

## بررسی تأثیر فصول بر شدت بیماری کووید-۱۹: تجزیه و تحلیل داده‌های اپیدمیولوژیک یکسان بیمارستانی با استفاده از ضریب همبستگی جاکارد

### چکیده

دریافت: ۱۴۰۲/۰۷/۰۵ ویرایش: ۱۴۰۲/۰۷/۱۱ پذیرش: ۱۴۰۲/۰۹/۲۳ آنلاین: ۱۴۰۲/۱۰/۰۱

آمنه جوانمرد<sup>۱</sup>، علیرضا صالحان<sup>۲\*</sup>

۱- گروه مهندسی کامپیوتر، دانشکده مهندسی، دانشگاه امام رضا (ع)، مشهد، مشهد، ایران.  
۲- گروه مهندسی کامپیوتر، دانشکده مهندسی، دانشگاه تربیت مدرس، تربت حیدریه، ایران.

**زمینه و هدف:** در سال ۱۹۶۰، ویروس‌های کرونا کشف شدند. موجودات زنده درشت‌پیکر از خانواده ویروس‌های پاکت‌دار که RNA تک‌رشته‌ای با منشاء جانوری دارند. ویروس‌های کرونا در انسان می‌تواند به بیماری تنفسی خفیف یا شدید تنفسی تبدیل شوند. در سال ۲۰۲۰، سازمان بهداشت جهانی ویروس کووید-۱۹ را یک بیماری همه‌گیر جهانی اعلام کرد. هدف این مطالعه استفاده از ضریب همبستگی جاکارد جهت تعیین شباهت الگوی رفتار بیماری کووید-۱۹ در فصول مختلف سال است.

**روش بررسی:** در این بررسی از سیستم‌های یادگیری ماشین و معیار تشابه در تعیین الگوی رفتار بیماری کووید-۱۹ در فصل‌های سال استفاده شد. مکان انجام مطالعه، بیمارستان موسی بن جعفر (ع) مشهد و زمان دقیق انجام مطالعه از اردیبهشت ۱۳۹۹ لغایت شهریور ۱۴۰۱ می‌باشد. علایم بیماران مبتلا با مجموعه داده تدوین شده مقایسه و تشابه بیماران در ماتریس شباهت تهیه و ضریب همبستگی جاکارد روی داده‌ها انجام شد. نهایتاً تحلیل سویه‌ها از ابتدای پیدایش تا آخرین سویه بررسی شد.

**یافته‌ها:** شاخص‌های عملکرد الگوریتم در روش تشابه جاکارد، معیار یادآوری با مقدار ۰/۹۴، معیار دقت با مقدار ۰/۸۶، معیار امتیاز F1 با مقدار ۰/۸۶ و معیار صحت با مقدار ۰/۷۶ را نشان داد. مهم‌ترین فاکتورهای مؤثر در بررسی، گلبول‌های سفید خون، پلاکت، RT PCR، CT SCAN، تنگی تنفس، تب، SPO2 و تعداد تنفس می‌باشند.

**نتیجه‌گیری:** در این مطالعه رفتار ویروس کووید-۱۹ با استفاده از الگوریتم‌های یادگیری ماشین با احتساب موقعیت جغرافیایی و فصلی در بیماران بررسی شد و یک الگوی واضح از ارتباط فصل‌ها در گسترش کووید-۱۹ مشخص گردید، به طوری که در هر فصل علایم مشخصی مشاهده شده است که با سویه همان فصل مطابقت دارد.

**کلمات کلیدی:** کووید-۱۹، بیماری، بومی‌شناسی، یادگیری ماشین، سویه‌ها.

\* نویسنده مسئول: تربت حیدریه، ۷ کیلومتر جاده تربت به مشهد، دانشگاه تربت حیدریه، اتاق ۲۱۰.  
تلفن: ۰۵۱-۵۱۲۴۰۱۶۵  
E-mail: salehan@torbath.ac.ir

### مقدمه

می‌توان به میزان بالای نوترکیبی ژنتیکی آن اشاره کرد که منجر به ظهور سویه‌های جدید خواهد شد.<sup>۱</sup> ویژگی عفونت‌های ویروسی در انسان و حیوانات بسیار متفاوت است. به طوری که آنفلوآنزای انسانی بیشترین شیوع را در فصل سرد سال نشان می‌دهد، برخی دیگر ویروس‌ها در تابستان اوج فصلی دارند.<sup>۲</sup> پس این احتمال می‌رود که ویروس کووید-۱۹ نیز ویژگی‌های فصلی را نشان دهد و تعیین اینکه

شیوع بیماری کووید-۱۹ بار سنگینی بر دوش جامعه پزشکی گذاشته است. با توجه به خصوصیت خاص ژنتیکی ویروس کرونا و همچنین تکثیر شدید آن، همه‌گیری و گسترش این ویروس برای ویروس‌شناسان زیاد دور از ذهن نبوده است.<sup>۱</sup> از جمله ویژگی‌های آن

فصول سرد سال، سیستم ایمنی بدن در یک وضعیت التهابی بالاتر قرار می‌گیرد که باعث افزایش بیماری‌های زمینه‌ای شده و منجر به حساسیت سیستم ایمنی به عفونت‌های ویروسی می‌گردد. ارتباط فصل‌ها و سلامتی، موضوعی است که توسط فیزیکی‌دان مشهور یونانی در کتاب «هوا، آب‌ها و مکان‌ها» به عکس‌العمل انسان در برابر تغییر فصل‌ها اشاره کرده و بیان نموده است که مرگ‌ومیر و بیماری‌های انسانی تا حد زیادی به الگوهای هوا بستگی دارد.<sup>۶</sup>

از سوی دیگر اخیراً مطالعات متعددی امکان تشخیص و پیش‌بینی کووید-۱۹ بر مبنای سیستم‌های هوشمند و ادغام آن با روش‌هایی مانند علائم بالینی، اقدامات پاراکلینیکی و روش‌های تشابه‌یابی بین بیماران را مورد تجزیه و تحلیل قرار داده‌اند. در ادامه برخی از این مطالعات مورد بررسی قرار خواهند گرفت.

