

استفاده از شبکه‌های عصبی مصنوعی و مقایسه با روشهای رگرسیونی در

پیش‌بینی تراکم استخوان زنان یائسه ایرانی

دکتر اکبر سلطانی*، دکتر محسن سادات صفوی*، دکتر علیرضا معیری*، دکتر مهدی نورایی*

دکتر آرش حسین نژاد*، دکتر باقر لاریجانی*

* دانشگاه علوم پزشکی و خدمات بهداشتی درمانی تهران، مرکز تحقیقات غدد درون‌ریز و متابولیسم

چکیده

به تازگی استفاده از شبکه‌های هوش مصنوعی (ANN) در پیش‌بینی اثرات متغیرهای متعدد و با روابط پیچیده بر روی یک متغیر خاص مورد توجه قرار گرفته است. عوامل محیطی و فیزیولوژیک متعددی در تخمین تراکم استخوان افراد دخیل دانسته شده‌اند. در این مطالعه، قدرت ANN در پیش‌بینی وضعیت تراکم استخوان زنان یائسه ایرانی با روشهای قدیمی مقایسه شد. ۲۱۵۸ خانم یائسه ارجاع شده به مرکز تراکم سنجش استخوان بیمارستان دکتر شریعتی به طور تصادفی به سه گروه training (۱۴۰۰ نفر)، Validation (۱۵۰ نفر) و test (۶۰۸ نفر) تقسیم شدند. ۱۰ متغیر سن، وزن، سن در هنگام یائسگی، استفاده از کورتیکواستروئیدها، استفاده از استروژن، تعداد حاملگی، سن در هنگام منارک، قد، میزان فعالیت و سیگار کشیدن در این افراد بررسی شدند. مدل‌های Robust رگرسیون خطی چند متغیره و ANN در گروه اول ساخته و در گروه دوم Validate شدند و در گروه سوم تفاوت بین متغیرهای نتیجه (T score) نواحی گردن فمور و مهره‌های کمری و مقادیر پیش‌بینی شده توسط دو مدل مقایسه شدند. برای تخمین قدرت پیش‌بینی استئوپروز از منحنی ROC با حد برابر ۲/۵- استفاده شد. مدل‌های ANN با ۴ متغیر و بیشتر در پیش‌بینی T score ناحیه گردن فمور و با ۵ متغیر و بیشتر در پیش‌بینی T score ناحیه کمری از مدل رگرسیون قوی‌تر بودند. این تفاوت نشانگر قدرت بیشتر مدل‌های مبتنی بر شبکه‌های هوش مصنوعی در پیش‌بینی موارد استئوپروز با استفاده از ریسک فاکتورهای مختلف و لزوم استفاده از این فن‌آوری در موارد غربالگری جمعیتی می‌باشد. (سال پنجم، شماره ۴، زمستان ۱۳۸۲، ص ۲۰۵ تا ۲۱۳)

کلواژه‌ها: شبکه‌های هوش مصنوعی، تراکم معدنی استخوان، زنان یائسه

مقدمه

زمینه‌ای در گروه‌های مختلف افراد پیشنهاد شده است،^(۵-۸) اما هیچ‌کدام از این روشها نتوانسته‌اند تأثیر مثبت خود را در برنامه‌های غربالگری به اثبات برسانند.^(۹) به علت گرانی و در دسترس نبودن وسائل اندازه‌گیری BMD استفاده از روشهای غربالگری جمعیتی برای این مورد امکان‌پذیر نمی‌باشد.^(۱۰)

شبکه‌های مصنوعی (ANN) ساختارهای پردازش اطلاعاتی جدیدی هستند که از روشهای مخصوص شبکه‌های عصبی بیولوژیک استفاده می‌کنند.^(۱۱) تفاوت اصلی ANN با سایر روشهای متداول در این است که در این روش نیازی به وجود

استئوپروز بیماری است که با شکستگی‌ها و دفورمیتی‌های پاتولوژیک استخوانها همراه است و داروهای جدید برای افزایش تراکم معدنی استخوان (BMD) و برگشت‌پذیری این بیماری پزشکان را به فکر استفاده از روش‌های پروفیلاکتیک برای جلوگیری از کاهش BMD انداخته است.^(۱۰) BMD موثرترین فاکتور پیش‌بینی کننده شکستگی‌های پاتولوژیک است^(۳) و حد T score کمتر از ۲/۵- انحراف معیار از جمعیت جوان به عنوان حد بیماری استئوپروز شناخته شده است.^(۴) روش‌های مختلفی برای پیش‌بینی BMD بر اساس متغیرهای

رضایت‌نامه مربوط به استفاده از اطلاعات شخصی را امضاء کردند و کمیته اخلاق پزشکی مرکز تحقیقات غدد و متابولیسم دکتر شریعتی این مطالعه را تأیید کرد.

