول شماره ۱ - برآوردهای پسین پارامترهای مدل متغیر پنهان بیزی و معیارهای دقت تشخیصی برای بتا گنادوتروپین جفتی انسانی در غیاب تست استاندارد مرجع کامل

پارامتر	میانگین	انحراف معیار	خطای MC	۲/۵ درصد	۹۷/۵ درصد
میانگین (گروه بیمار)	۱/۳۸	۰/۱۰	۰/۰۲	۱/۲۳	۱/۵۲
میانگین (گروه سالم)	۱/۱۴	۰/۰۲	۰/۰۰۲	۱/۰۴	۱/۲۱
واریانس (گروه بیمار)	۰/۲۲	۰/۰۵	۰/۰۰۱	۰/۱۶	۰/۳۳
واریانس (گروه سالم)	۰/۲۷	۰/۱۱	۰/۰۹	۰/۱۸	۰/۳۸
حساسیت	۰/۷۸	۰/۰۹	۰/۰۱	۰/۶۶	۰/۸۳
ویژگی	۰/۸۳	۰/۰۲	۰/۰۰۱	۰/۷۴	۰/۸۹
مساحت زیر منحنی راک	۰/۷۲	۰/۰۵	۰/۰۰۸	۰/۶۴	۰/۸۸
$\Delta$	۰/۳۲	۰/۰۸	۰/۰۰۱	۰/۲۶	۰/۴۱



نمودار شماره ۱ - نمودار خود هم‌بستگی پارامترهای مدل پنهان بیزی و معیارهای دقت تشخیصی بتا گنادوتروپین جفتی انسانی



نمودار شماره ۲- منحنی راک برآورد شده به روش بیزی برای بتا گنادوتروپین جفتی انسانی در غیاب تست استاندارد مرجع کامل

## بحث

در این پژوهش، داده‌های مربوط به دیابت بارداری مورد تحلیل قرار گرفت. این بیماری یک معضل بهداشتی رو به افزایش در سراسر دنیا است و ممکن است منجر به بروز بیماری‌هایی چون پره اکلامپسی، ماکروزومی نوزاد و مشکلاتی در انجام سزارین شود (۱۲). درمان دیابت بارداری در کاهش ماکروزومی، وزن بالای نوزاد، دیستوشی شانه، فشار خون بالا در بارداری و اختلالات پره اکلامپسی نقش مهمی دارد. تست تحمل گلوکز خوراکی با ۷۵ گرم گلوکز به‌عنوان استاندارد طلایی برای تشخیص دیابت بارداری بین هفته‌های ۲۴-۲۸ بارداری انجام می‌شود. بنابراین به دلیل آن که این تست تقریباً در اواخر بارداری انجام می‌شود، نمی‌تواند برای پیش‌بینی زود هنگام دیابت بارداری و انجام درمان و مداخله‌های به موقع مناسب باشد. از این رو دستیابی به نشانگرهای زیستی که بتوانند دیابت بارداری را در اوایل دوره بارداری برای پیشگیری از عوارض نامطلوب تشخیص دهند، دارای اهمیت است. بتا گنادوتروپین جفتی انسانی نشانگری است که برای تشخیص سندروم داون در اوایل سه ماهه دوم بارداری مورد استفاده قرار می‌گیرد (۱۳). با این حال در برخی پژوهش‌های قبلی، به‌کارگیری این نشانگر در تشخیص دیابت بارداری نیز مورد بررسی قرار گرفته است. در این مطالعه نیز دقت تشخیصی این نشانگر در غیاب استاندارد طلایی برای پیش‌بینی زودرس دیابت