Marateb و همکارانش، یک سیستم تشخیص قابل اعتماد کامپیوتری برای طبقه‌بندی بیماران مبتلا به کووید-۱۹ از غیر کووید در مرکز تحقیقاتی خورشید اصفهان را مطرح کردند.<sup>۷</sup> پنومونی، جمعیت‌شناسی، علائم و آزمایش خون در سیستم پیشنهادی مورد استفاده قرار گرفته است. آن‌ها از سه مجموعه داده مختلف جهت تجزیه و تحلیل داده‌ها استفاده کرده‌اند. مجموعه داده آنها شامل ۶۱ ویژگی و ۶۰ رکود بیمار است. روش پیشنهادی آنها شامل انتخاب ویژگی، نمونه‌گیری‌های متعدد و تکنیک‌های یادگیری گروهی است. در مورد پیش‌بینی موارد عفونت کووید-۱۹ در مقایسه با پنومونی‌های مشابه، بررسی‌هایی انجام شده است. به عقیده Jiang و همکارانش، بیماری عفونی کووید-۱۹ هنوز نامشخص است.<sup>۸</sup> پیش‌بینی دقیق بیماری تنها زمانی مشخص می‌شود که اپیدمی تمام شده باشد. آنها یک روش تشخیص کرونا ارائه دادند که براساس مدل پیش‌بینی، بیماری را شناسایی می‌کند، اما عوامل دیگری مانند تجزیه و تحلیل درمان پزشکی و اقدامات پیشگیرانه را بررسی نکرده‌اند. داده‌ها با توجه به ایالت یا منطقه‌ای که در حال تجزیه و تحلیل آن بودند متفاوتند، زیرا تعداد عوامل در یک منطقه خاص نیز از فردی به فرد دیگر متفاوت است. آنها بر روی آنالیز ایالتی تحقیق کردند که نتایج متنوع به‌طور جمعی برای هر حالت نمایان شده است. در خصوص معیارهای تشابه، Fethi مقایسه‌ای بین ۱۳ معیار تشابه را انجام داده است.<sup>۹</sup> وی مطالعه تجربی را به دو مرحله تکنیک‌های مبتنی بر کاربر و مبتنی بر آیتم تقسیم کرده است. یافته‌های وی نشان می‌دهند

آیا بیماری کووید-۱۹ فصلی است یا خیر، نقش خیلی زیادی در برنامه‌ریزی تولید واکسن خواهد داشت. مطابق با تحقیقات انجام شده، بادهای می‌توانند عامل پخش ویروس کووید-۱۹ و باعث انتقال بیماری شوند. ویروس‌ها در دمای معمولی  $23^{\circ}\text{C}$ ، از شش تا ۱۲ روز و در دمای  $4^{\circ}\text{C}$  تا ۱۰۰ روز در آب باقی می‌مانند. همچنین رطوبت ماندگاری ویروس را در محیط زیاد می‌کند.<sup>۴</sup>

هدف از این مقاله، بررسی مشخصات بیماران کووید-۱۹ و تعیین رفتار این ویروس در فصول مختلف سال می‌باشد. در این راستا بر روی داده‌های واقعی که از بیماران بیمارستان موسی‌ابن‌جعفر (ع) مشهد جمع‌آوری شده است، تمرکز خواهد شد و سعی می‌شود با بهره‌گیری از معیارهای تشابه که جزء الگوریتم‌های پایه مورد استفاده در تکنیک یادگیری ماشین هستند، میزان شباهت رفتار ویروس در فصول مختلف سال تعیین گردد. نتایج تحلیل داده‌های موجود می‌تواند برای کمک به تصمیم‌گیری دقیق و سریع در علم پزشکی مورد استفاده قرار گیرند. سایر بخش‌های این مقاله به شرح ذیل ساماندهی شده‌اند.

در بخش دوم، مطالعات مرتبط به معیار تشابه و میزان ارتباط بین پایداری بیماری کووید-۱۹ با آب و هوا و بومی‌شناسی بررسی می‌شود. بخش سوم، به بررسی روش پیشنهادی شامل روش تشخیص بیماری کووید-۱۹ و معرفی معیار تشابه جهت پیش‌بینی بیماری می‌پردازد و در انتها، مقایسه سویه‌ها و ارتباط آن‌ها با فصل‌های سال براساس معیار تشابه جاکارد که به تعیین میزان تشابه بیماران با مجموعه داده تدوین شده پرداخته است، انجام می‌گیرد. بخش چهارم به ارزیابی مدل و کارایی روش و همچنین تحلیل نتایج اختصاص دارد. در نهایت، در بخش پنجم و ششم بحث و نتیجه‌گیری انجام خواهد شد.

مطالعات مرتبط: با توجه به گسترش و پیدایش سویه‌های جدید ویروس کووید-۱۹، تحقیقات زیادی در مورد این بیماری همه‌گیر انجام شده است. در مطالعات پیشین نشان داده شده است که دما و رطوبت باعث ترشح بیشتر موکوس می‌شود و این امر باعث عدم حذف ویروس توسط مکانیزم‌های دفاعی میزبان شده و در نتیجه، باعث پایداری بیشتر ویروس در راه‌های تنفسی بیمار می‌شود. تابش خورشید با ایجاد تغییراتی در ریتم ملاتونین، درک فصلی بدن بیمار را تغییر می‌دهد.<sup>۵</sup> این ویروس منجر به نوسانات ژنی خاص می‌شود. در

مجموعه نمونه با توجه به اندیس جاکارد از تقسیم تعداد اشتراک دو مجموعه بر تعداد اجتماع آن‌ها به دست می‌آید.<sup>۱۰</sup> این الگوریتم معمولاً در حوزه‌های یادگیری ماشینی، داده‌کاوی و تحلیل داده برای مقایسه تشابه میان دو مجموعه مورد استفاده قرار می‌گیرد.

$$\text{sim}(u_a, u_b) = \frac{|I_{u_a} \cap I_{u_b}|}{|I_{u_a} \cup I_{u_b}|} \quad (1)$$

نتایج حاصل از ضریب همبستگی جاکارد بر روی مجموعه داده بیماران براساس معیارهای ارزیابی یادآوری، دقت، امتیاز FI و صحت مورد ارزیابی قرار گرفت تا میزان تشابه کلی بین داده‌های جمع‌آوری شده تعیین گردد (جدول ۱). مطابق با ارزیابی صورت گرفته، معیار یادآوری دارای مقدار ۰/۹۴، معیار دقت دارای مقدار ۱، معیار امتیاز FI دارای مقدار ۰/۸۶ و معیار صحت دارای مقدار ۰/۷۶ می‌باشند که این نتایج نشان می‌دهند تشابه زیادی بین ویژگی‌های تمامی مبتلایان به کووید-۱۹ وجود دارد.

تجزیه و تحلیل نتایج از لحاظ معیار تشابه: با تبدیل داده‌ها به مقادیر قابل فهم یادگیری ماشینی و ایجاد ماتریسی از شباهت بیماران، نهایتاً به مقایسه میزان شباهت بیماران با استفاده از معیار تشابه جاکارد جهت دانستن این‌که کووید-۱۹ از فصول سال پیروی می‌کند یا خیر؟ پرداخته شد. در بررسی کلی، در ابتدای پیدایش این بیماری به علت عدم آگاهی کافی و جمع‌آوری ناقص ویژگی‌های بیماران، تشابه زیادی بین نمونه‌ها وجود ندارد. ولی در ادامه، تشابه بیماران به‌وضوح قابل مشاهده می‌باشد. حتی میزان بالای مشابهت بین ویژگی‌های بیماران در تابستان، نشان می‌دهد که رفتار فصلی این ویروس در تابستان بسیار قابل توجه است.