شرکت کنندگان به طور تصادفی به سه گروه آموزش (training)، معتبرسازی (Validation) و آزمون (testing) تقسیم شدند. متغیرهای ورودی بر اساس مطالعات قبلی^(۱۵) و همچنین آنالیز مقدماتی بر روی داده‌های این مطالعه انتخاب شدند. آنها شامل ۱۰ متغیر سن، وزن، سن در هنگام یائسگی، استفاده از کورتیکواستروئیدها، استفاده از استروژن، تعداد حاملگی، سن در هنگام منارک، قد، میزان فعالیت و استفاده از سیگار بودند. متغیرهای خروجی T score نواحی گردن فمور و ستون فقرات کمری (L_1 تا L_4) بودند. برای بررسی قدرت پیش‌بینی مدل‌های مختلف به وسیله تعداد متفاوت متغیرهای ورودی، این متغیرها بر اساس ترتیب نزولی تأثیر بر روی متغیرهای خروجی در مدل رگرسیون قدم به قدم (stepwise) مرتب شدند. برای افزایش دقت آنالیز رگرسیون، مواردی که مقادیر غیرطبیعی برای متغیرهای مختلف داشتند و یا تأثیر زیادی بر روی نتیجه مدل می‌گذاشتند شناسایی شده و از آنالیز حذف شدند. برای شناسایی این موارد در متغیرهای اسمی از Box plot و در متغیرهای عددی از روش‌های leverage استفاده شد. روش وزندهی loess برای بدست آوردن یک مدل رگرسیون Robust و مقاوم به outlier استفاده شد.^(۱۶)

برای تولید شبکه‌های هوش مصنوعی، ابتدا پرسپترون‌های چند لایه (MLP) با یک یا دو لایه مخفی و دو نورون در لایه خارجی در گروه اول آموزش داده شدند. پس از بدست آوردن موثرترین ترکیب ممکن برای پیش‌بینی متغیرهای خروجی، از آن ترکیب در بقیه آنالیزها استفاده شد. آموزش شبکه در هنگامی که خطای پیش‌بینی در گروه دوم با استفاده از خطای مربع میانگین به حداقل رسید متوقف شد.^(۱۷)

در قدم بعدی قدرت مدل‌های ANN و رگرسیون در گروه

رابطه جبری قابل کشف بین متغیرهای اولیه و ثانویه وجود ندارد. در عوض، این شبکه‌ها خود رابطه منطقی بین متغیرها را طی روند training کشف می‌کنند.^(۱۲) شبکه‌های هوش مصنوعی به خصوص در مواردی که رابطه‌های پیچیده و غیرخطی بین متغیرهای اولیه و ثانویه وجود دارد قدرت خود را نشان می‌دهند.^(۱۳)

ساختار استخوانی بدن انسان تحت ساخت و ساز مداوم در طول دوران زندگی است و عوامل مختلفی بر این فرآیند موثر می‌باشند.^(۱۴) به علت احتمال غیرخطی بودن رابطه این متغیرها، استفاده از ANN برای پیش‌بینی وضعیت تراکم استخوان می‌تواند بر روش‌های خطی برتری داشته باشد. در این مطالعه ما به بررسی قدرت روش ANN و مقایسه آن با روش‌های متداول قدیمی برای پیش‌بینی BMD در زنان یائسه ایرانی پرداخته‌ایم.

