

مقایسه رگرسون لجستیک و تحلیل ممیزی در شناسایی عوامل تعیین کننده دیابت نوع ۲ در میان افراد پره‌دیابتی مناطق روستایی شهرستان کرمانشاه

اقبال زند کریمی^۱؛ علیرضا افشاری صفوی^{۱*}؛ منصور رضایی^۲؛ غضبان رجیبی^۳

چکیده

زمینه: عدم کنترل به موقع دیابت منجر به عوارض جبران‌ناپذیری در سایر بافت‌های بدن از جمله قلب، کلیه و چشم می‌گردد. با توجه به این‌که افراد پره‌دیابتی مستعدترین افراد برای ابتلا به این بیماری هستند، در این مطالعه به بررسی مهم‌ترین عوامل تعیین‌کننده ابتلا به دیابت در این گروه از افراد با استفاده از دو روش پیشرفته آماری رگرسون لجستیک و تحلیل ممیزی پرداخته شد.

روش‌ها: داده‌های این مطالعه از ۱۷ مرکز بهداشتی درمانی روستایی شهرستان کرمانشاه جمع‌آوری گردید. ۱۰۰ بیمار دیابتی و ۱۰۰ بیمار پره‌دیابتی وارد مطالعه شدند. اطلاعات جمعیت‌شناختی، شاخص توده بدن، قندخون ناشتا (FBS)، تست تحمل گلوکز خوراکی (OGTT)، فشارخون، چربی خون و میزان تحرک روزانه در دو چک لیست جداگانه گردآوری شد. مدل‌های رگرسون لجستیک و تحلیل ممیزی به منظور شناسایی عوامل تعیین‌کننده بر داده‌ها برازش و از نمودار راک جهت مقایسه قدرت پیش‌بینی مدل‌ها استفاده شد.

یافته‌ها: قدرت پیش‌بینی رگرسون لجستیک و تحلیل ممیزی به ترتیب برابر ۰/۸۸۴ و ۰/۸۰ برآورد شد. متغیرهای جنس ($P=0/027$) و FBS ($P<0/001$) در رگرسون لجستیک و متغیرهای سن ($P=0/014$)، FBS ($P<0/001$) و OGTT ($P<0/001$) در تحلیل ممیزی معنادار بودند. مدل رگرسون لجستیک با ۷۹ درصد از حساسیت بالاتری برخوردار بود. نتیجه‌گیری: در این مطالعه رگرسون لجستیک در تفکیک بیماران دیابتی از پره‌دیابتیک، قدرت بیشتری نشان داد. در جوامعی که گروه‌های مورد و کنترل قرابت زیادی دارند، کشف تفاوت‌های آن‌ها نیازمند روش‌های نیرومندتر است. بنابراین استفاده از این روش‌ها در مطالعات پزشکی توصیه می‌گردد.

کلیدواژه‌ها: رگرسون لجستیک، تحلیل ممیزی، دیابت نوع دو، پره‌دیابت، منحنی راک

«دریافت: ۱۳۹۱/۱۰/۲۵ پذیرش: ۱۳۹۲/۳/۲۱»

۱. گروه آمار زیستی، دانشکده بهداشت، دانشگاه علوم پزشکی کرمانشاه

۲. مرکز تحقیقات توسعه اجتماعی و ارتقای سلامت، دانشگاه علوم پزشکی کرمانشاه

۳. واحد پیشگیری و مبارزه با بیماری‌ها، مرکز بهداشت شهرستان کرمانشاه، دانشگاه علوم پزشکی کرمانشاه

* عهده‌دار مکاتبات: کرمانشاه، دانشگاه علوم پزشکی کرمانشاه، دانشکده بهداشت، گروه آمار زیستی و اپیدمیولوژی، تلفن: ۰۹۳۸۸۷۸۷۸۴۵

Email: ar8as12@gmail.com

مقدمه

پیشنهاد می‌کند که افراد چاق با بیشتر از ۱۲۰ درصد وزن ایده‌آل یا شاخص توده بدن (BMI) بیشتر یا مساوی ۲۷ کیلوگرم بر مترمربع، افرادی که اقوام درجه یک آنان دیابتی می‌باشند، جمعیت‌های در معرض خطر با سابقه تولد نوزاد با وزن بیش از ۴/۱ کیلوگرم، کسانی که فشارخون بالای ۹۰/۱۴۰ دارند و کسانی که HDL کم‌تر

براساس پیش‌بینی کارشناسان سازمان جهانی بهداشت، شیوع دیابت نوع ۲ در ایران در سال ۲۰۲۵ در حدود ۶/۸ درصد و برابر ۵۲۱۵۰۰۰ نفر خواهد بود و این تعداد در سال ۲۰۳۰ از مرز ۶/۴ میلیون نفر خواهد گذشت. کمیته بین‌المللی تشخیص و طبقه‌بندی دیابت

در مطالعه‌ای که با هدف مقایسه دو روش تحلیل ممیزی و آنالیز تشخیص لجستیک در پیش‌بینی عوامل تعیین‌کننده دیابت بر روی ۲۰۰۰ داده جمع‌آوری شده از مرکز تحقیقات و توانبخشی دیابت بنگلادش انجام گرفت، در مدل تحلیل ممیزی، منطقه، فعالیت بدنی، شاخص توده بدن (BMI)، سن و وراثت به‌عنوان عوامل تعیین‌کننده دیابت و در آنالیز ممیزی غیرخطی، جنسیت، سن، منطقه، فعالیت بدنی و BMI به‌عنوان عوامل خطر ساز در نظر گرفته شد. سطح تحصیلات در هر دو روش، تأثیری در شیوع دیابت نداشت (۹).