تشابه بیماران کرونایی در فصول مختلف سال: طبق گزارش سازمان بهداشت جهانی یا WHO، بیش از ۲۰۰ نوع ویروس کووید-۱۹ وجود دارد.

محققان این ویروس را به چهار دسته تقسیم کرده‌اند: آلفا، بتا، گاما و دلتا، که تفاوت اصلی در تنوع کووید-۱۹ در شدت بیماری در مبتلایان است.<sup>۱۱</sup> از نظر مولکولی، تفاوت‌هایی در روند ژنتیکی و همچنین گیرنده‌های سلولی ویروس وجود دارد. تاکنون هفت سویه آن در فصل‌های مختلف در ایران مشاهده شده است که به تحلیل نتایج آن به کمک معیار تشابه جاکارد (Jaccard Similarity Index, JSI) می‌پردازیم:

تشابه سویه ۱ در بیماران کرونایی: اولین سویه بیماری کووید-۱۹

روش‌های پالایش مشارکتی مبتنی بر کاربر و مبتنی بر آیتم معیار تشابه یکسانی را ندارند.

## روش بررسی

در همه‌گیری‌ها، شناسایی سریع افراد مبتلا به کووید-۱۹ بسیار بااهمیت است، زیرا براساس اطلاعات جمع‌آوری شده می‌توان از شیوع مجدد بیماری جلوگیری نمود. انواع ویروس کرونا در فصول مختلف، خواص فصلی بارزی دارند، ولی در بعضی دیگر این ویروس‌ها از الگوی چندان ثابتی پیروی نمی‌کنند. به‌همین منظور از الگوریتم‌های یادگیری ماشینی و هوش مصنوعی با استفاده از داده‌های یکپارچه جهت پیش‌بینی و تشخیص الگوی بیماری استفاده می‌شود. بدین گونه که الگوریتم، علایم و داده‌های بیماران مبتلا به کووید-۱۹ را با داده‌های آموزش داده‌شده تجزیه و تحلیل می‌کند و تشابه آنها را استخراج نموده و با استفاده از این معیارها، پیش‌بینی علایم بیماری در هر فصل انجام می‌شود. در این مطالعه با داده‌های بومی امکان مقایسه و تحلیل دقیق‌تری از ارتباط فصول بر شدت و روند بیماری ارایه می‌شود. در راستای این مطالعه ابتدا به بررسی معیارهای ارزیابی مدل، سپس تحلیل نتایج معیار تشابه و در ادامه تشابه بیماران در فصول مختلف و مقایسه بین آن‌ها پرداخته خواهد شد. مجموعه داده مورد استفاده در این مقاله که در وب‌سایت Data.World بارگذاری شده است، براساس داده‌های اپیدمیولوژیک بیماران مبتلا به کووید-۱۹ بیمارستان موسی بن جعفر (ع) در مشهد جمع‌آوری شده است. این مجموعه دارای ۱۳۷۹ رکورد و ۷۷ ویژگی بیماران مبتلا به ویروس است که شامل علایم بالینی بیمار، فاکتورهای خونی، همچنین وضعیت بیمار مبنی بر ترخیص یا فوت، تست RT PCR و CT Scan قفسه صدری می‌باشند. معیارهای علمی در تشخیص بیماری کووید-۱۹ در مجموعه داده استفاده شده عبارتند از علایم بالینی، یافته‌ها و گزارش‌های تصویربرداری، تست PCR؛ و تست‌های آزمایشگاهی.

بررسی معیارهای ارزیابی: در این بررسی، داده‌ها از تجربیات گذشته آموزش داده شده‌اند، سپس با استفاده از روش‌های تشابه جاکارد، تصمیم خودکار گرفته می‌شود. ضریب همبستگی جاکارد: ضریب جاکارد معیاری برای مقایسه شباهت یا تفاوت مجموعه نمونه‌های داده است (فرمول ۱). در این رابطه میزان شباهت دو

تشابه سویه ۳ در بیماران کرونایی: سویه سوم از پاییز سال ۱۳۹۹ شروع شد و به‌عنوان نوع ویروس B.1.1.413 شناخته شده است. بیشترین علائم در بین ۱۴۰ نفر مبتلا به کووید-۱۹ در فصل پاییز، تب (۷۳/۵۷٪)، تنگی تنفس (۶۳/۵۴٪)، سرفه (۶۴/۲۸٪)، ضعف و بی‌حالی (۴۸/۵۷٪)، لرز (۵۷/۱۴٪)، گلودرد (۳۶/۴۲٪) و بیماری زمینه‌ای (۳۴/۲۸٪) مشاهده شد (نمودار ۱، قسمت ج).

تشابه سویه ۴ در بیماران کرونایی: سویه چهارم از بهار سال ۱۴۰۰ شروع شد و به‌عنوان نوع ویروس B.1.1.7 شناخته شده است. بیشترین علائم در بین ۶۶ نفر مبتلا به کووید-۱۹ در فصل بهار، تب (۶۱/۷۶٪)، تنگی تنفس (۵۲/۹۴٪)، سرفه (۴۱/۱۷۶٪)، لرز (۲۶/۴۷٪) و بیماری زمینه‌ای (۱۰۰٪) مشاهده شد (نمودار ۱، قسمت د).

