

پیش‌بینی عوامل مؤثر بر میزان ابتلا به بیماری کووید ۱۹ در بیماران بستری با روش مدل‌سازی آماری

زهرا رحیمی^۱ ID، خدیجه حیدری زاده^۲ ID، فردین حیدری^۳ ID، کیمیا کریمی^{۲*} ID

۱-دانشجوی دکتری، مرکز مطالعات و توسعه، دانشگاه علوم پزشکی لرستان، خرم آباد، ایران

۲-استادیار، مرکز تحقیقات عوامل اجتماعی مؤثر بر سلامت، دانشگاه علوم پزشکی لرستان، خرم آباد، ایران

۳-کارشناس پرستاری، کمیته تحقیقات دانشجویی، دانشگاه علوم پزشکی لرستان، خرم آباد، ایران

یافته / دوره ۲۳ / شماره ۵ / زمستان ۱۴۰۰ / مسلسل ۹۰

چکیده

دریافت مقاله: ۱۴۰۰/۲/۶ پذیرش مقاله: ۱۴۰۰/۱۱/۱۰

مقدمه: همه‌گیری کووید ۱۹ اثر چشمگیری بر بهزیستی افراد و ملل مختلف در مسائل اجتماعی اقتصادی روان شناختی و بهداشت عمومی داشته است. با توجه به اهمیت پیش‌بینی پیک‌های شیوع و بروز بیماری کووید ۱۹ مطالعه با هدف پیش‌بینی عوامل مؤثر بر میزان ابتلا در بیماران بستری با روش رگرسیون لجستیک چند متغیره در شهر خرم آباد انجام شده است.

مواد و روش‌ها: مطالعه از نوع توصیفی-تحلیلی است. داده‌های مبتلایان کووید ۱۹ در پیک اول شروع بیماری که شامل ۴۴۲۵ نفر مراجعین به بیمارستان شهدای عشایر خرم‌آباد در سال ۱۳۹۹ بودند و از این تعداد ۲۹۷۸ نفر تست کووید انجام داده بودند، مورد بررسی قرار گرفتند. چک لیست متغیرها به صورت یک ابزار محقق ساخته به بررسی متغیرهای مؤثر با مراجعه روزانه و مطالعه پرونده بیماران، اطلاعات موجود به دست آمد. تجزیه و تحلیل اطلاعات با استفاده از روش‌های توصیفی، رگرسیون لجستیک چند متغیره و روش ماشین بردار پشتیبان در نرم افزار R برای مدل بندی عوامل خطر مرتبط با تشخیص بیماری انجام شد.

یافته‌ها: بیشترین میزان ابتلا متعلق به گروه سنی ۴۰-۴۹ سال (۱۹٪) و کمترین میزان ابتلا مربوط به گروه سنی ۱۰-۱۹ سال (۱٪) بود. گروه سنی ۶۰ تا ۶۹ سال با ۶ درصد جمعیت و ۲۵ درصد مرگ و میر بود. مدل ایجاد شده نشان می‌دهد از جمله مؤثرترین علائم، سن، تب، کاهش سطح هوشیاری، سطح اکسیژن خون و سابقه ابتلا به بیماری‌های قلبی می‌باشد.

بحث و نتیجه‌گیری: با اطلاعات به دست آمده می‌توان با تشخیص متغیرهای مؤثر، افراد پرخطر را شناسایی کرد و در پیشگیری از بروز بیماری و شیوع بیشتر در گروه‌های پرخطر استفاده کرد که در کنترل پاندمی کووید ۱۹ کمک شایانی می‌تواند داشته باشد.

واژه‌های کلیدی: پیش‌بینی، عوامل مؤثر کووید ۱۹، مدل‌سازی آماری.

*آدرس مکاتبه: خرم آباد، دانشگاه علوم پزشکی لرستان، مرکز مطالعات و توسعه آموزش پزشکی، واحد رهبری آموزشی.

پست الکترونیک: karami.k@lums.ac.ir

مقدمه

کرونا و ویروس‌ها، خانواده بزرگی از ویروس‌ها هستند که می‌توانند عفونت‌های تنفسی و بیماری‌های شدید دیگری مانند مرس را به دنبال داشته باشند (۱).

ویروس جدید از این خانواده به نام کووید ۱۹ در اواخر سال ۲۰۱۹ در چین کشف شد و در ۱۱ مارس سال ۲۰۲۰ توسط سازمان بهداشت جهانی به عنوان یک اپیدمی اعلام شد (۲).

مهار این بیماری در حال حاضر بزرگترین و مهم‌ترین مشکل برای بهداشت عمومی جهان است (۶). شیوع بیماری از فردی به فردی دیگر و از شهری به شهری و از کشوری به کشور دیگر در اثر جابجایی‌ها ادامه دار شد، به صورتی که در تاریخ ۱۷ سپتامبر ۲۰۲۰ بیش از ۱۳ میلیون مبتلا در جهان شناسایی شدند (۲،۳). تعداد مبتلایان به این بیماری در تاریخ ۲۷ ژوئیه ۲۰۲۱ به حدود ۱۹۵ میلیون نفر در جهان رسیده است. ایران نیز از آلودگی به این ویروس در امان نبوده و اپیدمی این ویروس نوظهور از سایر ویروس‌های خانواده کرونا ویروس گسترده‌تر است، که این موضوع بیانگر قدرت فوق‌العاده بالای این ویروس می‌باشد (۴،۵). تعداد مبتلایان به این بیماری تا تاریخ ۲۷ ژوئیه ۲۰۲۱ به حدود ۳ میلیون نفر در ایران رسیده است.

میزان شیوع و شدت اثر کووید ۱۹، به عملکرد سیستم ایمنی فرد در برابر ویروس و البته شرایط خود ویروس بستگی دارد. نوع ویروس، جهش و تعداد آن و عواملی مانند سن، جنسیت، نحوه تغذیه، وضعیت جسمانی فرد و ارتباط بین سیستم عصبی و غدد درون ریز با سیستم ایمنی، که بر عملکرد سیستم ایمنی تأثیر گذارند، در تعیین شدت بیماری نقش دارند (۶).