روش کار

شرکت کنندگان در این مطالعه ۲۳۴۶ خانم یائسه بررسی شده در مرکز سنجش تراکم استخوان بیمارستان دکتر شریعتی در فاصله سالهای ۱۳۷۹ تا ۱۳۸۲ بودند. برخی از این افراد موارد مشکوک به استئوپروز بودند که توسط پزشکان مختلف برای بررسی بیشتر و تأیید تشخیص ارجاع شده بودند و گروه دیگر زنان یائسه سالمی بودند که در این مرکز تحت غربالگری قرار گرفته بودند. BMD نواحی ستون فقرات کمری و فمور با استفاده از روش DXA و با دستگاه دانسیتومتر لونا DPXMD و به وسیله یک اپراتور آموزش دیده طبق توصیه‌های کمپانی سازنده اندازه‌گیری شد. دستگاه دانسیتومتر بوسیله فانتوم‌های مناسب و به صورت هفتگی کالیبره می‌شد.

یک پرسشنامه استاندارد شده (پرسشنامه NOF^(۱۵)) برای کلیه افراد پیش از انجام BMD تکمیل شد. اطلاعات مربوط به بیماران استئوپروتیک با تشخیص قبلی و سابقه استفاده از داروهای محرک استخوان‌سازی (مانند آلدرونا، کلسیتونین و فلوراید) از مطالعه حذف شد. کلیه افراد تحت مطالعه فرم

جدول ۳ سطح زیر منحنی ROC برای مدل‌های مختلف را نمایش می‌دهد. مدل‌های ANN با ۵ متغیر و بیشتر برای پیش‌بینی استئوپروز در ناحیه کمری و با ۳ متغیر و بیشتر برای پیش‌بینی استئوپروز گردن فمور از مدل‌های رگرسیون خطی قوی‌تر بودند.

بحث

محققین معدودی عملکرد ANN را در پیش‌بینی BMD بررسی کرده‌اند. در مطالعه‌ای که به وسیله دکتر Rae و همکاران انجام شد، مدل‌های ANN و رگرسیون لجستیک برای تشخیص استئوپروز بوسیله اولتراسونوگرافی کمیته در ۲۷۴ فرد مشکوک مقایسه شدند.^(۲۲) در یک مطالعه دیگر نیز عملکرد ANN در پیش‌بینی استئوپروز وابسته به محل در افراد سالم و بیماران (با بیماری شناخته شده از نظر تأثیرگذاری بر روی تراکم استخوان) بر روی ۱۱۵۲ فرد بررسی شد.^(۲۳) در هر دوی این مطالعات نیز ANN بر روش‌های رگرسیون برتری داشت و محققین استفاده از این روش را برای موارد مشابه توصیه کردند. شباهت عملکرد مدل‌های ANN و رگرسیونی در تعداد کمتر متغیرها می‌تواند به علت ترتیب ورود متغیرها به مدل‌ها بر اساس نتایج آنالیز رگرسیون قدم به قدم باشد. با وجودیکه این روش یک روش استاندارد برای ورود متغیرها به مدل‌های رگرسیونی است، لزوماً این ترتیب معادل ترتیب ارزش متغیرها در مدل ANN نیست. در حقیقت هیچ روش استاندارد مشابهی برای مدل‌های ANN وجود ندارد. حتی در بعضی مطالعات اثبات شده است که کم کردن بعضی از متغیرها می‌تواند باعث بهبود عملکرد شبکه‌ها شود.^(۲۱)

در این مطالعه نتایج آنالیز منحنی‌های ROC به طور دقیق نتایج آنالیز کای دو غیر متمرکز را تکرار نکردند. این می‌تواند به علت cutoff در نظر گرفته شده (T score برابر ۲/۵-) و توزیع داده‌های مطالعه ما باشد. با این وجود، از دیدگاه بالینی، تشخیص صحیح ۸۷ درصد موارد استئوپروز کمری و ۷۷ درصد موارد

سوم مقایسه شد. مقایسه تفاوت بین مقادیر پیش‌بینی شده توسط دو مدل و مقادیر واقعی متغیرهای خروجی در هر بیمار با استفاده از روش تفاوت مربع میانگین‌ها انجام شد. این آمار از توزیع کای دو غیر مرکزی (non-centric chi-square) تبعیت می‌کند. علاوه بر این برای بدست آوردن نتایج ملموس‌تر از نظر بالینی، قدرت تشخیص مدل‌ها برای جدا کردن بیماران استئوپروتیک (دارای T score کمتر از ۲/۵-) از افراد نرمال بوسیله آنالیز منحنی ROC انجام شد. برای این منظور تفاوت سطح زیر منحنی در مدل‌های مختلف با یکدیگر مقایسه شد. برای آموزش شبکه‌های ANN از نرم‌افزار MATLAB (مدل ۶/۵، ۲۰۰۲) و برای آنالیزهای آماری از نرم‌افزار STATA (مدل ۸، ۲۰۰۱) استفاده شد.