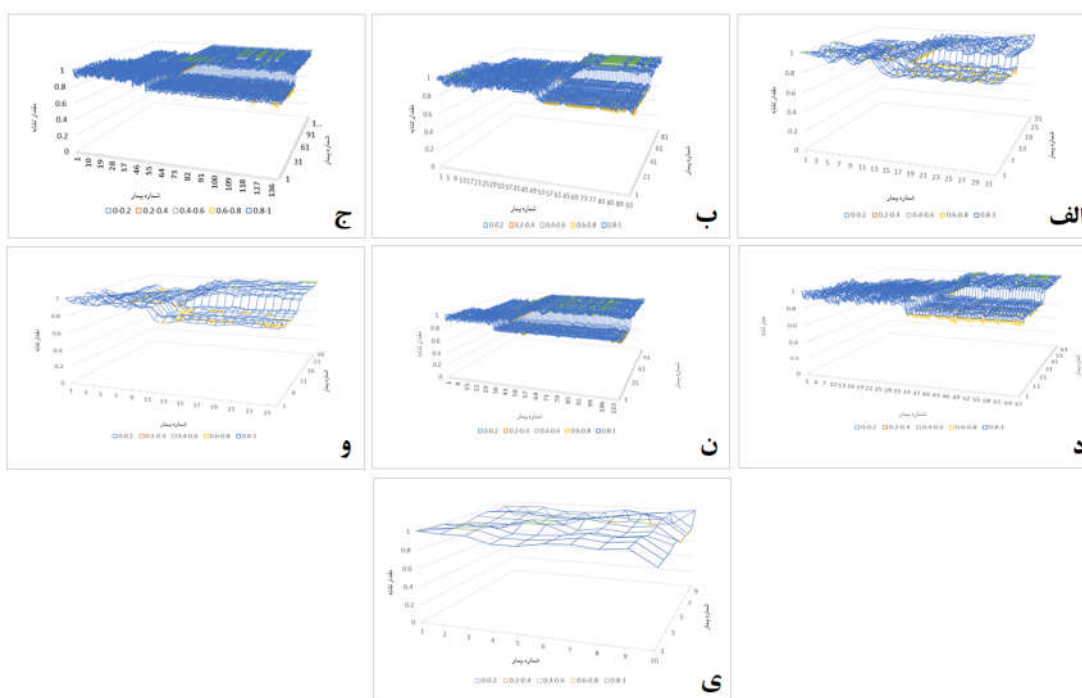
تشابه سویه ۵ در بیماران کرونایی: سویه پنجم از تابستان سال ۱۴۰۰ شروع شد. اکثر مراجعین در سویه پنجم به این بیمارستان مراجعه کرده‌اند و آمار فوتی در این سویه بیشتر بوده است. بیشترین علائم در بین

جدول ۱: معیارهای ارزیابی روش تشابه جاکارد

	Preccition	F1	ACC	RECALL
Jaccard	۰/۷۶	۰/۸۶	۰/۸۱	۰/۹۴

از بهار سال ۱۳۹۹ شروع شد که به‌عنوان ویروس Base شناخته شده است. در فصل بهار، بیشترین علائم از بین ۳۲ نفر مبتلا به کووید-۱۹ در تب (۵۹/۳۸٪)، تنگی تنفس (۴۰/۶۲٪)، سرفه (۳۱/۲۵٪) و بیماری زمینه‌ای (۶۸/۷۵٪) مشاهده شد (نمودار ۱، قسمت الف).

تشابه سویه ۲ در بیماران کرونایی: سویه دوم از تابستان سال ۱۳۹۹ شروع شد و به‌عنوان نوع ویروس B.1.36 شناخته شده است. بیشترین علائم در بین ۹۲ نفر مبتلا به کووید-۱۹ در فصل تابستان، تنگی تنفس (۵۰٪)، سرفه (۴۴/۵۶٪)، ضعف و بی‌حالی (۳۱/۵۲٪) و بیماری زمینه‌ای (۵۶/۵۲٪) مشاهده شد (نمودار ۱، قسمت ب).



نمودار ۱: مقایسه میزان تشابه بیماران کرونایی با استفاده از معیار تشابه جاکارد در: الف) سویه ۱، ب) سویه ۲، ج) سویه ۳، د) سویه ۴، ه) سویه ۵، و) سویه ۶ و ز) سویه ۷

آمار فوتی این سویه بسیار پایین تر مشاهده شده است. بیشترین علایم در بین ۱۱ نفر مبتلای قطعی به کووید-۱۹ در فصل تابستان، تب (۷۲/۷۳٪)، تنگی تنفس (۳۶/۳۶٪)، بدن درد (۴۵/۴۵٪) و بیماری زمینه‌ای (۷۲/۷۲٪) مشاهده شد (نمودار ۱، قسمت ی).