بارداری با مدل متغیر پنهان بیزی ارزیابی شد. یافته‌ها نشان دادند که  $\beta$ -hCG قادر است ۷۸ درصد مادران در معرض خطر ابتلا به دیابت بارداری را در هفته ۱۷-۱۴ بارداری به درستی تشخیص دهد. از طرفی مساحت زیر منحنی راک (۰/۷۲) گویای توان پیش‌بینی خوب و قدرت رده‌بندی بالا این نشانگر در نبود استاندارد طلایی است. یافته‌های این مطالعه با برخی مطالعه‌های پیشین هم‌خوانی دارد. برای مثال، هور و همکاران در مطالعه‌ای روی ۱۵۵۳ زن باردار در سال ۲۰۱۷ نشان دادند  $\beta$ -hCG پیشگوی مناسبی برای دیابت بارداری در اوایل سه ماهه دوم می‌تواند باشد (۱۰). همچنین بر اساس یافته‌های سایرین و همکاران در سال ۲۰۰۸ میلادی، ۶۹/۶ درصد زنان باردار مبتلا به دیابت توسط  $\beta$ -hCG به درستی به‌عنوان بیمار شناخته شده بودند. از طرفی در مطالعه آن‌ها مقدار محاسبه شده برای AUC، ۰/۵۶۵ بود (۱۴). در مقابل، سانکن و بارتلس روی ۱۹۳۶۵ زن باردار به نتیجه معنی‌داری در مورد رابطه بین  $\beta$ -hCG و دیابت بارداری بین هفته ۲۰-۱۵ بارداری دست نیافتند (۱۵). در مطالعه دیگری توسط کاواک و همکاران،  $\beta$ -hCG تنها قادر بود ۵۷/۵ درصد افراد را به درستی به‌عنوان بیمار تشخیص دهد. همچنین مقدار محاسبه شده برای AUC (۰/۵۸) بیانگر قدرت ضعیف این نشانگر در پیش‌بینی زود هنگام دیابت بارداری و تفکیک رده‌ها بود (۱۶). به‌طور کلی در همه مطالعه‌های یاد شده، بررسی عملکرد  $\beta$ -hCG در حضور استاندارد طلایی و با روش‌های کلاسیک

صورت گرفته بود. نکته مهم دیگر آن است که تنها در دو مطالعه معیارهای دقت تشخیص  $\beta$ -hCG محاسبه شده بود. بر این اساس این پژوهش، نخستین مطالعه‌ای است که به این موضوع از نگاه دیگر و با روش تحلیلی پیشرفته‌تر پرداخته است.

پیشرفت تکنولوژی در علوم پزشکی، پژوهشگران را ترغیب به استفاده از روش‌های تشخیصی مختلف برای شناسایی و پیش‌بینی به موقع بیماری‌های مختلف کرده است. با این حال، تشخیص بسیاری از بیماری‌ها وابسته به استفاده از تست‌های تشخیصی یا نشانگرهای زیستی است. در ارزیابی دقت تست تشخیصی، یافته‌های به‌دست آمده از آن را با یافته‌های تست مرجع کامل مقایسه می‌کنند. در بهترین حالت، هر دو تست تشخیصی و مرجع کامل روی همه بیماران انجام می‌شوند، اما گاهی به دلایلی مانند گران بودن، غیر اخلاقی بودن، عدم دسترسی یا غیر قابل اجرا روی همه افراد مورد مطالعه نمی‌توان از تست مرجع کامل استفاده کرد. نادیده گرفتن این موضوع در ارزیابی دقت تست تشخیصی، منجر به ایجاد سوگیری در برآورد پارامترهای دقت تشخیص می‌شود. این سوگیری می‌تواند مشکلات جدی را در حیطه پزشکی تشخیصی مانند اتخاذ تصمیم‌گیری‌های اشتباه در درمان، تأخیر در ارایه درمان‌های بهتر و همچنین تحمیل بار مالی سنگین به همراه داشته باشد. به‌طور کلی برای حل این معضل، پژوهشگران آماری راه‌کارهای مختلفی از جمله روش جانهمی، روش تصحیح استاندارد مرجع ناقص و روش ساخت استاندارد مرجع را پیشنهاد داده‌اند (۱۷). در روش ساخت استاندارد مرجع، یکی از رویکردهای مورد استفاده تحلیل کلاس پنهان است. یک شرط اساسی در این مدل، برقراری فرض استقلال شرطی بین یافته‌های تست‌های تشخیصی است. در صورت عدم برقراری این فرض، لحاظ کردن همبستگی بین تست‌ها ضروری است. برای مثال دندوکوری و جوزف در سال ۲۰۰۱ میلادی از مدل کلاس پنهان بیزی برای ارزیابی دقت تست‌های تشخیصی دوحالتی با مد نظر قرار دادن وابستگی شرطی بین تست‌ها در غیاب استاندارد طلایی استفاده کردند (۱۸). بارادو و همکاران در سال ۲۰۱۶ مدل کلاس پنهان بیزی آمیخته را برای انتخاب ترکیب بهینه‌ای از نشانگرهای زیستی پیوسته مربوط به بیماری آلزایمر با حضور تست مرجع ناقص ارایه دادند (۱۹). در سال ۲۰۱۶ جعفرزاده و همکاران نیز تحلیل کلاس پنهان با اثرات تصادفی را در ترکیب تست‌های تشخیصی پیوسته در شرایط بدون استاندارد طلایی به‌کار بردند (۲۰). در این مطالعه برای مدل‌بندی یک تست تشخیصی با شرطی کردن روی وضعیت پنهان بیماری از روش پیشنهادی