یافته‌ها

از میان کلیه شرکت‌کنندگان در مطالعه، تعداد ۲۰۶ مورد به علت کمبود اطلاعات پایه (۳۴ نفر) و استفاده از داروهای ضد استئوپروز (۱۷۲ نفر) از مطالعه حذف شدند. تعداد ۲۱۵۸ خانم باقی‌مانده به صورت تصادفی در سه گروه آموزش (۱۴۰۰ نفر)، معتبرسازی (۱۵۰ نفر) و آزمون (۶۰۸ نفر) تقسیم شدند. علاوه بر این ۵۵ نفر به علت مشکلات مربوط به outlier از آنالیز رگرسیون در گروه اول حذف شدند. متغیرهای ورودی و خروجی مدل‌ها در گروه‌های مختلف در جدول ۱ نمایش داده شده است. تفاوت معنی‌داری در هیچ یک از متغیرها بین گروه‌های مختلف وجود نداشت. در جدول ۲ تفاوت بین مقادیر پیش‌بینی شده و واقعی T score نواحی مختلف بوسیله دو مدل نشان داده شده است. برای پیش‌بینی BMD ناحیه کمری، مدل‌های ANN با سه متغیر و بیشتر به طور معنی‌داری از مدل‌های رگرسیون قوی‌تر بودند. برای پیش‌بینی T score ناحیه گردن فمور، مدل‌های ANN با ۴ متغیر و بیشتر برتری خود را بر مدل‌های رگرسیون نشان دادند.

جدول ۱: متغیرهای ورودی و خروجی برای گروه‌های سه‌گانه استفاده شده در آموزش، معتبرسازی و آزمون مدل‌های شبکه‌های عصبی و رگرسیونی

مقادیر عددی [†]			نوع متغیر	نام متغیر (واحد)*
گروه آموزش (۱۴۰۰ نفر)	گروه معتبرسازی (۱۵۰ نفر)	گروه آزمون (۶۰۸ نفر)		
۵۸/۶(۸/۵)	۵۸/۷(۸/۳)	۵۹/۳(۸/۲)	پیوسته عددی	سن (سال)
۶۷/۱(۱۲/۵)	۶۷/۳(۱۲/۴)	۶۶/۸(۱۲/۴)	پیوسته عددی	وزن (کیلوگرم)
۴۶/۸(۵/۹)	۴۶/۹(۶/۰)	۴۷/۳(۵/۸)	پیوسته عددی	سن در هنگام یائسگی (سال)
۵۴۶(۸۹/۸) ۶۲(۱۰/۲)	۱۳۰(۸۶/۷) ۲۰(۱۳/۳)	۱۲۶۷(۹۰/۵) ۱۳۳(۹/۵)	اسمی رتبه‌ای ۰ = عدم استفاده (در ۳ ماه اخیر) ۱ = استفاده (در سه ماه اخیر)	استفاده از کورتیکواستروئیدها
۵۷۳(۹۴/۲) ۳۵(۵/۸)	۱۴۲(۹۴/۶) ۸(۵/۴)	۱۳۲۷(۹۴/۸) ۷۳(۵/۲)	اسمی رتبه‌ای ۰ = عدم استفاده (در ۳ ماه اخیر) ۱ = استفاده (در سه ماه اخیر)	استفاده از استروژن‌ها
۵/۹(۳/۳)	۵/۶(۳/۳)	۵/۶(۳/۱)	گسسته عددی	تعداد حاملگی‌ها
۱۳/۷(۱/۵)	۱۳/۲(۱/۴)	۱۳/۶(۱/۴)	پیوسته عددی	سن در هنگام منارک (سال)
۱۵۵/۱(۶/۸)	۱۵۳/۶(۷/۱)	۱۵۴/۹(۸/۱)	پیوسته عددی	قد (سانتی‌متر)
۳۰۸(۵۰/۷) ۳۰۰(۴۹/۳)	۷۲(۴۸/۰) ۷۶(۵۲/۰)	۶۶۷(۴۷/۷) ۷۳۳(۵۲/۳)	اسمی رتبه‌ای ۰ = کمتر از ۱۵ دقیقه در روز ۱ = بیشتر از ۱۵ دقیقه در روز	میزان فعالیت
۵۹۷(۹۸/۲) ۱۱(۱/۸)	۱۴۶(۹۷/۳) ۴(۲/۷)	۱۳۶۵(۹۷/۵) ۳۵(۲/۵)	اسمی رتبه‌ای ۰ = عدم استفاده (در ۱ سال اخیر) ۱ = استفاده (در ۱ سال اخیر)	وضعیت سیگار کشیدن
-۱/۷۸ (۱/۴۷)	-۱/۸۴ (۱/۳۹)	-۱/۸۱ (۱/۳۷)	پیوسته عددی	T score در ناحیه کمری [‡]
-۱/۱۵ (۱/۲۱)	-۱/۲۴ (۱/۱۶)	-۱/۲۰ (۱/۲۰)	پیوسته عددی	T score در ناحیه گردن فمور [‡]