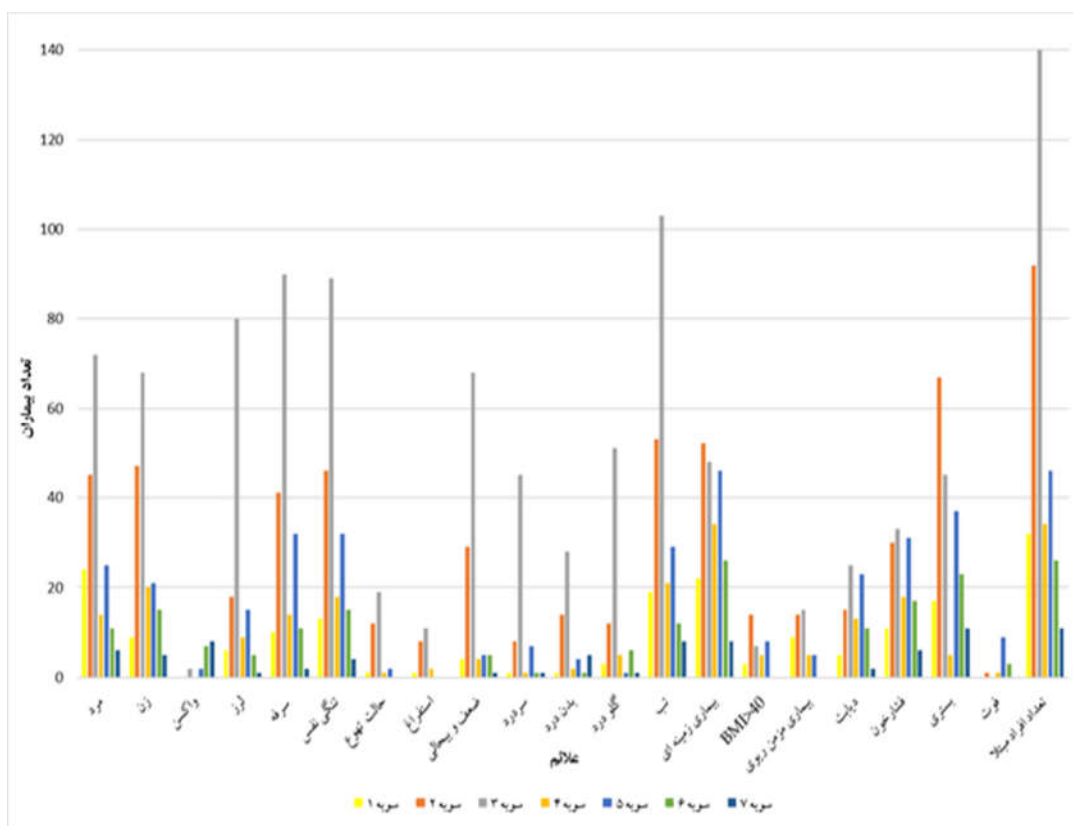
## یافته‌ها

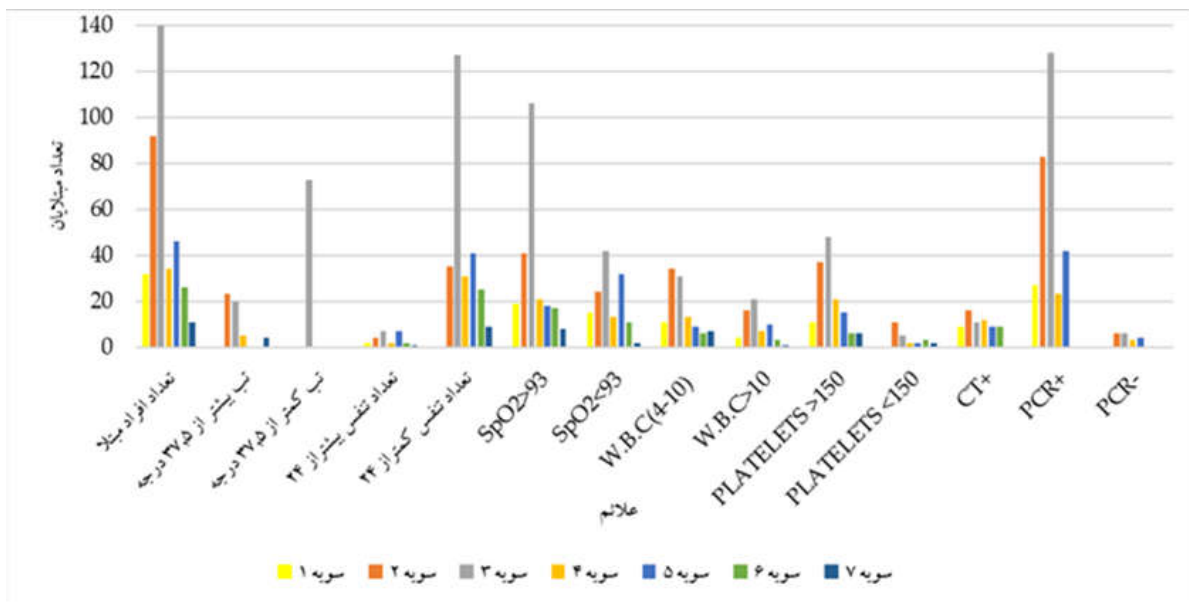
مقایسه علایم بیماری در سویه‌های مختلف: علایم بیماران در اکثر سویه‌ها تقریباً مشابه بودند. در سویه‌های چهار، پنج و شش بیماری زمینه‌ای مشترک و به میزان (۱۰۰٪) و پس از آن علایم تب، تنگی تنفس و سرفه دارای بیشترین میزان گزارش بود. نمودارهای ۲ و ۳، به مقایسه ۳۵ ویژگی مختلف ۷ سویه فوق‌الذکر برای ۱۴۰ بیمار و مبتلای به کووید-۱۹ می‌پردازند. همانطور که در این دو نمودار قابل

مبتلایان قطعی کووید-۱۹ در فصل تابستان، تب (۶۳/۰۴٪)، تنگی تنفس (۶۹/۵۶٪)، سرفه (۶۹/۵۶٪) و بیماری زمینه‌ای (۱۰۰٪) مشاهده شد (نمودار ۱، قسمت ن).

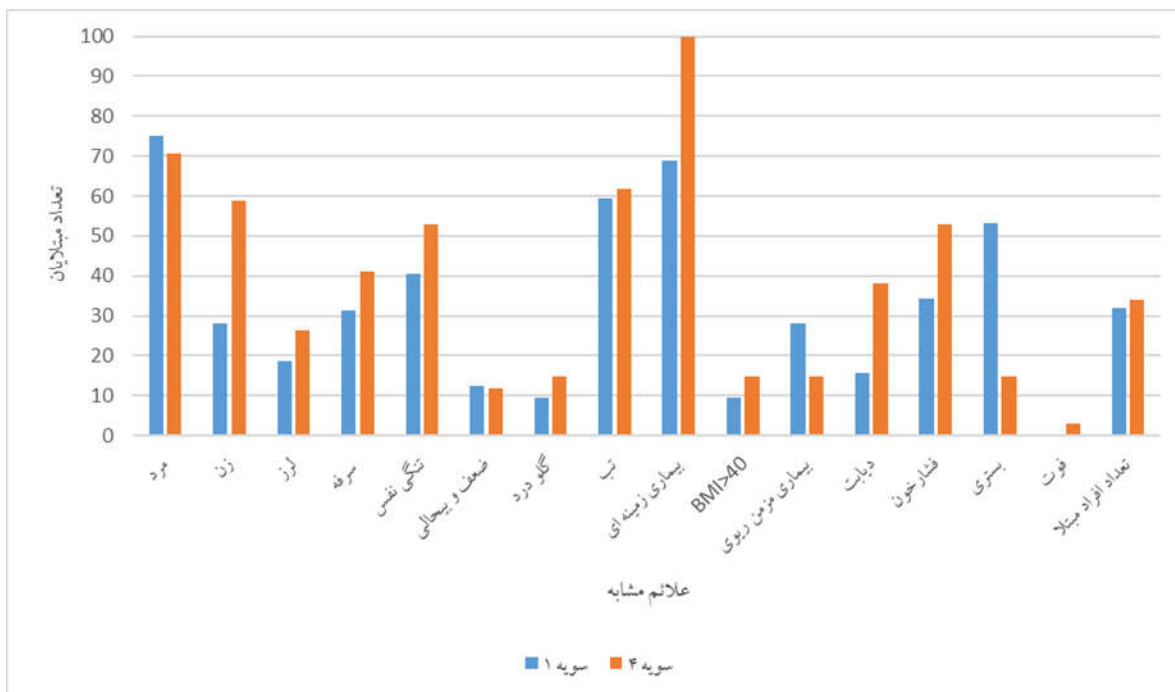
تشابه سویه ۶ در بیماران کرونایی: سویه ششم از پاییز سال ۱۴۰۰ شروع شد و پس از سویه پنجم بالاترین میزان مبتلایان به کرونا در این بیمارستان را به خود اختصاص داده است. بیشترین علایم در بین ۳۶ نفر مبتلا قطعی کووید-۱۹ در فصل پاییز، تب (۴۶/۱۵٪)، تنگی تنفس (۵۷/۶۹٪)، سرفه (۴۲/۳۰٪) و بیماری زمینه‌ای (۱۰۰٪) مشاهده شد (نمودار ۱، قسمت و).