اگرچه بیماری‌های عفونی می‌توانند همه افراد را تحت تأثیر قرار دهند، اما همانطور که گفته شد، جنسیت می‌تواند به طور قابل توجهی بر پاسخ‌های ایمنی و روند بیماری در بدن انسان تأثیر بگذارد (۷). این بیماری بر روی

مردان نسبت به زنان تأثیر بیشتری می‌گذارد و تعداد بیشتری از مردان را درگیر می‌کند (۸). فرضیه‌های متفاوتی برای توضیح تفاوت‌های جنسی در میزان شیوع و مرگ و میر ناشی از بیماری کووید ۱۹، مانند تفاوت‌های ذاتی در سیستم ایمنی بدن فرد، محیط زیست و مصرف سیگار پیشنهاد شده است (۹). اما هیچ کدام از این فرضیه‌ها تاکنون ثابت نشده‌اند. اگر چه تفاوت در سیستم ایمنی ممکن است سبب تفاوت در شیوع مرگ و میر بین مردان و زنان شود، اما هیچ مطالعه‌ای برای توصیف این پیشنهاد که چگونه سیستم‌های ایمنی مرد و زن ممکن است نسبت به سارس و کووید، واکنش متفاوتی داشته باشند، وجود ندارد و از طرفی فرضیه تأثیر استعمال سیگار بر وضعیت بیمار نیز مورد اختلاف است (۱۰).

در حال حاضر هیچ مطالعه بررسی شده‌ای وجود ندارد که به طور مستقیم خطر بستری شدن در بیمارستان به علت کووید ۱۹، را در میان افراد سیگاری تخمین بزند. با این حال، ۲۷ مطالعه مشاهده‌ای نشان داد که افراد سیگاری ۱۸/۵ درصد یا حدود یک چهارم از بزرگسالان بستری را تشکیل می‌دهند (۱۱).

یکی دیگر از عوامل تأثیرگذار بر کووید ۱۹، رده سنی است. افراد مسن و همچنین افرادی که مشکلات جدی سلامت و یا بیماری زمینه‌ای دارند به دنبال ابتلا به کووید ۱۹، در معرض خطرات شدیدتری هستند (۱۲،۱۳). همچنین مطالعه بر روی افراد دارای بیماری‌های زمینه‌ای بیانگر آن است که در این افراد نه تنها احتمال خطر ابتلا به بیماری بیشتر است بلکه احتمال مرگ و میر ناشی از بیماری نیز بیشتر است (۱۴). افرادی که چندین بیماری مزمن دارند، در صورت آلوده شدن به کووید ۱۹، علائم خطرناک تری را تجربه می‌کنند. این بیماری‌ها شامل دیابت نوع ۲، چاقی شدید و بیماری‌های مزمن قلبی است (۱۵،۱۶).

چک‌لیست متغیرها به صورت یک ابزار محقق ساخته با توجه به موارد فوق تهیه و به روایی محتوایی ۱۰ نفر از متخصصین این زمینه رسید. پس از هماهنگی با محیط پژوهش با مراجعه روزانه و مطالعه پرونده بیماران، اطلاعات موجود استخراج گردید. جهت رعایت حریم خصوصی بیماران از افشا اطلاعات شخصی بیماران خودداری نموده و تنها به تجزیه و تحلیل اطلاعات بالینی آن‌ها پرداخته شد. برای تجزیه و تحلیل اطلاعات از روش‌های آماری از جمله آمار توصیفی جهت توصیف اطلاعات جمعیت شناختی و برآورد میزان شیوع کلی بیماری در گروه‌های سنی متفاوت و از روش رگرسیون لجستیک چندمتغیره و روش ماشین‌بردار پشتیبان در نرم افزار R، برای مدل‌سازی عوامل خطر مرتبط با تشخیص بیماری استفاده شده است و نتایج مقایسه شدند. به منظور بررسی ارتباط متغیرها با نتیجه تست از آزمون کای اسکور استفاده شده است.

یافته‌ها

تحلیل حاضر در سه بخش آمار توصیفی، مدل‌سازی با روش رگرسیون لجستیک برای شناخت علایم مؤثر در ابتلا به بیماری و روش رگرسیون SVM ماشین‌بردار پشتیبان به مدل‌سازی داده‌ها پرداخته است. افرادی که نتیجه تست آن‌ها مثبت بوده به تعداد ۱۲۲۲ نفر بوده است. ۱۶۷۶ نفر بدون در نظر گرفتن پاسخ تست برای درمان در بخش ایزوله و ویژه و ۵۳۲ نفر در بخش ویژه بستری شدند. بیشترین مدت زمان بستری ۵۸ روز و کمترین آن ۱ روز بود. در طی این دوره زمانی دو ماه و نیمه، ۱۵۸۶۰ بیمار در این مرکز به علت بیماری کووید ۱۹ بستری بودند متوسط زمان اقامت افراد مبتلا در بیمارستان ۶ روز بوده است. تعداد موارد مرگ و میر در این بازه ۲۸۰ نفر بوده است.

نمودار توزیع سنی افراد مبتلا به بیماری (نمودار ۱) حاکی از آن است که بیش‌ترین میزان ابتلا متعلق به گروه سنی ۴۰-۴۹ سال با ۱۹ درصد ابتلا و کم‌ترین میزان ابتلا

بنابراین با توجه به تاثیر متغیرهای مختلف مانند سن، جنس و غیره بر میزان شیوع و بروز بیماری‌های مربوط به عفونت‌های تنفسی و اهمیت بررسی این موارد برای پیش‌بینی علائم شایع در تشخیص بیماران مبتلا شده به کووید ۱۹، این مطالعه با هدف پیش‌بینی عوامل مؤثر بر میزان ابتلا به بیماری کووید ۱۹ در بیماران بستری در بیمارستان شهدای عشایر خرم‌آباد انجام گرفته است.