چوی و همکاران در سال ۲۰۰۶ استفاده شد (۱۱). آن‌ها مدل متغیر پنهان را با روش بیزی برای ارزیابی دقت دو تست تشخیصی پیوسته هم‌بسته استفاده کرده بودند. بر اساس ادعایی که داشتند، بیان کردند زمانی این روش عملکرد خوبی دارد که مقدار  $\Delta$  کوچک (یعنی کم‌تر از ۰/۵) باشد. بر اساس  $\Delta$  محاسبه شده در این مقاله (۰/۳۲) نیز می‌توان نتیجه گرفت که روش مورد استفاده برای تحلیل این داده‌ها مناسب بود. نکته دیگری که باید به آن اشاره کرد این است که در شرایط بدون استاندارد طلایی زمانی که متغیرهای آشکار و پنهان هر دو از نوع گسسته باشند (دوحالتی و یا چندحالتی) مدل متغیر پنهان را تحت عنوان مدل کلاس پنهان<sup>۱</sup> نام‌گذاری کرده‌اند. تحلیل متغیر پنهان یک روش منعطف بوده که قادر به ترکیب انواع پاسخ تست‌های تشخیصی دوحالتی، چندحالتی و پیوسته است.

روش برآوردی به‌کار رفته در این مطالعه، روش بیزی بود. با توجه به باریک بودن فاصله باورمند بیزی و نیز کوچک بودن خطای MC برای همه پارامترهای مدل می‌توان نتیجه گرفت که برآوردها از دقت خوبی برخوردار بودند. به‌طور کلی می‌توان بیان کرد که تقریباً در همه مطالعه‌های انجام شده در زمینه ارزیابی دقت تست‌های تشخیصی در غیاب استاندارد مرجع کامل از این رویکرد برای برآورد پارامترهای دقت تشخیص استفاده شده بود (۲۴-۲۱). بیش‌تر پژوهشگران نیز هم راستا با یافته‌های این مطالعه معتقد بودند که استفاده از روش بیزی، برآوردهای معتبرتری را در مورد دقت تست تشخیصی به‌دست می‌دهد. یکی از مزیت‌های روش بیزی آن است که کمک می‌کند تا اطلاعات بر اساس اطلاعات پیشین درباره دقت تست تشخیصی وزن‌دهی شود. به بیان بهتر، در این رویکرد می‌توان اطلاعات پیشین را با داده‌های موجود ترکیب کرد که این امر منجر به پردقت‌تر شدن برآوردها می‌شود. از طرفی با مشاهده یک داده جدید و استدلال در مورد توزیع احتمالی، تصمیم‌های بهینه‌ای را می‌توان اتخاذ نمود. همچنین در استنباط بیزی، فاصله‌های باورمند برخلاف فاصله‌های اطمینان در روش کلاسیک دارای ویژگی‌های خوب همگرایی هستند. زمانی که مقدارهای معیارهای تست تشخیصی از پیش دقیقاً مشخص نیست، استفاده از پارامترهای ثابت ممکن است نامعتبر باشد. بنابراین به‌جای محدود کردن پارامترها اطراف یک مقدار ثابت، می‌توان از توزیع پیشین آگاهی بخش در زمانی که اطلاع پیشینی در دسترس باشد؛ استفاده کرد، اما در عمل مواردی