در مطالعه رحیمی و همکاران که بر روی ۱۷۲۰ مادر باردار به‌منظور تعیین شیوع دیابت حاملگی و عوامل خطر ساز مرتبط با آن در زنان باردار انجام گرفت، مشخص شد مادران دیابتی مسن‌تر و چاق‌تر بیش از مادران غیردیابتی، سابقه قبلی ابتلا به دیابت بارداری، سابقه خانوادگی دیابت، سقط، PCO، ماکروزومی و مرده‌زایی داشته‌اند (۱۰).

با توجه به شیوع روزافزون دیابت، عوارض متعدد این بیماری و عدم اطلاع کافی از ابتلا به بیماری در میان افراد جامعه، در این مطالعه به بررسی قدرت مدل‌های رگرسیون لجستیک و تحلیل ممیزی در تشخیص عوامل تعیین‌کننده دیابت از جمله BMI، FBS، تست تحمل گلوکز خوراکی (OGTT)، فشارخون (HT) و چربی خون در میان بیماران پره‌دیابتی پرداخته شد.

مواد و روش‌ها

مطالعه مقطعی حاضر به‌صورت تحلیلی-توصیفی در سال ۱۳۹۱ انجام گرفت. جامعه مورد مطالعه را بیماران دیابتی و پره‌دیابتی که دارای پرونده در مراکز بهداشت روستایی شهرستان کرمانشاه بودند تشکیل می‌دادند. حداقل حجم نمونه مورد نیاز با سطح اطمینان ۹۵ درصد و توان آماری ۹۰ درصد و بر اساس اطلاعات به‌دست‌آمده از مطالعات گذشته (۱۰)، ۷۲ نفر برای هر گروه تعیین شد که به‌منظور جلوگیری از کاهش توان

از ۳۵mg/dl دارند از نظر خطر ابتلا به دیابت غربال شوند. در این بین افراد پره‌دیابت مستعدترین افراد برای ابتلا به دیابت شناخته می‌شوند. در مطالعه آقای میبیدی و همکاران که بر روی ۳۹۵۷ فرد بالای ۲۰ سال به‌منظور بررسی اثر تغییر وزن آن‌ها بر میزان اختلالات قندخون انجام گرفته بود، پس از سه سال پیگیری ۱۴۸ نفر به دیابت مبتلا شده بودند که بیشتر آن‌ها از حالت پره‌دیابت به دیابت بود (۱). پره‌دیابت مرحله‌ای است که سطوح قندخون به بالاتر از حد نرمال رسیده ولی هنوز در حدی نیست که به آن دیابت اطلاق شود و بعضاً از آن به‌عنوان نقص تحمل گلوکز (IGT) یا نقص گلوکز ناشتا (IFG) نیز یاد می‌شود (۲). مطالعات نشان داده است زنانی که دارای سابقه دیابت حاملگی می‌باشند، بیشتر در معرض (IGT) و دیابت نوع ۲ می‌باشند (۳). در پره‌دیابت، قندخون دو ساعته ۱۹۹-۱۴۰ میلی‌گرم بر دسی‌لیتر می‌باشد. در صورتی که اگر این میزان ۲۰۰ میلی‌گرم بر دسی‌لیتر و یا بیشتر باشد، فرد مبتلا به دیابت خواهد بود.

با توجه به افزایش شیوع دیابت در جهان و ایران و صدمات جبران‌ناپذیر بیماری بر دیگر اعضای بدن از جمله چشم، قلب، کلیه و... و همچنین هزینه‌های اقتصادی و اجتماعی سنگینی که برای فرد و جامعه به‌دنبال دارد (۴)، می‌توان با تشخیص زودهنگام و کنترل این بیماری تا حد ممکن از عوارض ناشی از آن کاست. شناسایی عوامل تعیین‌کننده ابتلا به دیابت در میان افراد پره‌دیابت به‌دلیل نزدیکی پارامترهای بالینی این دو گروه نیازمند روش‌های آماری نیرومندتر می‌باشد. رگرسیون لجستیک و تحلیل ممیزی از مدل‌هایی هستند که کارایی آن‌ها به‌ویژه در مطالعات گوناگون پزشکی مورد تأیید قرار گرفته است (۵-۸).