مواد و روش‌ها

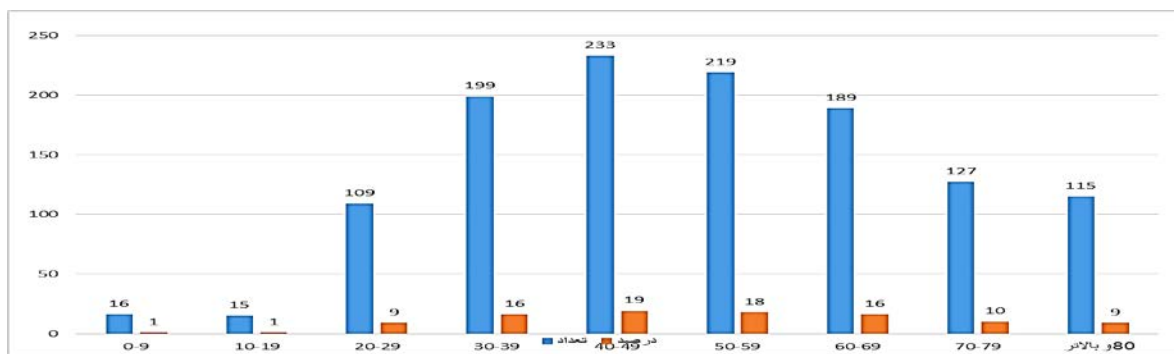
در این مطالعه که به روش آمار توصیفی انجام شده است، تعداد مراجعین مشکوک به ابتلا بیماری کووید ۱۹ از ابتدای اسفند ۱۳۹۸ تا انتهای فروردین سال ۱۳۹۹، در حدود ۴۴۲۵ نفر بوده که به صورت خودمراجعتی یا مراجعه از طریق اورژانس به بیمارستان شهدای عشایر، بیمارستان عمومی تشخیص و درمان بیماری در پیک اول، صورت گرفته از این تعداد، تعداد ۲۹۷۸ نفر در زمان انجام پژوهش مورد تست کووید ۱۹ (تست خون و PCR) قرار گرفته‌اند و نتیجه تست آن‌ها مشخص بوده است.

متغیرهای مدل شامل سن، جنس، سابقه تماس با بیمار مبتلا به کووید ۱۹، جواب مثبت تست کووید ۱۹، علائم اولیه: تب، سرفه، درد عضلانی، دیسترس تنفسی، کاهش سطح هوشیاری، از بین رفتن یا کاهش حس بویایی، از بین رفتن یا کاهش حس چشایی، تشنج، سردرد، سرگیجه، پارزی اندام‌ها، پلژی اندام‌ها، درد قفسه سینه، التهاب یا ضایعه پوستی، علایم گوارشی نظیر: دل درد، تهوع، استفراغ، اسهال، بی‌اشتهایی و سابقه مصرف سیگار، سابقه سوء مصرف مواد مخدر، میزان Po2، کانسرو، بیماری‌های مزمن کبدی، دیابت، بیماری‌های مزمن خونی، نقص ایمنی (اکتسابی یا مادرزادی)، بارداری، بیماری‌های قلبی، بیماری‌های مزمن کلیوی، آسم، سایر بیماری‌های مزمن ریوی بجز آسم، اختلالات مزمن عصبی، سایر بیماری‌های مزمن، سابقه ابتلا به بیماری فشار خون می‌باشند.

مربوط به گروه سنی ۱۰-۱۹ سال با یک درصد ابتلا است. همچنین در مقایسه آمار و اطلاعات با جمعیت شهری و روستایی شهرستان خرم آباد که در زمان نمونه‌گیری ۴۸۰۷۳۹ نفر بوده است و استخراج اطلاعات میزان جمعیت در گروه‌های سنی متفاوت و مقایسه میزان ابتلا در گروه‌های سنی با توجه به جمعیت هر گروه سنی به این نتیجه می‌رسیم که گروه سنی ۴۰-۴۹ سال ۱۵ درصد جمعیت و ۱۹ درصد مبتلایان به بیماری کووید ۱۹، را به

خود اختصاص می‌دهند. از طرفی میانگین سنی افراد فوتی ۶۳ سال، میانه سنی ۶۸ سال، ماکزیمم سنی ۹۷ سال و کم‌ترین سن مربوط به چند نوزاد (چند ماهه)، است. بیشترین آمار فوتی‌ها در دو دسته ۶۹-۶۰ سال و ۷۹-۷۰ سال می‌باشد که با توجه به آمار تعداد افراد هر رده سنی بیش‌ترین میزان مرگ و میر مربوط به گروه سنی ۶۹-۶۰ سال با ۶ درصد جمعیت و ۲۵ درصد مرگ و میر است (جدول ۱).

نمودار ۱. نمودار توزیع سنی و درصد ابتلا به بیماری در گروه‌های سنی مختلف



جدول ۱. جدول توزیع سن جمعیت خرم‌آباد، افراد مبتلا به کووید ۱۹ و مرگ و میر

سن	توزیع سنی جمعیت شهرستان خرم‌آباد	توزیع سنی جمعیت افراد مبتلا به کووید ۱۹	میزان مرگ و میر در هر رده سنی
تعداد	درصد	تعداد	درصد
۹-۰	۸۳۳۹۳	۱۶	۰/۰۵
۱۹-۱۰	۶۶۵۵۵	۱۵	۰/۰۱
۲۹-۲۰	۷۱۳۲۶	۱۰۹	۰/۰۳
۳۹-۳۰	۱۰۰۰۲۱	۱۹۹	۰/۰۳
۴۹-۴۰	۷۱۶۴۷	۲۳۳	۰/۰۶
۵۹-۵۰	۴۲۴۲۵	۲۱۹	۰/۱۲
۶۹-۶۰	۲۷۶۵۲	۱۸۹	۰/۲۵
۷۹-۷۰	۱۰۵۲۴	۱۲۷	۰/۱۷
۸۰ و بیشتر	۷۱۹۶	۱۱۵	۰/۰۳

برای بررسی عوامل خطرزای مرتبط با بیماری کووید ۱۹ از مدل رگرسیون لجستیک استفاده شد. برای تجزیه و تحلیل اطلاعات از اندازه اثر ضرایب رگرسیونی، خطای استاندارد، مقادیر z-value، استفاده شده است. داده‌های جمع‌آوری شده با استفاده از نرم افزار R تجزیه و تحلیل شده است. همچنین به کمک بررسی مقدار p-value هر متغیر میزان اهمیت آن‌ها در پیش‌بینی مورد سنجش قرار داده شد. متغیرهایی با مقدار p-value کوچک‌تر از ۰/۰۵ در فرآیند مدل‌سازی مفید تلقی می‌شوند که در این جا با علامت * نشان داده شده‌اند (جدول ۲).