* متغیرها به ترتیب نزولی بر اساس همکاری در واریانس تخمین متغیرهای خروجی در آنالیز رگرسیون قدم به قدم مرتب شده‌اند.

† اعداد به صورت میانگین (انحراف معیار) برای متغیرهای عددی و به صورت شیوع (درصد) برای متغیرهای اسمی نمایش داده شده‌اند.

‡ متغیرهای ریفرانس (ارجاعی)

جدول ۲: مقایسه تفاوت بین مقادیر متخیرهای ریفرانس و مقادیر پیش‌بینی شده بوسیله مدل‌های رگرسیونی و شبکه‌های

عصبی مصنوعی (ANN)*

T score گردن فمور			T score کمری			تعداد متغیرها
P value	رگرسیون	ANN	P value	رگرسیون	ANN	
۰/۲۱۹	-۰/۰۱ ± ۱/۰۳	۰/۰۳ ± ۱/۰۲	۰/۷۶۷	۰/۰۲ ± ۱/۲۹	-۰/۰۲ ± ۱/۲۹	۲
۰/۰۷۳	۰/۰۲ ± ۱/۰۱	۰/۰۲ ± ۰/۹۸	۰/۰۴۸ [‡]	۰/۰۲ ± ۱/۲۵	۱/۰۱ ± ۱/۲۰	۳
۰/۰۳۸ [‡]	۰/۰۱ ± ۰/۰۹۸	۰/۰۲ ± ۰/۹۲	۰/۰۳۲ [‡]	۰/۰۴ ± ۱/۲۵	۰/۰۲ ± ۱/۱۸	۴
۰/۰۳۲ [‡]	۰/۰۲ ± ۰/۹۶	-۰/۰۱ ± ۰/۹۰	۰/۰۳۱ [‡]	-۰/۰۳ ± ۱/۲۴	-۰/۰۳ ± ۱/۱۷	۵
۰/۰۲۳ [‡]	۰/۰۲ ± ۰/۹۵	-۰/۰۱ ± ۰/۸۸	۰/۰۲۵ [‡]	-۰/۰۲ ± ۱/۲۴	۰/۰۰ ± ۱/۱۵	۶
۰/۰۱۹ [‡]	-۰/۰۴ ± ۰/۹۴	۰/۰۲ ± ۰/۸۵	۰/۰۲۲ [‡]	۰/۰۳ ± ۱/۲۳	-۰/۰۲ ± ۱/۱۵	۷
۰/۰۱۸ [‡]	-۰/۰۲ ± ۰/۹۳	-۰/۰۳ ± ۰/۸۳	۰/۰۲۲ [‡]	-۰/۰۲ ± ۱/۲۲	۰/۰۱ ± ۱/۱۴	۸
۰/۰۱۶ [‡]	۰/۰۰ ± ۰/۹۲	۰/۰۱ ± ۰/۸۱	۰/۰۱۶ [‡]	۰/۰۰ ± ۱/۲۱	-۰/۰۲ ± ۱/۱۱	۹
۰/۰۱۸ [‡]	۰/۰۱ ± ۰/۹۳	۰/۰۰ ± ۰/۸۲	۰/۰۱۷ [‡]	۰/۰۱ ± ۱/۲۳	-۰/۰۱ ± ۱/۱۱	۱۰

* نتایج به صورت میانگین ± انحراف معیار تفاوت بین مقادیر پیش‌بینی شده و ریفرال نمایش داده شده‌اند.