در مطالعه‌ای که هیگنور و همکاران بر روی ۸۶۷ بیمار دیابتی و پره‌دیابتی انجام دادند، تحلیل رگرسیون لجستیک رابطه معناداری را بین بروز دیابت و پره‌دیابت با قندخون ناشتا، رژیم غذایی و سطوح فعالیت بدنی نشان داد (۸).

می‌شود. به طوری که متغیر با یک ضریب β در مدل تأثیر دارد.

تحلیل ممیزی: تحلیل ممیزی دارای تابعی به صورت زیر است:

$$y = b_0 + b_1x_1 + b_2x_2 + \dots + b_3x_3$$

که در آن مقادیر b_1, b_2, \dots مقدار ضریب تشخیصی یا وزن هر متغیر را نشان می‌دهد و معادل β در تحلیل رگرسیون خطی است. همانند تحلیل عاملی متغیرهایی در مدل باقی می‌مانند که در دو گروه معنادار و ضرایب تشخیصی بیشتر از $0/3$ داشته باشند (۱۳). تحلیل ممیزی تکنیک چندمتغیره است که با جدا کردن مجموعه‌های متمایز مشاهده‌ها و با تخصیص مشاهده‌های جدید به دسته‌های از پیش تعریف‌شده سر و کار دارد. مسأله آماری در مورد به وجود آوردن یک قانون (تابع تشخیص) بر مبنای اندازه‌های حاصل از افراد می‌باشد. با استفاده از این قانون می‌توان افراد جدید را که مشخص نیست به کدام جمعیت تعلق دارند به یکی از جمعیت‌ها منتسب کرد. از معروف‌ترین توابع مورد استفاده در تحلیل ممیزی می‌توان به تابع ممیز فیشر اشاره کرد. در روش دیگر، قاعده رده‌بندی با به حداقل رساندن تابع متوسط هزینه (ECM) به دست می‌آید (۱۴).

ابتدا به منظور دستیابی به یک دید کلی از مشاهدات از آمار توصیفی ساده و نمودارهای آماری استفاده گردید. نرمال بودن متغیرهای کمی سن و چربی خون با آزمون کولموگروف-اسمیرنوف بررسی شد. متغیر چربی خون پس از استفاده از تبدیل لگاریتم، نرمال شد. با توجه به مشاهده تعدادی داده گمشده در مطالعه از روش رگرسیونی نسبت به داده‌های گمشده متغیرهای کمی استفاده و در نهایت مطالعه با ۱۹۲ مشاهده اجرا گردید. همچنین متغیر میزان تحرک به سه رده کم‌تر از ۳۰ دقیقه، بین ۳۰-۶۰ دقیقه و بیشتر از ۶۰ دقیقه رده‌بندی شد. ابتدا از مدل رگرسیون لجستیک به منظور شناسایی عوامل تعیین‌کننده و ارزیابی دقت مدل در تفکیک بیماران دیابتی از پره‌دیابتی استفاده شد. در این مدل، انتخاب متغیرها

مطالعه در برخورد با داده‌های گمشده احتمالی، این مقدار برای هر گروه ۱۰۰ نفر در نظر گرفته شد. بنابراین ۱۰۰ بیمار دیابتی نوع ۲ به‌عنوان گروه مورد و ۱۰۰ پره‌دیابتی به‌عنوان گروه شاهد انتخاب شدند. دیابتی بودن بر اساس تشخیص پزشک و داروهای کاهش قندخون تجویزی به بیمار تعریف گردید (۱۱). با استفاده از روش نمونه‌گیری با تخصیص متناسب، حجم نمونه مورد نیاز هر یک از مراکز بهداشتی درمانی روستایی شهرستان کرمانشاه مشخص شد. ابزار جمع‌آوری داده‌ها دو چک لیست شامل اطلاعات جمعیت‌شناختی (سن و جنس)، شاخص توده بدن، قندخون ناشتا، تست تحمل گلوکز خوراکی، فشارخون، چربی خون و میزان تحرک روزانه بود که توسط کارکنان‌های مراکز روستایی از روی پرونده بیماران دیابتی (مورد) و پره‌دیابتی (شاهد) که به صورت تصادفی انتخاب می‌شدند، تکمیل گردید. آموزش‌های لازم جهت تکمیل چک‌لیست‌ها در جلسه‌ای در محل مرکز بهداشت شهرستان کرمانشاه به کارکنان‌ها به‌منظور جلوگیری از تورش در جمع‌آوری اطلاعات داده شد. پس از تکمیل چک‌لیست‌ها و ورود اطلاعات در کامپیوتر، تجزیه و تحلیل با استفاده از SPSS 16 صورت پذیرفت.