جدول ۲. اهمیت متغیرها پس از مدل سازی رگرسیون

متغیر	ضریب در مدل	خطای استاندارد	z-value	p-value	درجه اهمیت
اینترسپت	-۰/۹۳۸	۰/۱۶	-۵/۸۳	۰/۰۰۰۰۰۰۰۵۵۱	***
سن	۰/۰۰۸۹	۰/۰۰۲	۴/۳۲	۰/۰۰۰۰۱۵	***
تب	۰/۲۷۶	۰/۰۷	۳/۴۸	۰/۰۰۰۴۸	***
سرفه	۰/۲۳۵	۰/۰۸	۲/۹۲	۰/۰۰۳۵	**
دیسترس	۰/۲۱۲۷	۰/۰۸	۲/۵۸	۰/۰۰۹۸	**
کاهش سطح هوشیاری	۰/۷۲۹	۰/۲۱	-۳/۴۵	۰/۰۰۰۵۵	***
سابقه مصرف سیگار	-۱/۱۲۲	۰/۳۹	-۲/۸۶	۰/۰۰۴۱	**
سطح اکسیژن خون بیش از ۹۳ درصد	-۰/۲۹۵	۰/۰۸	-۳/۶۰	۰/۰۰۰۳۱۱	***
بیماری قلبی	-۰/۴۸	۰/۱۳	-۳/۴۶	۰/۰۰۰۵۳۲	***
اعصاب	-۰/۹۹۹	۰/۵۶	-۱/۷۶	۰/۰۷۷	.
سایر عوامل	-۰/۳۱۸	۰/۱۸	-۱/۷۵	۰/۰۷۹	.
فشارخون	-۰/۳۴۶	۰/۱۸	-۱/۸۳	۰/۰۶۶	.
درجه اهمیت:	۰/۱(۱)	۰/۰۵(۰)	۰/۰۱(*)	۰/۰۰۱(**)	۰/۰۰۰۰۰۰۰۰۰۰۰۰۰۰۰۰(***)

روی داده‌های گروه آموزش انجام می‌شود. یک فرض بسیار قوی در ماشین بردار پشتیبان این است که داده‌ها به صورت جداپذیر خطی هستند. در این صورت خط جداکننده بین کلاس‌ها با استفاده از الگوریتم‌های مختلفی مانند پرسپترون رسم می‌شود. در این روش مرز خطی بین دو کلاس به گونه‌ای محاسبه می‌شود که: ۱- تمام نمونه‌های کلاس +۱ در یک طرف مرز و تمام نمونه‌های کلاس -۱ در طرف دیگر مرز واقع شوند و ۲- مرز تصمیم‌گیری به گونه‌ای می‌باشد که فاصله نزدیک‌ترین نمونه‌های آموزشی هر دو کلاس از یکدیگر در راستای عمود بر مرز تصمیم‌گیری تا جایی که ممکن است حداکثر شود. در حالت کلی دو نوع ماشین بردار پشتیبان وجود دارد که استفاده از آن‌ها به نوع پیامد بستگی دارد. اگر پیامد مد نظر پیوسته باشند، از نوع رگرسیون و اگر گسسته باشند از نوع رده بندی استفاده می‌شود.

با توجه به این که بیماران مشکوک به ابتلا به بیماری کووید ۱۹ در یکی از دو کلاس مبتلا و عدم ابتلا قرار دارند، یعنی متغیر پاسخ، از نوع کیفی دو حالت است بنابراین نوع مدل به کار گرفته شده برای ماشین بردار پشتیبان از نوع طبقه بندی است. به منظور اعتبار سنجی نتایج حاصل از روش ماشین بردار پشتیبان، روش اعتبار سنجی آموزش و آزمون به کار برده شده است. به این صورت که داده‌ها به طور تصادفی

ماشین بردار پشتیبان (Support vector machine;

SVM)، به عنوان یک روش آماری قوی و کارآمد در تشخیص و پیش‌بینی پیامدهای بالینی بر اساس ترکیباتی از متغیرهای پیش‌بین کاربرد دارد. هدف این بخش از پژوهش، استفاده از SVM برای تشخیص عوامل مؤثر بر ابتلا به بیماری کووید ۱۹ است. یکی از روش‌های پیشرفته و قوی آماری در زمینه تشخیص و طبقه بندی صحیح پیامدها بر اساس عوامل خطر مرتبط که اخیراً کاربرد فراوانی در علوم پزشکی پیدا کرده است، روش ماشین بردار پشتیبان (SVM) است. با توجه به نتایج قابل قبول این روش در کاربردهای پزشکی در این مطالعه بر آن شدیم تا از این روش برای طبقه بندی وضعیت استفاده کنیم. لذا هدف از این پژوهش استفاده از روش ماشین بردار پشتیبان برای تشخیص عوامل خطر مرتبط با ابتلا به بیماری بر اساس متغیرهای ذکر شده بالا می‌باشد.

فرآیند ساخت مدل شامل دو مرحله آموزش و آزمون (TEST&TRAIN)، است که در مرحله آموزش، مدل الگو و روابط موجود بین پیامد و پیش‌بین‌ها را می‌آموزد و سپس قابلیت تعمیم‌دهی مدل بر ارزش داده شده توسط داده‌ها مورد ارزیابی قرار می‌گیرد. مزیت این روش نسبت به روش‌های کلاسیک آماری این است که در این روش نیاز به پیش‌فرض خاصی در مورد داده‌ها نیست، یعنی شناسایی الگوی ارتباط از

افزار R و بسته نرم افزاری AUC و پس از تعیین شاخص‌های مرکزی و پراکندگی مورد تجزیه و تحلیل قرار گرفت. پس از بررسی نرمال بودن داده‌ها با استفاده از آماره‌های توصیفی، برای مقایسه متغیرهای کمی نرمال و غیر نرمال در دو گروه به ترتیب از آزمون‌های تی مستقل و آزمون من ویتنی (Mann-Whitney)، استفاده شد. سطح معنی‌داری در آزمون‌ها ۰/۰۵ در نظر گرفته شده است.