† مقادیر P بر اساس مقایسه جفتی مربع میانگین تفاوت بین مقادیر پیش‌بینی شده و ریفرال (تست کای دو غیر متمرکز) محاسبه شده‌اند.

‡ نتایج تفاوت معنی‌دار در سطح $P = ۰/۰۵$

جدول ۳: سطح زیر نمودار برای مدل‌های رگرسیونی و شبکه‌های عصبی مصنوعی (ANN)

ناحیه گردن فمور			ناحیه کمری			تعداد متغیرها
P value	رگرسیون	ANN	P value	رگرسیون	ANN	
۰/۵۵۷	۰/۷۹۹ ± ۰/۰۲۱	۰/۸۰۱ ± ۰/۰۲۱	۰/۳۲۷	۰/۷۰۹ ± ۰/۰۲۱	۰/۷۰۹ ± ۰/۰۲۱	۲
۰/۰۴۸ [*]	۰/۸۰۹ ± ۰/۰۱۷	۰/۸۱۸ ± ۰/۰۱۹	۰/۱۳۶	۰/۷۲۱ ± ۰/۰۲۰	۰/۷۲۱ ± ۰/۰۲۰	۳
۰/۰۲۴ [*]	۰/۸۱۷ ± ۰/۰۲۰	۰/۸۳۱ ± ۰/۰۱۶	۰/۰۸۳	۰/۷۳۰ ± ۰/۰۱۶	۰/۷۳۶ ± ۰/۰۱۵	۴
۰/۰۱۶ [*]	۰/۸۲۶ ± ۰/۰۱۸	۰/۸۳۹ ± ۰/۰۱۷	۰/۰۳۵ [*]	۰/۷۳۷ ± ۰/۰۱۷	۰/۷۴۹ ± ۰/۰۱۷	۵
۰/۰۱۰ [*]	۰/۸۳۲ ± ۰/۰۲۱	۰/۸۴۷ ± ۰/۰۱۵	۰/۰۱۴ [*]	۰/۷۴۱ ± ۰/۰۱۷	۰/۷۵۶ ± ۰/۰۱۵	۶
۰/۰۰۸ [*]	۰/۸۳۴ ± ۰/۰۲۰	۰/۸۵۵ ± ۰/۰۱۶	۰/۰۱۰ [*]	۰/۷۴۳ ± ۰/۰۱۹	۰/۷۶۴ ± ۰/۰۱۸	۷
۰/۰۰۳ [*]	۰/۸۳۴ ± ۰/۰۱۹	۰/۸۶۰ ± ۰/۰۱۴	۰/۰۰۸ [*]	۰/۷۴۳ ± ۰/۰۲۰	۰/۷۷۰ ± ۰/۰۱۳	۸
۰/۰۰۱ [*]	۰/۸۳۴ ± ۰/۰۲۲	۰/۸۶۷ ± ۰/۰۱۶	۰/۰۰۴ [*]	۰/۷۴۴ ± ۰/۰۲۲	۰/۷۷۵ ± ۰/۰۱۴	۹
۰/۰۰۱ [*]	۰/۸۳۴ ± ۰/۰۲۳	۰/۸۶۷ ± ۰/۰۱۵	۰/۰۰۵ [*]	۰/۷۴۴ ± ۰/۰۲۳	۰/۷۷۴ ± ۰/۰۱۸	۱۰

* نتایج تفاوت معنی‌دار در سطح $P = ۰/۰۵$

آنالیزهای خطی مقایسه شد. نتایج نشان دهنده بهبود غیرمعنی‌دار در کاهش واریانس تخمین‌ها بود و لذا نتایج این آنالیزها از قسمت نتایج حذف شدند.