رگرسیون لجستیک: رگرسیون لجستیک یکی از ابزارهای آماری است که به منظور مدل‌سازی و تحلیل داده‌ها از آن استفاده می‌شود. شکل کلی این مدل به صورت زیر است (۱۲):

$$\text{Log} \left(\frac{\pi}{1-\pi} \right) = \alpha + \sum \beta x$$

χ بیانگر متغیرهای پیشگو است، که در اینجا فشارخون، چربی خون، BMI، جنسیت، OGTT، FBS و میزان تحرک را شامل می‌شود. β ضریب برآوردشده مدل برای متغیر مستقل مربوطه و π احتمال ابتلا یا عدم ابتلا به بیماری است که در این مطالعه به این صورت تعیین می‌شود که اگر فرد پره‌دیابتی مبتلا به دیابت شده باشد، این مقدار برابر یک و در غیر این صورت عدد صفر را می‌گیرد. اگر هر کدام از عوامل تعیین‌کننده در مقایسه دو گروه (مورد و شاهد) معنادار شد در رابطه بالا وارد

تعداد ۱۲۳ نفر (۶۴/۱٪) زن و ۶۹ نفر (۳۵/۹٪) مرد در مطالعه حضور داشتند. میانگین سنی در گروه پره‌دیابتی ۵۲/۱±۱۲/۲۰ سال و میانگین سنی در گروه بیماران ۵۵/۸۹±۱۱/۶۴ سال بود (جدول ۱).

مدل رگرسیون لجستیک برای چهار متغیر پیشگوی سن، FBS (با سه سطح)، BMI (با سه سطح) و میزان تحرک (با سه سطح) به داده‌ها برازش داده شد. سطح اول هر متغیر کیفی به‌عنوان طبقه مرجع برای تحلیل در نظر گرفته شد.

به‌روش Forward LR انجام گرفت. در ادامه از مدل تحلیل ممیزی به‌منظور کشف عوامل تعیین‌کننده مهم و ارزیابی دقت این روش استفاده شد. در این مدل نیز انتخاب متغیرها براساس روش Wilks' lambda انجام گرفت.

یافته‌ها

از ۱۹۲ نمونه مورد بررسی، ۹۲ نفر (۴۷/۹٪) پره‌دیابتی و ۱۰۰ نفر (۵۲/۱٪) دیابتی نوع ۲ بودند. از این

جدول ۱- فراوانی بیماران دیابتی و پره‌دیابتی به تفکیک متغیرهای پیشگو

پره دیابتی (درصد) فراوانی	دیابتی (درصد) فراوانی	عوامل تعیین کننده	
۲۷ (۲۹/۳)	۴۲ (۴۲)	مرد	جنسیت
۶۵ (۷۰/۷)	۵۸ (۵۸)	زن	
۲۲ (۲۳/۹)	۱۰ (۱۰)	<۱۰۰	FBS
۵۶ (۶۰/۹)	۱۶ (۱۶)	۱۲۵-۱۰۰	
۱۴ (۱۵/۲)	۷۴ (۷۴)	>۱۲۵	
۲۶ (۳۲/۱)	۱۴ (۱۴/۷)	<۱۴۰	OGTT
۳۸ (۴۶/۹)	۳۸ (۴۰)	۱۴۰-۲۰۰	
۱۷ (۲۱/۰)	۴۳ (۴۵/۳)	>۲۰۰	
۱۹ (۲۰/۷)	۳۰ (۳۰)	<۲۵	BMI
۴۹ (۵۳/۳)	۳۸ (۳۸)	۲۵-۳۰	
۲۴ (۲۶/۱)	۳۲ (۳۲)	>۳۰	
۵۶ (۶۰/۹)	۶۲ (۶۲)	<۳۰	تحرک
۲۷ (۲۹/۳)	۱۷ (۱۷)	۳۰-۶۰	
۹ (۹/۸)	۲۱ (۲۱)	>۶۰	
۴۳ (۴۶/۷)	۴۵ (۴۵)	<۱۲۰	HT
۳۶ (۳۹/۱)	۳۸ (۳۸)	۱۲۰-۱۴۰	
۱۳ (۱۴/۱)	۱۷ (۱۷)	>۱۴۰	

$$\text{logi}[p(y=1)]^{(1)} = ۱/۲۷۵ + ۰/۵۱(\text{Age}) - ۲/۱۵۷(\text{FBS}_1) - ۳/۵۱۱(\text{FBS}_2) - ۱/۱۶۴(\text{BMI}_2) - ۱/۸۱۹(\text{taha}r_1) - ۲/۳۱۳(\text{taha}r_2)$$

(FBS₁): میزان قند خون ناشتا بین ۱۰۰-۱۲۵
 (FBS₂): میزان قند خون ناشتا بیشتر از ۱۲۵
 (BMI₂): شاخص توده بدنی بیش از ۳۰
 (taha₁): میزان تحرک بین ۳۰ تا ۶۰ دقیقه
 (taha₂): میزان تحرک بیش از ۶۰ دقیقه