<sup>۱</sup>Latent class model

بهداشتی- درمانی مناسب آن به‌ویژه به زنان در معرض خطر بالا، می‌توان از پیشرفت و رخداد عوارض نامطلوب ناشی از این بیماری جلوگیری کرد و فایده‌های چشم‌گیری در پیامدهای بارداری و سلامت آینده هم مادر و هم فرزندان داشته باشد. بر اساس یافته‌های به‌دست آمده در این پژوهش، با توجه به این‌که بتا گنادوتروپین جفتی انسانی توانایی کافی در پیش‌بینی زود هنگام دیابت بارداری در اوایل سه ماهه دوم بارداری داشت، بنابراین استفاده از آن را در مطالعه‌های مشابه آینده در شرایط بدون استاندارد طلائی می‌توان پیشنهاد کرد. از جمله محدودیت‌های مهم این پژوهش، ثبت ناقص اطلاعات مربوط به برخی متغیرهای کمکی در پرونده پزشکی بیماران مانند فشار خون، سابقه پره‌اکلمپسی، سابقه دیابت فامیلی، سابقه اختلال متابولیسم گلوکز، سابقه زایمان زودرس و عدم تعمیم یافته‌ها به جامعه کل مادران باردار مبتلا به دیابت بارداری بود.

### تشکر و قدردانی

این مقاله برگرفته از بخشی از رساله دکترای تخصصی آمار زیستی در دانشگاه تربیت مدرس بود. نویسندگان این مقاله از حمایت دانشگاه تربیت مدرس و نیز دانشگاه علوم پزشکی شهید بهشتی صمیمانه تشکر و قدردانی می‌نمایند.

وجود دارد که اطلاعات پیشین، اغلب مبهم یا ناکافی هستند. در این مطالعه، با توجه به محدود بودن اطلاعات پیشین در مورد پارامترهای مدل، از پیشین ناآگاهی بخش برای محاسبه توزیع پسین استفاده شد (۲۵). در سال ۲۰۱۴ کولین و هاین مروری بر استفاده از هر دو روش بیزی و کلاسیک برای ارزیابی دقت تست‌های تشخیصی از نوع‌های مختلف دو حالتی، ترتیبی و پیوسته در نبود استاندارد طلائی داشتند. بررسی آن‌ها نشان داد که محاسبه برآوردهای عددی با استفاده از روش‌های کلاسیک بسیار پیچیده بوده و ممکن است منجر به عدم هم‌گرایی توزیع پارامترها شود. در حالی‌که روش‌های بیزی در برآورد عددی عملکرد بهتری دارند (۲۶). در این مطالعه از روش MCMC با استفاده از الگوریتم نمونه‌گیری گیبز استفاده شد. از آن‌جا که در روش بیزی، محاسبه مستقیم برآوردگر پسین هر پارامتر به علت پیچیدگی توزیع پسین توأم دشوار بوده و مشکلاتی مانند بالا بودن بعد انتگرال، محاسبه مستقیم را غیر ممکن می‌کند؛ از این‌رو برای غلبه بر این مشکلات، استفاده از روش MCMC پیشنهاد شده است (۶).

### نتیجه‌گیری

به‌طور کلی، با در نظر گرفتن پرهزینه بودن درمان بیماری دیابت بارداری، شناسایی زود هنگام، آموزش و آرایه خدمات