تشابه سویه ۷ در بیماران کرونایی: سویه هفتم از تیرماه سال ۱۴۰۱ شروع شد و به‌عنوان ویروس B.1.617.2\_EX شناخته شده است. در حال حاضر بیماران به این نوع ویروس مبتلا می‌شوند. به خاطر واکنش‌های انبوه در دوزهای مختلف، خطر ابتلا به این ویروس کم‌تر و

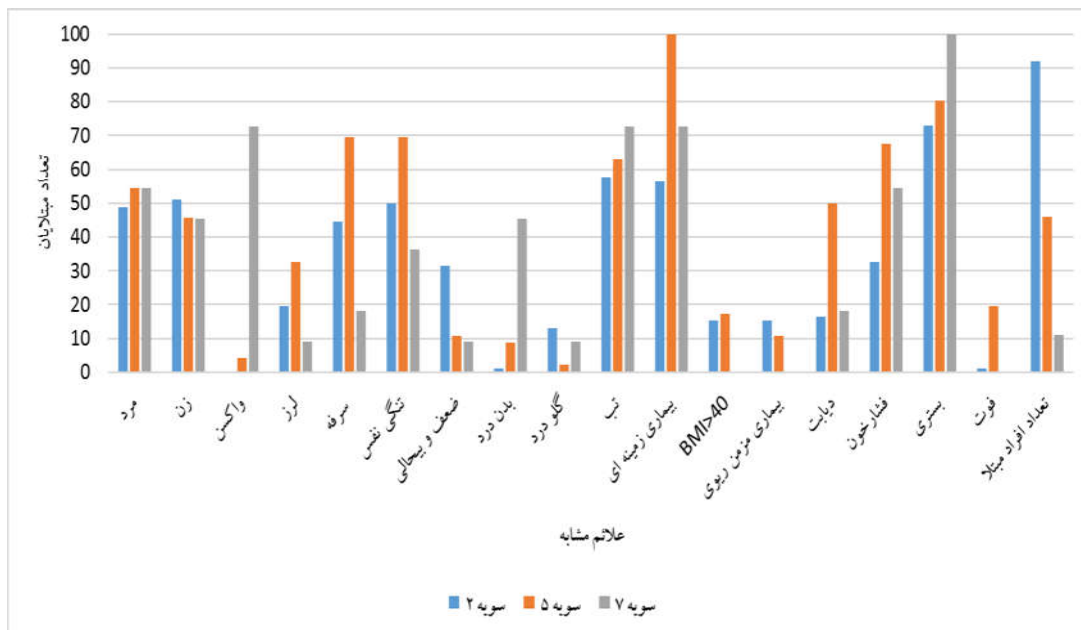




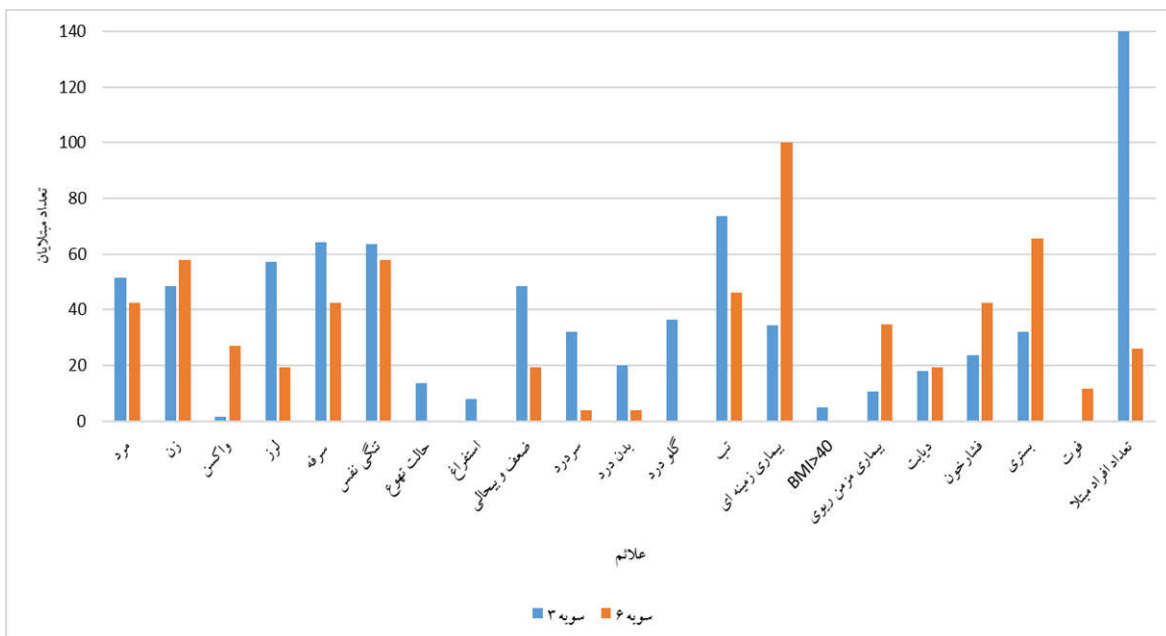
نمودار ۳: مقایسه ۱۴ علائم مبتلایان به کووید-۱۹ در سویه‌های مختلف