با در نظر گرفتن این نکته که تقریباً تمامی مدل‌های پیش‌بینی کننده موجود از روش‌های خطی و لجستیک برای آنالیز استفاده می‌کنند، استفاده از روابط غیرخطی کشف شده در سیستم‌های ANN می‌تواند در طراحی برنامه‌های موثرتر برای غربالگری افراد مشکوک به استئوپروز کاربرد داشته باشد. استفاده از این مدل‌ها به خصوص در برنامه‌های غربالگری جمعیتی که بهبود تخمین زدن به مقدار کم می‌تواند باعث تغییرات شدید در کاهش اشتباهات شود برتری خود را نشان می‌دهند. این بهبود باعث هدفدارتر شدن برنامه‌های غربالگری برای یافتن افراد با ریسک بالای استئوپروز و در نتیجه بهبود پوشش و هزینه و اثر بخشی این برنامه‌ها می‌شود.

مانند سایر مطالعات این مطالعه نیز محدودیتهایی دارد. اولاً، متغیرهای ورودی ما محدود به متغیرهای تکمیل شده در پرسشنامه بودند. استفاده از سایر متغیرهای آنتروپومتریک و استئوپروز گردن فمور با کمک مدل ANN ۱۰ متغیره اهمیت بالینی بسزایی دارد.

بیوشیمیایی می‌تواند قدرت همه مدل‌ها را افزایش دهد. دوماً، با وجود استفاده از تعداد زیاد شرکت کنندگان در این مطالعه، حد مشخصی برای آموزش موثر شبکه‌های عصبی وجود ندارد و استفاده از تعداد بیشتر افراد می‌تواند کیفیت مدل‌ها را بیشتر کند. در آخر، بعضی از متغیرها احتمالاً ارتباط خطی با متغیرهای خروجی ندارد. بطور مثال، نشان داده شده است که رابطه بین سن و BMD از یک مدل درجه سوم پیروی می‌کند.^(۲۴) برای بررسی این اثرات، ما آنالیزهای رگرسیون اضافی با هدف تکمیل منحنی‌های درجه سوم بر روی اطلاعات انجام دادیم. در این آنالیزها از متغیرهای سن، وزن و قد و استفاده شد و نتایج آنها با

References

1. Akesson K. New approaches to pharmacological treatment of osteoporosis. Bulletin of World Health Organization 2003; 81:657-64.
2. Baran DT, Faulkner KG, Genant HK, et al. Diagnosis and management of osteoporosis: guidelines for the utilization of bone densitometry. Calcified Tissue International 1997; 61:433-40.
3. Siris ES, Miller PD, Barrett-Connor E, et al. Identification and fracture outcomes of undiagnosed low bone mineral density in postmenopausal women: results from the National Osteoporosis Risk Assessment. JAMA 2001; 286:2815-22.
4. Assessment of fracture risk and its application to screening for postmenopausal osteoporosis. Report of a WHO Study Group. World Health Organization Technical Report Series 1994; 843:1-129.
5. Fulton JP. New guidelines for the prevention and treatment of osteoporosis. National Osteoporosis Foundation. Medicine and Health, Rhode Island 1999; 82:110-1.
6. Michaelsson K, Bergstrom R, Mallmin H, et al. Screening for osteopenia and osteoporosis: selection by body composition. Osteoporosis International 1996; 6:120-6.