به کمک P مرز طبقه بندی مشخص شد، یعنی اگر P برابر با ۰/۰۵ باشد احتمال‌های بالاتر از ۰/۰۵ که توسط مدل پیش‌بینی شده‌اند، با کلاس تست مثبت و زیر ۰/۰۵ با کلاس منفی برچسب گذاری شدند. دقت‌ها توسط ماتریس درهم‌ریختگی محاسبه شده‌اند، مشاهده شد که بهترین حالت با پارامترهای $\text{cost}=1$ و $\text{gamma}=10$ می‌باشد. حال با این مدل و چند مدل نمونه دیگر با پارامترهای متفاوت پیش‌بینی و نتایج بررسی شد. جهت بررسی دقت مدل در پیش‌بینی داده‌ها از اعتبارسنجی متقاطع ۱۰ بخشی استفاده شد. که در این جا دقت مدل در حدود ۶۰ درصد به دست آمد. متغیرهای مورد استفاده در این مدل شامل سن، سرفه، دی‌سترس تنفسی، هوشیاری، سابقه مصرف سیگار، سطح اکسیژن خون، بیماری قلبی، فشارخون می‌باشد. در این بخش انتخاب بهترین متغیرها جهت پیش‌بینی در پنج مرحله و به کمک معیار AIC یا criterion Akaike information صورت می‌گیرد. بهترین مدل دارای کم‌ترین AIC می‌باشد (جدول ۳).

به دو قسمت ۸۰ درصد (داده آموزش) و ۲۰ درصد (داده آزمون) تقسیم شدند. پس از برازش مدل بر روی داده‌های آموزش، برای اعتبارسنجی نتیجه به دست آمده و بررسی قابلیت تعمیم دهی مدل، بر روی داده‌های آزمون از سطح زیر منحنی (Receiver operating characteristic) (ROC) استفاده شده است. با توجه به این که در این روش داده‌های گروه آزمون کاملاً از روند مدل‌سازی کنار گذاشته می‌شوند و در صورت داشتن حجم نمونه پایین، این کار باعث کاهش توان مطالعه می‌شود. یک روش دیگر برای سنجش اعتبار مدل، استفاده از روش اعتبارسنجی k باره است، که در این روش کل داده‌ها به k دسته تقریباً هم اندازه تقسیم می‌شوند. $k-1$ دسته برای آموزش مدل به کار می‌رود و مدل بر روی k امین دسته آزمون می‌شود. این کار k بار تکرار می‌شود، به طوری که هر بار یکی از دسته‌ها به عنوان آزمون و بقیه به عنوان آموزش استفاده شود. میانگین شاخص‌های اعتبارسنجی حاصل از k دسته آزمون به عنوان شاخص نهایی گزارش می‌شود.

در مطالعه حاضر، برای اعتبارسنجی مدل و ارزیابی قابلیت تعمیم‌دهی نتایج به داده‌های جدید، با استفاده از روش اعتبارسنجی ۱۰ باره، داده‌ها به ۱۰ قسمت تقریباً مساوی تقسیم شدند و پس از برازش مدل بر روی داده‌های آموزش، مدل بر روی داده‌های آزمون، مورد ارزیابی قرار گرفت و در نهایت، متوسط شاخص‌های مربوط به ارزیابی دقت طبقه بندی کننده حاصل از ۱۰ گروه آزمون گزارش شد. تحلیل‌ها با نرم

جدول ۳. پیش‌بینی آزمون‌ها به وسیله SVM

پیش‌بینی به وسیله SVM			
کرنل	کاست	گاما	دقت
رادپال	۰/۱	۰/۵	۰/۵۸۹۹
رادپال	۰/۱	۰/۱	۰/۵۸۹۹
رادپال	۱	۰/۵	۰/۵۸۳۲
رادپال	۱	۱	۰/۵۸۹۹
لینر	۱	۰/۵	۰/۵۸۹۹
لینر	۱	۱	۰/۵۸۹۹

از کل داده‌ها ۸۰ درصد آموزشی و ۲۰ درصد آزمایشی بوده‌اند

دقت‌ها توسط ماتریس درهم‌ریختگی محاسبه شده‌اند

مدل روی تمامی متغیرها اجرا شده است (به غیر از a1.paresis و a1.poost و patient)

بحث و نتیجه گیری

اغلب افراد آلوده به کووید ۱۹ پس از دوره کمون تظاهرات بالینی یک بیماری تنفسی خفیف تا متوسط مانند تب، سرفه خشک، احساس خستگی، درد عضلانی و تنگی نفس را دارند (۱۷). موارد حاد ابتلا به این بیماری مشکلات گوارشی نظیر اسهال، نارسایی حاد تنفسی، اختلالات انعقادی خون و نارسایی کلیوی گزارش شده است (۱۸).

بر اساس این مطالعه تب ($P=0/00048$)، سرفه ($P=0/0035$)، دیسترس تنفسی ($P=0/0098$) و میزان پایین po_2 ($P=0/0041$) علائم اولیه ابتلا به کووید ۱۹ محسوب می‌شوند و این علائم مرتبط با احتمال ابتلا به این بیماری هستند. در مطالعه گوان و همکاران نیز تب ۸۷/۹ درصد و سرفه با ۶۷/۷ درصد شایع‌ترین علائم ابتلا به کووید ۱۹ شناسایی شدند، در حالی که اسهال ۳/۷ درصد و استفراغ با ۵ درصد از عوامل نادر بودند (۱۷). که نتایج این مطالعه با مطالعه حاضر همسو بود.