### منابع

- Schünemann HJ, Oxman AD, Brozek J, Glasziou P, Jaeschke R, Vist GE, et al. Grading quality of evidence and strength of recommendations for diagnostic tests and strategies. *BMJ*. 2008; 336: 1106-10.
- Parker LA, Chilet-Rosell E, Hernández-Aguado I, Pastor-Valero M, Gea S, Lumberras B. Diagnostic Biomarkers: Are We Moving from Discovery to Clinical Application? *Clin Chem*. 2018; 64: 1657-67.
- van Smeden M, Naaktgeboren CA, Reitsma JB, Moons KG, de Groot JA. Latent class models in diagnostic studies when there is no reference standard: a systematic review. *Am J Epidemiol*. 2013; 179: 423-31.
- Asselineau J, Paye A, Bessède E, Perez P, Proust-Lima C. Different latent class models were used and evaluated for assessing the accuracy of campylobacter diagnostic tests: overcoming imperfect reference standards? *Epidemiol Infect*. 2018; 146: 1556-64.
- Waikar SS, Betensky RA, Emerson SC, Bonventre JV. Imperfect gold standards for biomarker evaluation. *Clin Trials*. 2013; 10: 696-700.
- Pereira GdA, Louzada F, Barbosa VdF, Ferreira-Silva MM, Moraes-Souza H. A general latent class model for performance evaluation of diagnostic tests in the absence of a gold standard: an application to Chagas disease. *Comput Math Methods Med*. 2012; 2012: 487502.
- Keith JM, Davey CM, Boyd SE. A Bayesian method for comparing and combining binary classifiers in the absence of a gold standard. *BMC Bioinformatics*. 2012; 13: 179.
- Wang XN, Zhou V, Liu Q, Gao Y, Zhou X-H. Evaluation of the accuracy of diagnostic scales for a syndrome in Chinese medicine in the absence of a gold standard. *Chin Med*. 2016; 11: 35.
- Engel B, Buist W, Orsel K, Dekker A, de Clercq K, Grazioli S, et al. A Bayesian evaluation of six diagnostic tests for foot-and-mouth disease for vaccinated and non-vaccinated cattle. *Prev Vet Med*. 2008; 86: 124-38.
- Hur J, Cho EH, Baek KH, Lee KJ. Prediction of gestational diabetes mellitus by unconjugated estriol levels in maternal serum. *Int J Med Sci*. 2017; 14: 123-7.
- Choi YK, Johnson WO, Collins MT, Gardner IA. Bayesian inferences for receiver operating characteristic curves in the absence of a gold standard. *Journal of Agricultural, Biological, and Environmental Statistics*. 2006; 11: 210.
- Yu Y, Xie R, Shen C, Shu L. Effect of exercise during pregnancy to prevent gestational diabetes mellitus: a systematic review and meta-analysis. *Journal Matern Fetal Neonatal Med*. 2018; 31: 1632-7.
- Amirian A, Kariman N, Hedayati M, Borumandnia N, Naeje Z. Relationship between levels of maternal serum unconjugated estriol and gestational diabetes mellitus. *The Iranian Journal of Obstetrics, Gynecology and Infertility*. 2019; 21: 52-60.
- Sayin NC, Canda MT, Ahmet N, Arda S, Süt N, Varol FG. The association of triple-marker test results with adverse

- pregnancy outcomes in low-risk pregnancies with healthy newborns. *Arch Gynecol Obstet*. 2008; 277: 47-53.
15. Sancken U, Bartels I. Biochemical screening for chromosomal disorders and neural tube defects (NTD): is adjustment of maternal alpha-fetoprotein (AFP) still appropriate in insulin-dependent diabetes mellitus (IDDM)? *Prenat Diagn*. 2001; 21: 383-6.
  16. Kavak ZN, Basgul A, Elter K, Uygur M, Gokaslan H. The efficacy of first-trimester PAPP-A and free Beta hCG levels for predicting adverse pregnancy outcome. *J Perinat Med*. 2006; 34: 145-8.
  17. Enøe C, Georgiadis MP, Johnson WO. Estimation of sensitivity and specificity of diagnostic tests and disease prevalence when the true disease state is unknown. *Prev Vet Med*. 2000; 45: 61-81.
  18. Dendukuri N, Joseph L. Bayesian approaches to modeling the conditional dependence between multiple diagnostic tests. *Biometrics*. 2001; 57: 158-67.
  19. Garcia Barrado L, Coart E, Burzykowski T, Initiative AsDN. Development of a diagnostic test based on multiple continuous biomarkers with an imperfect reference test. *Stat Med*. 2016; 35: 595-608.
  20. Jafarzadeh SR, Johnson WO, Gardner IA. Bayesian modeling and inference for diagnostic accuracy and probability of disease based on multiple diagnostic biomarkers with and without a perfect reference standard. *Stat Med*. 2016; 35: 859-76.
  21. Henkelman RM, Kay I, Bronskill MJ. Receiver operator characteristic (ROC) analysis without truth. *Med Decis Making*. 1990; 10: 24-9.
  22. Branscum AJ, Johnson WO, Hanson TE, Gardner IA. Bayesian semiparametric ROC curve estimation and disease diagnosis. *Stat Med*. 2008; 27: 2474-96.
  23. Zhou XH, Castelluccio P, Zhou C. Nonparametric estimation of ROC curves in the absence of a gold standard. *Biometrics*. 2005; 61: 600-9.
  24. Hanson TE, Branscum AJ, Gardner IA. Multivariate mixtures of Polya trees for modeling ROC data. *Statistical Modelling*. 2008; 8: 81-96.25.
  25. Gonçalves L, Subtil A, de Oliveira MR, do Rosário V, Lee P-W, Shaio M-F. Bayesian latent class models in malaria diagnosis. *PLoS one*. 2012; 7: e40633.
  26. Gonçalves L, Subtil A, de Oliveira MR, do Rosário V, Lee P-W, Shaio M-F. Bayesian latent class models in malaria diagnosis. *PLoS One*. 2012; 7: e40633.
  27. Collins J, Huynh M. Estimation of diagnostic test accuracy without full verification: a review of latent class methods. *Stat Med*. 2014; 33: 4141-69.