7. Cadarette SM, Jaglal SB, Kreiger N, et al. Development and validation of the Osteoporosis Risk Assessment Instrument to facilitate selection of women for bone densitometry. *CMAJ* 2000; 162: 1289-94.
8. Lydick E, Cook K, Turpin J, et al. Development and validation of a simple questionnaire to facilitate identification of women likely to have low bone density. *American Journal of Managed Care* 1998; 4:37-48.
9. Clinical practice guidelines for the diagnosis and management of osteoporosis. Scientific Advisory Board, Osteoporosis Society of Canada. *CMAJ* 1996; 155: 1113-33.
10. Houston TP, Elster AB, Davis RM, Deitchman SD. The U.S. Preventive Services Task Force Guide to Clinical Preventive Services, Second Edition. AMA Council on Scientific Affairs. *American Journal of Preventive Medicine* 1998; 14: 374-6.
11. Cross SS, Harrison RF, Kennedy RL. Introduction to neural networks. *Lancet* 1995; 346:1075-9
12. Almeida JS. Predictive non-linear modeling of complex data by artificial neural networks. *Current Opinion in Biotechnology* 2002; 13: 72-6.
13. Sargent DJ. Comparison of artificial neural networks with other statistical approaches: results from medical data sets. *Cancer* 2001; 91:1636-42.
14. Lappe JM. Pathology of osteoporosis and fracture. *Nursing Clinics of North America* 2001; 36:393-400.
15. Suzuki T. Risk factors for osteoporosis in Asia. *Journal of Bone and Mineral Metabolism* 2001; 19:133-41.
16. Cleveland WS. Robust locally weighted regression and smoothing scatter-plots. *Journal of the American Statistical Association* 1979; 74:829-36.
17. Lisboa PJ. A review of evidence of health benefit from artificial neural networks in medical intervention. *Neural Networks* 2002; 15:11-39.
18. Tu JV. Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes. *Journal of Clinical Epidemiology* 1996; 49:1225-31.
19. Dayhoff JE, De Leo JM. Artificial neural networks: opening the black box. *Cancer* 2001; 91:1615-35.
20. Baxt WG. Application of artificial neural networks to clinical medicine. *Lancet* 1995; 346:1135-8.
21. Tetko IV, Villa AEP, Livingstone DJ. Neural network studies: 2. Variable selection. *Journal of Chemical Information and Computer Sciences* 1996; 36:794- 803.
22. Rae SA, Wang WJ, Partridge D. Artificial neural networks: a potential role in osteoporosis. *Journal of Royal Society of Medicine* 1999; 92:119-22.

23. Mohamed EI, Maiolo C, Linder R, et al. Artificial neural network analysis: a novel application for predicting site-specific bone mineral density. *Acta Diabetologica* 2003; 40:S19-22.
24. Wu XP, Liao EY, Zhang H, et al. Establishment of BMD reference plots and determination of peak BMD at multiple skeletal regions in mainland Chinese women and the diagnosis of osteoporosis. *Osteoporosis International* 2004; 15:71-9.

Archive of SID

Artificial neural networks in prediction of osteoporosis among Iranian post-menopausal women

Soltani A. MD*, Sadatsafavi M. MD*, Moayeri A. MD*, Nooraei M. MD*
Hossin-nezhad A. MD*, Larijani B. MD*

It is shown that artificial neural networks (ANN) are promising tools in learning complex interplay of factors on a particular outcome. We performed this study to compare the predictive power of ANN and conventional methods in prediction of bone mineral density (BMD) in Iranian post-menopausal women.

A database of 10 variables (age, weight, age at menopause, corticosteroid use, estrogen use, number of pregnancies, age at menarche, tea consumption, activity and smoking) on 2158 participants (who underwent screening for osteoporosis in an endocrinology research center) was randomly divided into training (1400), validation (150), and test (608) groups. Robust multivariate linear regression and ANN models were developed and validated on the training and validation sets and outcomes (femoral neck and lumbar T scores) were predicted and compared on the test group using different numbers of input variables. Results were evaluated by comparing mean square of differences between predicted and reference values (non-centric chi-square test) and by measuring area under the receiver operating characteristic curve (AUROC) around cutoff value of -2.5.

For models with less than 3 input variables in femoral neck and 4 variables in lumbar area, performance of linear and ANN models was almost the same. With more variables imported into models, ANN outperformed linear regression models. AUROC varied in 2 to 10 variable models as follows: for ANN in lumbar area: from 0.709 to 0.774, linear models in lumbar area: from 0.709 to 0.744, ANN in femoral neck: from 0.801 to 0.867, linear models in femoral neck: from 0.799 to 0.834. All models performed better on prediction of femoral neck BMD than lumbar BMD. The best performance of ANN model yielded a sensitivity of 85.6% and 77.7% and specificity of 74.6% and 65.5% for femoral neck and lumbar area, respectively; whereas the regression model at the best had the sensitivity of 72.5% and 75.2% and specificity of 75.1% and 66.7% for these two places, in that order.

Superior performance of neural networks than linear models can be utilized to achieve more accurate predictions based on easily obtainable and questionnaire based variables. These models demonstrate their advantage especially in mass screening applications, when even a slight enhancement in performance results in significant decrease in number of misclassifications.

KEY WORDS: *Artificial neural networks (ANNs), Bone mineral density, Post-menopausal women*

* Endocrinology and Metabolism Research Center, Tehran University of Medical Sciences and health services, Tehran, Iran.