در مطالعه حاضر به بررسی تأثیر رده سنی بر میزان ابتلا به کووید ۱۹ نیز پرداخته شد. براساس نتایج این مطالعه بیشترین ابتلا مربوط به گروه سنی ۴۰ تا ۴۹ سال با ۱۹ درصد ابتلا و کمترین ابتلا مربوط به گروه سنی ۱۰ تا ۱۹ سال با یک درصد ابتلاست. همچنین بیشترین آمار فوتی در دو دسته ۶۰ تا ۶۹ سال و ۷۰ تا ۷۹ سال می‌باشد که با توجه به آمار تعداد افراد در هر رده سنی بیشترین میزان مرگ و میر مربوط به گروه سنی ۶۹ سال با ۶ درصد جمعیت و ۲۵ درصد مرگ و میر می‌باشد. در واقع سالمندان به دلایل متعددی از جمله ضعف سیستم ایمنی و بیماری‌های مزمن زمینه‌ای، رژیم چند دارویی، کم‌توجهی و کم‌توانی در رعایت کامل و درست اصول بهداشت فردی و مراقبت از خود، بهداشت نامناسب زندگی، تنهایی و تجرد، ساده انگاری بیماری‌ها، عدم تمایل به مصرف دارو، تغذیه نامناسب و فراموشی در زمره آسیب پذیرترین اقشار

جامعه در مواجهه با کووید ۱۹ قرار می‌گیرند (۱۹). در مطالعه Bonanad و همکاران نیز به بررسی تأثیر سن بر میزان شدت و مرگ و میر بیماری پرداخته شد. در مطالعه عنوان شده نیز مانند مطالعه حاضر مرگ و میر در بیماران کمتر از ۵۰ سال کمتر از ۱ درصد بود. از طرفی بالاترین میزان مرگ و میر در بیماران بالای ۸۰ سال مشاهده شد. همچنین بالاترین میزان خطر مرگ و میر در بیماران ۶۰ تا ۶۹ ساله در مقایسه با ۵۰ تا ۵۹ سال مشاهده شد. از طرفی بیماران بالای ۸۰ سال در مقایسه با بیماران ۷۰ تا ۷۹ ساله ۶۰ درصد بیشتر در معرض خطر مرگ بودند (۲۰). نتایج این مطالعه با مطالعه حاضر همسو بود.

در بررسی رابطه بین بیماری‌های زمینه‌ای و میزان ابتلا، شدت و مرگ و میر ناشی از کووید ۱۹ در مطالعه حاضر بیماری‌های زمینه‌ای مانند فشار خون مزمن ($P=0/066$)، اختلالات عصبی مزمن ($P=0/077$) و همچنین کاهش سطح هوشیاری ($P=0/00055$) از جمله علائم اولیه مرتبط با احتمال ابتلا به بیماری کووید ۱۹ هستند. در مطالعه احمدی و همکاران نیز بیماری‌های قلبی عروقی ($p=0/001$) و نارسایی کلیوی ($p=0/001$)، از عوامل مرتبط با خطر مرگ ناشی از کووید ۱۹ شناسایی شدند (۲۱). نتایج مطالعه ذکر شده نیز با مطالعه حاضر همسو بود.

یکی دیگر از عوامل تأثیر گذار بر شدت بیماری کووید ۱۹، مصرف دخانیات است. در مطالعه حاضر نتایج حاکی از این بود که مصرف سیگار ($P=0/0041$)، با شدت ابتلا به بیماری کووید ۱۹ مرتبط است. در واقع ویروس کووید ۱۹ به گیرنده‌ای به نام پذیرنده آنژیوتانسین ۲، متصل شده و توسط آن به سلول وارد می‌شود. سپس از طریق آنژیوتانسین ۲، به آنژیوتانسین ۱-۷، موجب آسیب ریوی می‌شود (۲۲). در مطالعه Cai و همکاران، مشخص شد که مصرف سیگار موجب افزایش معناداری در این فاکتور می‌شود؛ که این افزایش ممکن

مصرف دخانیات وجود دارد. بنابراین برای تعیین دقیق اثرات مصرف دخانیات بر احتمال ابتلا به کووید ۱۹ انجام مطالعات بیشتری ضرورت دارد.

محدودیت‌های مطالعه

انجام مطالعه حاضر محدودیت‌هایی نیز دارد. از جمله می‌توان به منفی بودن کاذب تست PCR در بیماران مراجعه کننده، عدم دسترسی به تاریخچه دقیق بیماران برای بررسی وجود بیماری‌های زمینه‌ای اشاره کرد. علاوه بر این برخی از متغیرهای مرتبط با بیماری کووید ۱۹ مانند کاهش حس بویایی و چشایی و بعضی از بیماری‌های زمینه‌ای مانند دیابت، سردرد، سرگیجه، پارزی اندام‌ها و ... در این مطالعه به صورت جزئی و تک به تک مورد بررسی قرار نگرفته‌اند و به صورت کلی تحت عنوان سایر عوامل به آن‌ها پرداخته شده است.

تشکر و قدردانی

پژوهش حاضر نتیجه طرح تحقیقاتی مصوب دانشکده پیراپزشکی دانشگاه علوم پزشکی لرستان با کد اخلاق IR.LUMS.REC.1399.355 می‌باشد از همه عزیزانی که در انجام این پژوهش ما را یاری رساندند صمیمانه تشکر و قدردانی می‌نماییم.

است موجب افزایش خطر اتصال ویروس به و ورود آن به ریه‌های افراد سیگاری و در نهایت بروز اختلال در ریه‌های آن‌ها شود (۲۳). نتایج پژوهش ذکر شده با مطالعه حاضر هم‌سو بود.

در بررسی سایر عوامل تأثیرگذار بر شدت ابتلا به کووید ۱۹ و مرگ و میر ناشی از آن در مطالعه حاضر مشخص شد که جنسیت رابطه غیر معنادار اما قابل توجهی بر مرگ و میر ناشی از این بیماری دارد. اما در مطالعه Marina و همکاران، مشخص شد که در تمام افراد دارای ۲۰ سال سن یا بیشتر، مرگ و میر مردان نسبت به زنان بیشتر بود (۲۴). که نتایج این مطالعه با مطالعه حاضر در تضاد بود. علت این اختلاف نتایج می‌تواند به این دلیل باشد که در پژوهش کنونی جنسیت به طور جداگانه مورد بررسی قرار نگرفته است. بنابراین برای اثبات دقیق تأثیر جنسیت بر شدت و میزان ابتلا به کووید ۱۹ مطالعات بیشتری باید انجام گیرد.