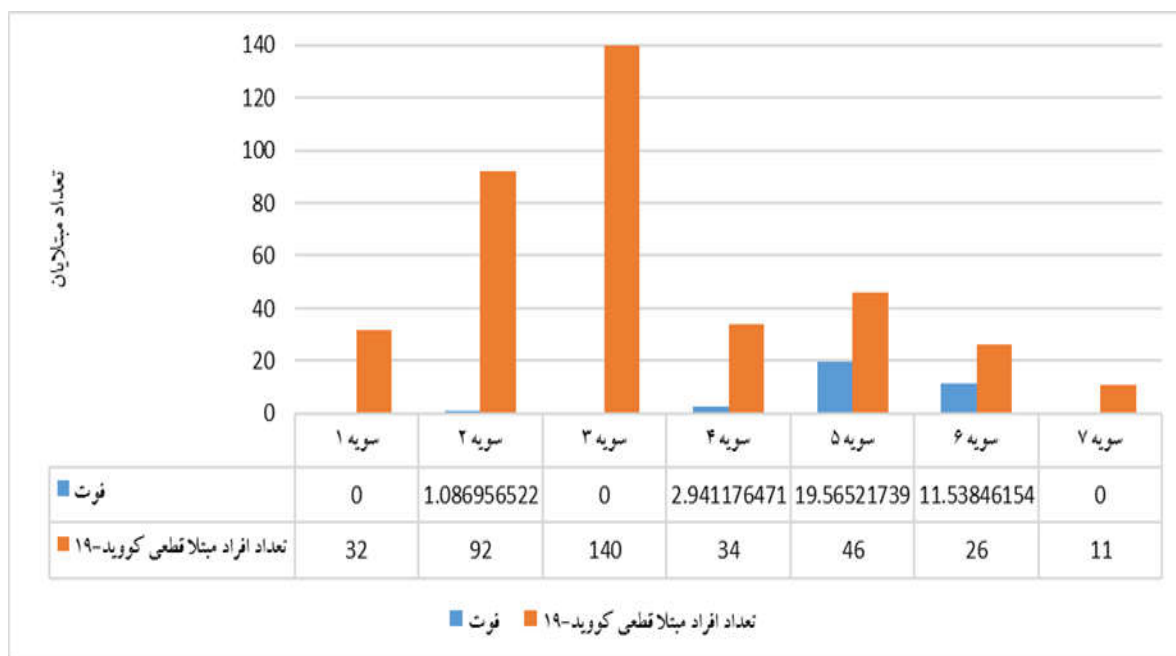
نمودار ۴: مبتلایان در فصل بهار



نمودار ۵: مبتلایان به کرونا در فصل تابستان



نمودار ۶: مبتلایان به کرونا در فصل پاییز



نمودار ۷: مقایسه درصد افراد فوت شده در سویه‌های مختلف

۵). همچنین سویه‌های ۵ و ۷ بسیار هم‌رفتاراند. - مبتلایان در فصل پاییز: در مقایسه بین سویه ۳ و ۶ در فصل پاییز، بیشترین تشابه برای علائم جنسیت مرد و زن، تب، سرفه و بیماری زمینه‌ای (دیابت)، همچنین تنگی نفس مشاهده شد (نمودار ۶).

- میزان افراد فوتی در سویه‌های مختلف: بیشترین تعداد مبتلایان فوتی در سویه پنج و پس از آن در سویه شش مشاهده شد (نمودار ۷).

## بحث

انتقال ویروس کووید-۱۹ به عوامل متعددی از جمله به تعامل انسان‌ها وابسته است و باید مورد بررسی قرار گیرد. شواهد مجموعه داده جمع‌آوری شده گویای این است که افراد مبتلا به کووید-۱۹، شمارش لنفوسیتی پایینی دارند و با نتایج مطالعات اخیر بسیار

مشاهده است، در اکثر علائم، بیشترین مقادیر مربوط به سویه سوم است. در این خصوص نکته قابل توجه آن است که علیرغم بالا بودن مقادیر علائم و حتی با وجود حداکثر تعداد افراد مبتلا در این سویه، اما کمترین میزان مرگ‌ومیر به این سویه اختصاص دارد. بیشترین میزان مرگ‌ومیر در بین کل سویه‌های مورد مطالعه نیز به سویه ۵ اختصاص دارد.

مقایسه بیماران مبتلا در فصل‌های مشابه: با استفاده از مجموعه داده بیماران کووید-۱۹ به تشابه سویه‌های مختلف در فصل‌های مشابه توسط معیار تشابه جاکارد پرداخته شد.

- مبتلایان در فصل بهار: در مقایسه بین سویه ۱ و ۴ در فصل بهار، بیشترین تشابه برای علائم جنسیت مرد، تب و بیماری زمینه‌ای گزارش شد (نمودار ۴).

- مبتلایان در فصل تابستان: در مقایسه بین سویه‌های ۲، ۵، و ۷ در فصل تابستان، بیشترین تشابه برای علائم در جنسیت مرد و زن، تب و بیماری زمینه‌ای، بستری و تنگی نفس حاصل گردید (نمودار

تعداد زیادی از مبتلایان دارای بیماری زمینه‌ای از جمله دیابت و فشارخون می‌باشند.

در بررسی سویه‌ها در فصل‌های یکسان سال، بیشترین شباهت در تابستان مشاهده شد.

شایعترین علامت بالینی در بین بیماران شامل تنگی نفس ۵۲/۳۹٪، تب ۵۱/۱۵٪، سرفه ۴۷/۰۸٪، لرز ۲۳/۰۱٪، ضعف و بی‌حالی ۲۰٪، گلودرد ۱۵/۹۲٪، بدن‌درد ۹/۲۰٪، حالت تهوع ۶/۳۷٪، سردرد ۶/۱۹٪ و در بعضی موارد کاهش اشتها، تعریق، خارش بدن، اسهال، و کاهش حس بویایی مشاهده شد که این موارد با نتایج مطالعات پیشین همخوانی دارد.

عارضه‌های اسهال و از دست دادن حس بویایی جزء علائم اضافی برخی از بیماران مبتلا به کووید-۱۹ بودند و همچنین ضعف و بی‌حالی، تب و لرز علائم اولیه اکثر سویه‌ها گزارش شد. به‌طورکلی به‌نظر می‌رسد این بیماری می‌تواند با جنسیت، سن و بیماری‌های زمینه‌ای فرد مرتبط باشد.

بیشترین و کمترین درگیری تنفسی به‌ترتیب در پیک چهارم، پنج و شش مشاهده شد.

شایعترین یافته پاراکلینیکی، تعداد لنفوسیت‌ها و میزان مرگ‌ومیر در سویه پنج با ۱۹/۵۶٪ بود.

بیماران بستری در پیک سوم بیشتر بودند، ولی مرگ‌ومیری مشاهده نشده است.

در سویه دوم بیشترین PCR- (۴۸/۹۱٪) مشاهده شد که دارای CT+ (۱۷/۳۹٪) بودند و در این سویه برخلاف مطالعات پیشین، مبتلایان، Spo2 بیشتر از ۹۳(۴۴/۵۷٪) داشتند.

در تمامی سویه‌های بیماری، افزایش پلاکت گزارش شده است. بیماری زمینه‌ای می‌تواند دلیلی برای افزایش میزان مرگ‌ومیر باشد. شاید افزایش و در دسترس بودن ظرفیت تست‌های تشخیصی در سویه‌های آخر و زمان طولانی ابتلا به آن دوره، از دلایل شناسایی حداکثری تعداد مبتلایان باشند، ولی بیشترین علائم بالینی در اکثر سویه‌ها مشابه بودند.

سپاسگزاری: این مقاله، حاصل یک پژوهش با کد اخلاق IR.IMAMREZA.REC.1402.012 می‌باشد که با حمایت دانشگاه تربت‌حیدریه، دانشگاه امام رضا (ع)، و با همکاری پرسنل و مدیران محترم بیمارستان موسی‌ابن‌جعفر (ع) مشهد انجام شده است.

همخوانی دارد. به علت تفاوت در خصوصیات و علائم بیماران در مناطق مختلف و عدم وجود مجموعه داده یکپارچه از یک اقلیم خاص، در این مقاله یک مطالعه مقایسه‌ای بر روی بیماران مبتلا به کووید-۱۹ در بیمارستان موسی‌ابن‌جعفر (ع) شهر مشهد انجام شد و در این راستا، یک مجموعه داده جمع‌آوری و برای استفاده در سایر تحقیقات، در منابع آنلاین بارگذاری گردید.