($P=0/895$) در مدل معنادار نبودند (جدول ۲). از طرفی حساسیت مدل رگرسیون ۷۹ درصد و اختصاصیت آن ۸۱/۱ درصد محاسبه گردید. همچنین قدرت مدل رگرسیون لجستیک در تفکیک بیماران دیابتی از پره‌دیابتی ۸۰/۱ درصد برآورد گردید (جدول ۳). در خصوص مدل تحلیل ممیزی، آماره آزمون لانداى ویلکز ($P<0/001$) معناداری مدل را نشان داد. همچنین آماره Box's M با ($P=0/568$) همگنی واریانس‌ها را در دو گروه تأیید کرد. تحلیل طی دو گام انجام گرفت که در انتهای گام دوم، متغیرهای جنس و FBS در مدل باقی ماندند. متغیرهای سن، OGTT، HT، BMI و میزان تحرک

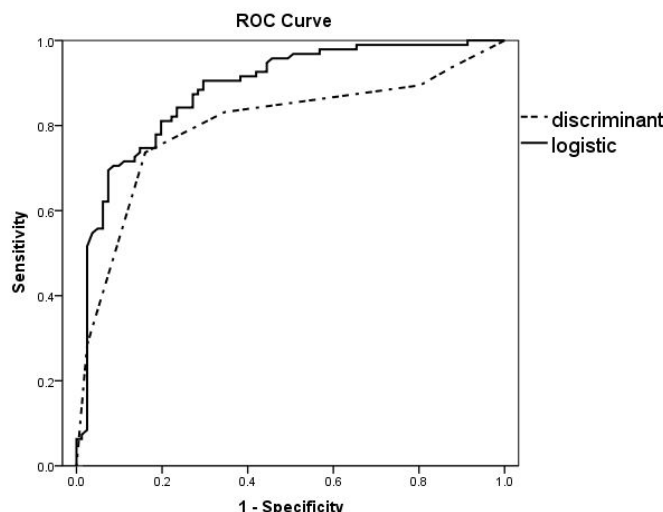
با برازش مدل رگرسیون لجستیک بر روی مشاهدات، متغیرهای پیشگو طی چهار مرحله وارد مدل شدند. آزمون هاسمر-لمشو ($P=0/016$) در مرحله آخر صحت برازش مدل را نشان داد. همچنین آماره (Nagelkerke R square=0.535) نشان می‌دهد که مدل تا چه اندازه توانسته تغییرات مربوط به متغیر وابسته (دیابتی-پره‌دیابتی) را تبیین کند. متغیرهایی که توسط مدل رگرسیون لجستیک به‌عنوان عوامل تعیین‌کننده شناخته شدند شامل سن، میزان تحرک، BMI و FBS بودند. متغیرهای جنس ($P=0/172$)، OGTT ($P=0/615$)، فشار خون ($P=0/781$) و چربی خون

جدول ۲- وضعیت متغیرهای پیشگو در مدل رگرسیون لجستیک

P value	انحراف استاندارد	آماره والد	ضرایب مدل	عوامل تعیین کننده
0/005	0/019	7/710	0/051	سن
<0/001	0	47/555	0	قندخون ناشتا
<0/001	0/581	13/789	-2/157	قندخون ناشتا (۱)
<0/001	0/516	46/250	-3/511	قندخون ناشتا (۲)
0/021	0	7/050	0	شاخص توده بدنی
0/866	0/582	0/029	0/099	شاخص توده بدنی (۱)
0/025	0/519	5/033	-1/164	شاخص توده بدنی (۲)
0/004	0	11/069	0	تحرک
0/004	0/629	8/374	-1/819	تحرک (۱)
0/001	0/713	10/520	-2/313	تحرک (۲)

جدول ۳- ارزش تشخیصی و درصد پیش‌بینی صحیح در دو مدل

درصد صحیح	پیش‌بینی شده		مشاهده شده	مدل
	دیابتی	پردیابتی		
0/79	17	64	پره‌دیابتی	رگرسیون لجستیک
81/1	77	18	دیابتی	
80/1	94	82	کل	
66/3	31	61	پره‌دیابتی	تحلیل ممیزی
83	83	17	دیابتی	
75	114	78	کل	



نمودار ۱- سطح زیرمنحنی راک برای دو مدل

بیماری نمی‌گردد. بلکه این بیماری معمولاً با اختلال تحمل گلوکز که اصطلاحاً مرحله پره‌دیابت و یا قبل از دیابت شناخته می‌شود، آغاز می‌گردد. پره‌دیابت که به‌عنوان عارضه عدم تحمل گلوکز نیز شناخته می‌شود، معمولاً هیچ علائم خاصی ندارد. این حالت اغلب قبل از ابتلاء جدی به دیابت نوع دوم بروز می‌کند. بیش از ۵۰ میلیون از افراد بالای ۲۰ سال در ایالات متحده آمریکا در مرحله پره‌دیابت هستند، یعنی قندخون آن‌ها بالاتر از سطح نرمال است، ولی در حدی بالا نیست که بتوان آن را دیابت در نظر گرفت. گرچه آمار دقیقی در این زمینه در ایران موجود نیست ولی به نظر می‌آید که در کشور ایران نیز افراد بسیاری مبتلا به پره‌دیابت هستند. مطالعات انجام‌شده در شهر تهران نشان‌دهنده ابتلا یک سوم از بالغین به دیابت و یا مرحله قبل از دیابت می‌باشد.