امروزه بیماری کووید ۱۹ به چالش بزرگ سلامت در سراسر جهان تبدیل شده است و تاکنون مرگ و میر زیادی را به دنبال داشته است. می‌توان گفت از بهترین راه‌های مبارزه با این بیماری، شناخت ابعاد و عوامل خطر مرتبط با آن و همچنین جلوگیری از آلودگی و انتشار از طریق اقدامات محافظتی و بهداشتی شخصی ضروری است. همچنین استفاده از سواد بهداشتی یکی از بهترین راه‌ها و ابزارهای حیاتی مقابله با این بیماری است. از طرفی شناخت عوامل مرتبط با احتمال ابتلا به این بیماری در رفع چالش ناشی از آن بسیار مهم است. مطابق با مطالعه حاضر عواملی مانند سن، جنسیت، ابتلا به بیماری‌های زمینه‌ای، اختلالات مزمن عصبی، تب، سرفه، دیسترس تنفسی، علائم گوارشی و مصرف دخانیات رابطه مستقیم با احتمال ابتلا به کووید ۱۹ دارند اما باید توجه نمود که موارد متناقضی در رابطه با

References

- Alipour A, Ghadami A, Alipour Z, Abdollahzadeh H. Preliminary validation of the Corona Anxiety Scale (CDAS) in an Iranian sample. *J. Health Psychol.* 1398. Year 8 No. 4: 163-175.(In Persian)
- Tantrakarnapa, K., Bhopdhornangkul, B., & Nakhaapakorn, K. (2020). Influencing factors of COVID-19 spreadig: a case study of Thailand. *J Public HealthGermany*).<https://doi.org/10.1007/s10389-020-01329-5>.
- Ing EB, Xu QA, Salimi A, Torun N. Physician deaths from corona virus (COVID-19) disease. *Occup Med (Lond)*. 2020 Jul 17;70(5):370-374. doi: 10.1093/occmed/kqaa088. PMID: 32409839; PMCID: PMC7239175.
- Tong ZD, Tang A, Li KF, Li P, Wang HL, Yi JP, et al. Potential Presymptomatic Transmission of SARS-CoV-2, Zhejiang Province, China, 2020. *Emerg Infect Dis*. 2020 May;26(5):1052-1054. doi:10.3201/eid2605.200198. Epub 2020 May 17. PMID: 32091386; PMCID: PMC7181913.
- Cascella M, Rajnik M, Aleem A, Dulebohn SC, Di Napoli R. Features, Evaluation, and Treatment of Coronavirus (COVID-19). 2021 Sep 2. In: *StatPearls [Internet]*. Treasure Island (FL): StatPearls Publishing; 2021 Jan-. PMID: 32150360.
- Ganji A, Mosayebi G, Khaki M, Ghazavi A. A Review of the 2019 Novel Coronavirus(Covid-19): Immunopathogenesis, Molecular Biology and Clinical Aspects. *J Arak Uni Med Sci*. 2020; 23 (1) :8-21. URL: <http://jams.arakmu.ac.ir/article-1-6296-fa.html>. (In Persian).
- Gausman J, Langer A. Sex and Gender Disparities in the COVID-19 Pandemic. *J Womens Health (Larchmt)*. 2020 Apr;29(4):465-466. doi: 10.1089/jwh.2020.8472. PMID: 32320331.
- Jin JM, Bai P, He W, Wu F, Liu XF, Han DM, et al. Gender Differences in Patients With COVID-19: Focus on Severity and Mortality. *Front Public Health*. 2020 Apr 29;8:152. doi: 10.3389/fpubh.2020.00152. PMID: 32411652; PMCID: PMC7201103.
- Peckham H, de Gruijter NM, Raine C, et al. Male sex identified by global COVID-19 meta-analysis as a risk factor for death and ITU admission. *Nature Communications*. 2020 Dec;11(1):6317. DOI: 10.1038/s41467-020-19741-6. PMID: 33298944; PMCID: PMC7726563.
- Cai, G. Bulk and Single-Cell Transcriptomics Identify Tobacco-Use Disparity in Lung Gene Expression of ACE2, the Receptor of 2019-nCov. Preprints 2020, 2020020051 (doi: 10.20944/preprints202002.0051.v2).
- Cai H. Sex difference and smoking predisposition in patients with COVID-19. *Lancet Respir Med*. 2020 Apr;8(4):e20. doi: 10.1016/S2213-2600(20)30117-X. Epub 2020 Mar 11. Erratum in: *Lancet Respir Med*. 2020 Apr;8(4):e26. PMID: 32171067; PMCID: PMC7103991..
- Williamson EJ, Walker AJ, Bhaskaran K, Bacon S, Bates C, Morton CE, et al.