# Application of Bayesian Latent Variable Model for Early Detection of Gestational Diabetes Mellitus without a Perfect Reference Standard Test by $\beta$ -Human Chorionic Gonadotropin

Amini M<sup>1</sup>, Kazemnejad A<sup>2</sup>, Zayeri F<sup>3</sup>, Amirian A<sup>4</sup>, Kariman N<sup>5</sup>

1- PhD Candidate of Biostatistics, Department of Biostatistics, Faculty of Medical Sciences, Tarbiat Modares University, Tehran, Iran

2- Professor of Biostatistics, Department of Biostatistics, Faculty of Medical Sciences, Tarbiat Modares University, Tehran, Iran

3- Professor of Biostatistics, Proteomics Research Center and Department of Biostatistics, Faculty of Allied Medical Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran

4- Instructor of Midwifery, Department of Midwifery, School of Nursing and Midwifery, Jiroft University of Medical Sciences, Jiroft, Iran

5- Associate Professor, Department of Midwifery and Reproductive Health, School of Nursing and Midwifery, Shahid Beheshti University of Medical Sciences, Tehran, Iran

**Corresponding author:** Kazemnejad A, kazem\_an@modares.ac.ir

(Received 25 February 2019; Accepted 27 July 2019)

**Background and Objectives:** Gestational diabetes mellitus (GDM) is a medical problem in pregnancy, and its late diagnosis can cause adverse effects in the mother and fetus. The purpose of this research was to estimate the accuracy parameters of a biomarker for early prediction of gestational diabetes in the absence of a perfect reference standard test.

**Methods:** This study was conducted in 523 pregnant women who presented to Mahdiah Hospital and Taleghani Hospital affiliated with Shahid Beheshti University of Medical Sciences, Tehran, Iran 2017-2018. As a predictor for detecting GDM, beta- human chorionic gonadotropin ( $\beta$ -hCG) measurements were recorded during 14-17th weeks' gestation in a checklist. The Bayesian latent variable model was used to estimate the sensitivity, specificity, and area under receiver operating characteristic curve (AUC). Bayesian parameter estimation was calculated using the R2OpenBUGS package in R version 3.5.3.

**Results:** The median gestational age was 33 years. In the absence of a perfect reference test, the applied model had a sensitivity, specificity, and AUC of 78% (95% credible interval (CrI): 0.66-0.83), 83% (95% CrI: 0.74-0.89), and 0.72 (95% CrI: 0.64-0.88) for  $\beta$ -hCG, respectively.

**Conclusion:** According to the results of this study,  $\beta$ -hCG may be an acceptable biomarker for early diagnosis of diabetes in pregnant women in the absence of a perfect reference test.

**Keywords:** Latent variable modeling, Receiver operating characteristic (ROC) curve, Bayesian approach, Gestational diabetes, Human chorionic gonadotropin