همواره پزشکان برای تشخیص به‌موقع پره‌دیابت، اهمیت بسیار قایلند چرا که معالجه و بهبود شرایط نامساعد در پره‌دیابت ممکن است از بروز مشکلات جدی در سلامت افراد پیشگیری کند. دورسی و همکاران نشان دادند گرچه افراد پره‌دیابت بیش از دیابتی، خواهان کنترل و کاهش وزن می‌باشند، با این حال میزان تحرک روزانه و کاهش میزان کالری و چربی در رژیم غذایی افراد دیابتی مشهودتر از پره‌دیابتی‌ها می‌باشد (۱۵).

بر اساس آماره لاندای ویلکز، مجوز ورود به مدل را نگرفتند. به‌عبارت دیگر این متغیرها نقشی در تفکیک بیماران پره‌دیابتی از دیابتی بازی نکردند. تابع ممیزی خطی فیشر برای تفکیک این دو گروه از بیماران به‌صورت زیر به‌دست آمد.

$$D = 0.356(\text{sex}) + 0.972(\text{GBS})$$

بر اساس این مدل می‌توان هر فرد جدید را با قرار دادن مقادیر خواسته‌شده در تابع فوق به یکی از گروه‌های دیابتی یا پره‌دیابتی منسوب کرد. منتهی قدرت این مدل در تفکیک بیماران دیابتی از پره‌دیابتی تنها ۷۵ درصد می‌باشد (جدول ۳). تحلیل ممیزی از حساسیت ۶۶/۳ درصدی و اختصاصیت ۸۳ درصدی برخوردار می‌باشد.

همچنین به‌منظور مقایسه بصری دو مدل رگرسیون لجستیک و تحلیل ممیزی از سطح زیرمنحنی راک استفاده گردید (نمودار ۱). این مقادیر برای دو مدل به‌ترتیب برابر ۰/۸۸۴ و ۰/۸۰ به‌دست آمد. بر اساس این نمودار نیز مدل رگرسیون لجستیک از قدرت بهتری در پیش‌بینی برخوردار بود.

بحث

دیابت از بیماری‌های مزمنی است که به‌دلیل اختلال در متابولیسم قند ایجاد می‌گردد. فرد یک‌روزه مبتلا به این

عوامل تعیین کننده بیماری دیابت مدت هاست در مطالعات و جوامع گوناگون مورد بررسی و تجزیه و تحلیل قرار گرفته است. با این حال تعداد مطالعاتی که این عوامل را در میان افراد پره دیابتی ارزیابی کند، به دلیل تشابه زیاد میان این گروه از افراد با بیماران دیابتی چندان زیاد و درخور توجه نبوده است. شناسایی این عوامل تعیین کننده در جوامعی که رفتاری تقریباً مشابه دارند نیازمند ابزار آماری دقیق و حساس می باشد. در مطالعه حاضر به مقایسه کارایی و صحت تشخیص مدل های پر کاربرد رگرسیون لجستیک و تحلیل ممیزی در شناسایی این عوامل پرداختیم.

در این مطالعه متغیرهای سن، FBS، BMI و تحرک در مدل رگرسیون لجستیک و متغیرهای جنس و FBS در مدل تحلیل ممیزی معنادار و در ساخت آن ها مورد استفاده قرار گرفتند. به عبارت دیگر مدل رگرسیون لجستیک نشان داد متغیرهای سن، قند خون ناشتا، شاخص توده بدنی و میزان تحرک روزانه در میان دیگر عوامل مورد بررسی در این مطالعه تنها عواملی هستند که قادرند پره دیابت را از دیابت متمایز گردانند. همچنین مدل تحلیل ممیزی نیز قادر است تنها از طریق جنسیت افراد و میزان قندخون ناشتا این تمایز را ایجاد نماید. با این حال این قدرت تفکیک در دو مدل چندان در خور توجه نمی باشد و برای رگرسیون لجستیک ۸۰/۱ درصد و برای تحلیل ممیزی ۷۵ درصد برآورد گردیده است.

این مطالعه همچنین نشان داد مدل رگرسیون لجستیک از حساسیت بیشتر (۷۹٪) و قدرت تشخیصی بالاتری (۸۰/۱٪) نسبت به مدل تحلیل ممیزی برخوردار است در حالی که معیار اختصاصیت در مدل ممیزی بیشتر (۸۳٪) بود.