- Factors associated with COVID-19-related death using OpenSAFELY. *Nature*. 2020 Aug;584(7821):430-436. doi: 10.1038/s41586-020-2521-4. Epub 2020 Jul 8. PMID: 32640463; PMCID: PMC7611074.
13. Ningthoujam R, Khomdram D. WHO statement - "Older people are at highest risk from COVID-19": Should the hypothesis be corroborated or rejected? *Med Hypotheses*. 2020 Jun 16;144:109896. doi: 10.1016/j.mehy.2020.109896. Epub ahead of print. PMID: 32585464; PMCID: PMC7297172
 14. Onder G, Rezza G, Brusaferro S. Case-Fatality Rate and Characteristics of Patients Dying in Relation to COVID-19 in Italy. *JAMA*. 2020 May 12;323(18):1775-1776. doi: 10.1001/jama.2020.4683. Erratum in: *JAMA*. 2020 Apr 28;323(16):1619. PMID: 32203977. onset to the acceleration phase of the US outbreak. *Journal of general internal medicine*. 2020;35(11):3285-3292.
 15. Farnoosh G, Ghanei M, Khorramdelazad H, Alishiri G, Farahani AJ, Shahriary A, et al. Are Iranian sulfur mustard gas-exposed survivors more vulnerable to SARS-CoV-2? Some similarity in their pathogenesis. *Disaster Med Public Health Prep*. 2020;14(6):826-832.(in Persian)
 16. Bailey SC, Serper M, Opsasnick L, Persell SD, O'Connor R, Curtis LM, et al. Changes in COVID-19 Knowledge, Beliefs, Behaviors, and Preparedness Among High-Risk Adults from the Onset to the Acceleration Phase of the US Outbreak. *J Gen Intern Med*. 2020 Nov;35(11):3285-3292. doi: 10.1007/s11606-020-05980-2. Epub 2020 Sep 1. PMID: 32875509; PMCID: PMC7462357..
 17. Guan WJ, Ni ZY, Hu Y, Liang WH, Ou CQ, He JX, et al. China Medical Treatment Expert Group for Covid-19. Clinical Characteristics of Coronavirus Disease 2019 in China. *N Engl J Med*. 2020 Apr 30;382(18):1708-1720. doi: 10.1056/NEJMoa2002032. Epub 2020 Feb 28. PMID: 32109013; PMCID: PMC7092819.
 18. Wang W, Tang J, Wei F. Updated understanding of the outbreak of 2019 novel coronavirus (2019-nCoV) in Wuhan, China. *J Med Virol*. 2020 Apr;92(4):441-447. doi: 10.1002/jmv.25689. Epub 2020 Feb 12. PMID: 31994742; PMCID: PMC7167192.
 19. Li J.-Y., You Z., Wang Q., Zhou Z.-J. Qiu, Y. Luo, et al. (2020) The Epidemic of 2019-Novel-Coronavirus (2019-nCov) Pneumonia and Insights for Emerging Infectious Diseases in the Future. *Microbes and Infections*, 22, 80-85. <https://www.sciencedirect.com/science/article/pii/S1286457920300307> <https://doi.org/10.1016/j.micinf.2020.02.002>
 20. Bonanad C, García-Blas S, Tarazona-Santabalbina F, Sanchis J, Bertomeu-González V, Fácila L, et al. The Effect of Age on Mortality in Patients With COVID-19: A Meta-Analysis With 611,583 Subjects. *J Am Med Dir Assoc*. 2020

- Jul;21(7):915-918. doi:
10.1016/j.jamda.2020.05.045. Epub 2020
May 25. PMID: 32674819; PMCID:
PMC7247470.
21. Ahmadi A., Ardeshiri S ., Rajab Nezhadi
V., Pajooresh A., Narimani Moghadam T,
Sabaghan M. .et al. Risk Factors for
Mortality in Hospitalized Patients with
COVID-19: A Cross-Sectional Study in
Southwestern Iran. Iranian Red Crescent
Medical Journal, 23(9). (in persian).
<https://doi.org/10.32592/ircmj.2021.23.9.12>
30
22. Ghoshooni H, Marefati N, Sahraei H,
Mahabadi M. The Effect of Smoking on
the Occurrence of COVID-19 and
Comparison with Non-Smokers. J Mar
Med. 2021; 3 (1) :13-20. (in Persian).URL:
<http://jmarmed.ir/article-1-111-en.html>
23. Cai G, Bossé Y, Xiao F, Kheradmand F,
Amos CI. Tobacco Smoking Increases the
Lung Gene Expression of ACE2, the
Receptor of SARS-CoV-2. Am J Respir
Crit Care Med. 2020 Jun 15;201(12):1557-
1559. doi:10.1164/rccm.202003-0693LE.
PMID: 32329629; PMCID: PMC7301735.
24. Marina S and Piemonti L.. Gender and Age
Effects on the Rates of Infection and
Deaths in Individuals with Confirmed
SARS-CoV-2 in European Countries.
SSRN Electronic journal.
2020,10.2139/ssrn.3576790.

Prediction of the Effective Factors on the Incidence of COVID-19 Disease in Hospitalized Patients using Statistical Modeling

Rahimi Z.¹, Heidarizadeh Kh.², Heydari F.³, Karami K.^{2*}

1. PhD student, Educational Development Center, Lorestan University of Medical Sciences, Khorramabad, Iran

2. Assistant Professor, Social Determinants of Health Research Center, Lorestan University of Medical Sciences, Khorramabad, Iran

3. Nursing expert, Student Research Committee, Lorestan University of Medical Sciences, Khorramabad, Iran, karami.k@lums.ac.ir

Received: 28 Sep 2021

Accepted: 30 Jan 2022

Abstract

Background: The COVID-19 pandemic has had a significant impact on the well-being of different individuals and nations in terms of socioeconomic, psychological, and public health issues. Regarding the importance of predicting the peak outbreaks and incidence of COVID-19, this study aimed to predict factors that affect the incidence of COVID-19 pandemic in hospitalized patients admitted to a hospital in Khorramabad city of Iran, using a multivariate logistic regression model.

Materials and Methods: In this descriptive-analytical study, the data from 4,425 COVID-19 patients in the first peak of the disease, who were referred to the Shohada-ye Ashayer Hospital, Khorramabad, Iran, in 2020, were examined, from whom 2,978 people had undergone COVID-19 test. Data were collected using a researcher-made checklist for variables and examination of patients' daily records. Data analysis was performed using descriptive methods, multivariate logistic regression model, and backup vector machine method in the R software for the modeling of risk factors associated with the disease diagnosis.

Results: The highest and lowest rates of incidence were observed in the age range of 40-49 (19%) and 19-10 years (1%). The age group 60-69 years accounted for 6% of the population with a mortality rate of 25%. Based on the implemented model, the most effective symptoms associated with the incidence of COVID-19 included age, fever, decreased level of consciousness, blood oxygen level, and a history of heart disease.

Conclusion: The present study showed that identification of the effective variables of the disease led to the identification of high-risk individuals. This method can be used to prevent disease incidence and prevalence in high-risk groups and is a great help in controlling the COVID-19 pandemic.

Keywords: COVID-19, Effective factors, Prediction, Statistical modeling.

***Citation:** Rahimi Z, Heidarizadeh Kh, Heydari F, and Karami K. Predicting the effective factors on the incidence of covid-19 disease in hospitalized patients by statistical modeling. *Yafte*. 2022; 23(5):87-98.