در مطالعه ای که در سال ۱۳۸۰ در کردستان انجام گرفت عواملی چون چاقی، فشارخون، جنس زن و سن بالای ۴۰ سال از عوامل خطر ابتلا به بیماری دیابت نوع ۲ شناخته شدند (۱۶) که تقریباً با نتایج این مطالعه همخوانی دارد. تفاوت های موجود نیز می تواند به دلیل

اختلاف میان جامعه گروه شاهد دو مطالعه باشد. مطالعه ای که در سال ۲۰۰۵ در بنگلادش انجام گرفت، نشان داد که رگرسیون لجستیک در دسته بندی صحیح بیماران دیابتی دقیق تر از تحلیل ممیزی رفتار می کند (۹) که با نتایج مطالعه حاضر همخوانی دارد.

سدهی و همکاران در مطالعه ای تحت عنوان مقایسه مدل های شبکه عصبی مصنوعی با رگرسیون لجستیک و تحلیل ممیزی در پیش بینی سندروم متابولیک به مقایسه سه روش پرداختند. در این مطالعه تعداد ۳۴۷ نفر از افراد شرکت کننده در آن مطالعه که در ابتدای فاز اول مطالعه بر اساس تعریف AtpIII مبتلا به سندروم متابولیک نبودند، به عنوان نمونه انتخاب شدند. یافته ها نشان دادند که تحلیل ممیزی از دقت بیشتری نسبت به رگرسیون لجستیک در پیش بینی سندروم متابولیک برخوردار است (۱۷). یکی از دلایل آن می تواند به خاطر نرمال بودن متغیرهای پیشگو و همچنین استفاده از متغیرهای کیفی کم تر در مطالعه باشد. زیرا تحلیل ممیزی به شرط برقراری فرض نرمال بودن، معمولاً دقیق تر از رگرسیون لجستیک پیش بینی می کند (۱۳).

در مطالعه کاظم نژاد و همکاران در سال ۲۰۰۸ که به منظور مقایسه شبکه عصبی پرسپترون و رگرسیون لجستیک در بررسی متابولیسم قند بیماران دیابتی انجام گرفت، مدل شبکه عصبی پرسپترون بر اساس سطح زیر منحنی راک از قدرت بیشتری در پیش بینی مدل برخوردار بود (۱۸).

در مطالعه کالرا و همکاران در سال ۲۰۰۳ که بر روی ۳۲۶۸ مرد به منظور کشف سرطان پروستات انجام گرفت، شبکه عصبی با سطح زیر منحنی ۰/۸۲۵ قدرت بهتری در پیش بینی نسبت به مدل های رگرسیون لجستیک، تابع ممیزی خطی و تابع ممیزی درجه دو نشان داد (۱۹) $(P < 0/002)$.

در مطالعه لی و همکاران در سال ۲۰۱۲ بر روی ۲۷۴ بیمار دیابتی نوع ۲ به منظور پیش بینی نوروپاتی پیش از موعد از مدل های درخت تصمیم، پرسپترون چندلایه و

لحظه خطر این بیماری آن‌ها را تهدید می‌کند. در این مطالعه، مدل رگرسیون لجستیک نشان داد افزایش سن، میزان تحرک، میزان قند خون ناشتا و همچنین شاخص توده بدنی می‌تواند عوامل تعیین‌کننده مهمی برای دیابتی شدن باشد. همچنین براساس یافته‌های ما متغیرهای جنسیت زن و قندخون ناشتا در مدل تحلیل ممیزی دارای این خصوصیت هستند. به‌ویژه رگرسیون لجستیک در شناسایی این عوامل بهتر از تحلیل ممیزی رفتار نمود. این مطالعه نشان داد، علاوه بر متغیرهایی که حضور آن‌ها در مدل تأیید شد، متغیرهای دیگری نیز وجود دارند که می‌توانند پاسخ‌گوی اختلاف میان دو جامعه باشند. لذا پیشنهاد می‌گردد در مطالعات آینده عوامل تعیین‌کننده دیگر نیز مورد بررسی قرار گیرند. استفاده از مدل‌های آماری پیشرفته، قدرت شناسایی عوامل تعیین‌کننده را به‌ویژه در جوامعی که بسیار مشابه هستند فراهم می‌آورد.

References

1. Aghaie Meybodi HR, Azizi F. [The assessment of relation between lipid distribution and weight change with diabetes incidence in a group of Tehran, district 13 population (Persian)]. *Pejouhesh dar Pezeshki*. 2008;32(2):105-13.
2. Wikner C, Gigante B, Hellenius M, Faire U, Leander K. The risk of type 2 diabetes in men is synergistically affected by parental history of diabetes and overweight. *PLoS One*. 2013;8(4):e61763.
3. Chew WF, Rokiah P, Chan SP, Chee WSS, Lee LF, Chan YM. Prevalence of glucose intolerance, and associated antenatal and historical risk factors among Malaysian women with a history of gestational diabetes mellitus. *Singapore Med J*. 2012;53(12):814-20.
4. Zhao X, Zhu X, Zhang H, Zhao W, Li J, Shu Y, et al. Prevalence of diabetes and predictions of its risks using anthropometric measures in southwest rural areas of China. *BMC Public Health*. 2012;12:821.
5. Guan P, Huang D, He M, Zhou B. Lung cancer gene expression database analysis incorporating prior knowledge with support vector machine-based classification method. *J Exp Clin Cancer Res*. 2009;28:103.
6. Harrison LCV, Luukkaala T, Petrovaara H, Saarinen TO, Heinonen TT, Jarvenpaa R, et al. Non-Hodgkin lymphoma response evaluation with MRI texture classification. *J Exp Clin Cancer Res*. 2009;28:87.
7. Wong CY, Gannon J, Bong J, Wong CO, Saha GB. Computer-assisted lateralization of unilateral temporal lobe epilepsy using Z-score parametric F-18 FDG PET images. *BMC Nucl Med*. 2007;7:5.
8. Hightower JD, Hightower CM, Vázquez BY, Intaglietta M. Incident prediabetes/diabetes and blood pressure in urban and rural communities in the Democratic Republic of Congo. *Vasc Health Risk Manag*. 2011;7:483-9.
9. Sultana F. Discriminant analysis and logistic discrimination: an application to diabetes mellitus data 2000. PhD. thesis in biostatistics. Institute of Statistical Research and Trainings, University of Dhaka. 2005.
10. Rahimi M, Dinari J, Najafi F. [The prevalence of gestational diabetes and its associated Determinants in pregnant women in Kermanshah (Persian)]. *Journal of Kermanshah University of Medical Sciences*. 2010;14(3):244-50.
11. Anjana RM, Pradeepa R, Deepa M, Datta M, Sudha V, Unnikrishnan R, et al. Prevalence of diabetes and prediabetes (impaired fasting glucose and/or impaired glucose tolerance) in urban and rural India: phase I results of the Indian Council of Medical Research-INDIA DIABetes (ICMR-INDIAB) study. *Diabetologia*. 2011;54(12):3022-7.

رگرسیون لجستیک استفاده شد. مدل شبکه عصبی از دقت تشخیصی بیشتری (۸۳٪) برخوردار بود. سطح زیر منحنی راک برای درخت تصمیم (۰/۸۸۶۳) و رگرسیون لجستیک (۰/۸۸۰۲) بیشتر از شبکه عصبی (۰/۸۵۳۶) بود (۲۰).

مدل‌های شبکه عصبی معمولاً در حالت‌هایی که روابط بین متغیرها بسیار پیچیده است و با حجم عظیمی از اطلاعات روبرو هستیم بهتر از روش‌های سنتی عمل می‌کند. با این حال تحلیل نتایج در مدل‌هایی چون رگرسیون لجستیک و تحلیل ممیزی راحت‌تر و قابل‌فهم‌تر می‌باشد.

نتیجه‌گیری

مطالعاتی که تاکنون انجام گرفته کم‌تر به بررسی وضعیت بیماران پره‌دیابتی پرداخته است. حال آن‌که این افراد نزدیک‌ترین فاصله را با فاز دیابتی شدن دارند و هر

12. Danesh-Pour MS, Mehrabi Y, Hedayati M, Azizi F. [Multivariable survey of factors correlated with metabolic syndrome using factor analysis (Persian)]. *Iranian Journal of Endocrinology and Metabolism*. 2006;30:139-46.
13. Habibpour K, Safari R. [A comprehensive guide to using the SPSS in survey research (Persian)]. 4th ed. Tehran: 2011; 816, 836.
14. Jobson DJ. *Applied multivariate data analysis*. 1st ed. New York : Springer-Verlag 1991.
15. Dorsey R, Songer T. Lifestyle behaviors and physician advice for change among overweight and obese adults with prediabetes and diabetes in the United States, 2006. *Prev Chronic Dis*. 2011;8(6):A132.
16. Bidarpour F, Holakouie K, Rahimi A, Esmailnasab N. [Evaluation of determinants for type 2 diabetes in the Diabetes Center of Kurdistan (Persian)]. *Journal of Kurdistan University of Medical Sciences*. 2003;7(26):15-20.
17. Sedehi M, Mehrabi Y, Kazemnejad A, Hadaegh F. [Comparison of neural network models with logistic regression and discriminant analysis in predicting the metabolic syndrome (Persian)]. *Iranian Journal of Endocrinology and Metabolism*. 2009;11(6):638-46.
18. Kazemnejad A, Batvandi Z, Faradmal J. Comparison of artificial neural network and binary logistic regression for determination of impaired glucose tolerance/diabetes. *East Mediterr Health J*. 2010;16(6):615-20.
19. Kalra P, Togami J, Bansal B S G, Partin AW, Brawer MK, Babaian RJ, et al. A neurocomputational model for prostate carcinoma detection. *Cancer*. 2003;98(9):1849-54.
20. Li CP, Zhi XY, Ma J, Cui Z, Zhu ZL, Zhang C, et al. Performance comparison between Logistic regression, decision trees, and multilayer perceptron in predicting peripheral neuropathy in type 2 diabetes mellitus. *Chin Med J (Engl)*. 2012;125(5):851-7.