نتیجه‌گیری: پیدایش و شیوع شدید کووید-۱۹، یک تهدید بالینی عمومی و بهداشتی در سراسر جهان عنوان شد. علیرغم آنکه با گذشت زمان، دانش در مورد این ویروس زیاد شده است، با این حال با پیدایش سویه‌های جدید بیماری، باز هم افزایش دانش در همه زمینه‌های این ویروس از جمله تأثیرات محیطی، آب و هوایی و بومی‌شناسی نیاز است تا الگو و رفتار آن به‌طور کامل شناسایی شود. یافته‌های اپیدمیولوژیک در مورد تأثیر عوامل هواشناسی فصلی در سرعت انتقال کووید-۱۹ مورد بررسی قرار گرفته شد. ویروس کووید-۱۹ یک ویروس غلاف‌دار است که از لایه‌ای روغنی پوشیده شده است. این لایه روغنی باعث حساسیت بیشتر ویروس در گرما شده و در شرایط سردتر لایه روغنی حالت سفت‌تر به خود می‌گیرد و باعث ماندگاری بیشتر ویروس خواهد شد و به‌همین دلیل بیشتر ویروس‌های غلاف‌دار خواص فصلی دارند. می‌توان به این نتیجه رسید که یک الگوی واضح از ارتباط فصل‌ها در گسترش کووید-۱۹ وجود دارد. در این تحقیق، در هر فصل علائم مشخصی از کووید-۱۹ مشاهده شد که با سویه همان فصل مطابقت دارد و بر همین اساس اگر علائم، الگوی دیگری را نشان دهند که با علائم سویه‌های جمع‌آوری شده مشابهتی نداشته باشد، می‌توان نتیجه گرفت با یک سویه یا بیماری جدید مواجه هستیم.

با شناسایی به‌موقع رفتار بیماری، می‌توان اقدامات پیشگیرانه از جمله رعایت فاصله اجتماعی، آموزش‌های بهداشتی و واکسیناسیون جمعی را انجام داد تا در مدیریت بهتر و کنترل شیوع بیماری بسیار مؤثر واقع شود. در این بخش به نتایجی که طی انجام تحلیل اکتشافی داده‌های مجموعه‌داده منتخب به‌دست آمده است، به اجمال اشاره می‌گردد:

با افزایش دمای هوا، علائم تنگی تنفس و سرفه در بیماری کووید-۱۹ گزارش شده است.

میانگین سن مبتلا به کووید-۱۹ در تمامی فصل‌ها ۴۸/۷ سال گزارش شد.

## References

1. Byun WS, Heo SW, Jo G, Kim JW, Kim S, Lee S, Park HE, Baek JH. Is coronavirus disease (COVID-19) seasonal? A critical analysis of empirical and epidemiological studies at global and local scales. *Environmental Research* 2021;196:110972.
2. Tavakoli, A, Vahdat, K, Keshavarz, M, Novel coronavirus 2019 (COVID-19): an emerging infectious disease in the 21st century, *sid.ir.*, 2020.
3. Shaman J, Kohn M. Absolute humidity modulates influenza survival, transmission, and seasonality. *Proceedings of the National Academy of Sciences* 2009;106(9):3243-8.
4. Mohammadi F, Pourzamani H, Karimi H, Artificial neural network and logistic regression modelling to characterize COVID-19 infected patients in local areas of Iran, *Biomedical journal, Elsevier* 2021.
5. Nelson RJ, Drazen DL. Melatonin mediates seasonal adjustments in immune function. *Reproduction Nutrition Development* 1999;39(3):383-98.
6. Rahnama, MR, Bazargan, M, Article on the analysis of spatio-temporal patterns of Covid 91 virus epidemic and its risks in Environmental risk management (Knowledge of former risks), 7(2), (2020), pp. 927-999
7. Marateb H, Reza, Ziaie N, Farzad, Mohebian M.R, Sami R, Javanmard Haghjooy Sh, Dehghan N.F, Akafzadeh-Savari M, Mansourian M, Angel M.M, Wolkewitz M and Binder H, Automatic Classification Between COVID-19 and Non-COVID-19 Pneumonia Using Symptoms, Comorbidities, and Laboratory Findings: *The Khorshid COVID Cohort Study* 2020
8. Jiang F, Deng L, Zhang L, Cai Y, Cheung CW, Xia Z. Review of the clinical characteristics of coronavirus disease 2019 (COVID-19). *Journal of general internal medicine* 2020;35:1545-9.
9. Fkih F. Similarity measures for Collaborative Filtering-based Recommender Systems: Review and experimental comparison. *Journal of King Saud University-Computer and Information Sciences* 2022;34(9):7645-69.
10. Patra BK, Launonen R, Ollikainen V, Nandi S. A new similarity measure using Bhattacharyya coefficient for collaborative filtering in sparse data. *Knowledge-Based Systems* 2015;82:163-77.
11. International Committee on Taxonomy of Viruses ICTV. (Accessed January 19, 2020, at <https://talk.ictvonline.org/>).

## Seasonal impact analysis on COVID-19 severity: insights from homogeneous epidemiological hospital data using the jaccard correlation coefficient

Ameneh Javanmard M.Sc.<sup>1</sup>  
Alireza Salehan Ph.D.<sup>2\*</sup>

1- Department of Computer Engineering, University of Imam Reza, Mashhad, Iran.

2- Department of Computer Engineering, University of Torbat Heydarieh, Heydarieh, Iran.

\* Corresponding author: Room 210, University of Torbat Heydarieh, the 7th KM of Torbat Heydarieh-Mashhad Road, Torbat Heydarieh, Iran.  
Tel: +98-51-51240165  
E-mail: salehan@torbath.ac.ir

### Abstract

Received: 27 Sep. 2023 Revised: 03 Oct. 2023 Accepted: 14 Dec. 2023 Available online: 22 Dec. 2023

**Background:** Coronaviruses were discovered in 1960. Large-sized living organisms from the Coronaviridae family, with single-stranded RNA of animal origin. Coronaviruses in humans can cause mild respiratory illness or severe respiratory illness. In 2020, the World Health Organization declared COVID-19 a global pandemic. The aim of this study is to use the Jaccard similarity coefficient to determine the similarity of COVID-19 behavior patterns in different seasons of the year.

**Methods:** This study used machine learning systems and similarity metrics to determine the behavior pattern of COVID-19 in different seasons of the year. The location of research was the Mousa ibn Ja'far Hospital in Mashhad, and the time was from May 2020 to August 2021. The symptoms of affected patients were compared with the compiled dataset, and the similarity of patients was prepared in a similarity matrix, and the Jaccard correlation coefficient was calculated on the data. Finally, the analysis of strains from the beginning of emergence to the latest strain was examined. The performance indicators of the algorithm in the Jaccard similarity method showed a recall metric with a value of 0.94, a precision metric with a value of 1, an F1 score with a value of 0.86, and remove accuracy metric with a value of 0.76. The most important factors in the investigation include white blood cells, platelets, RT-PCR, CT SCAN, shortness of breath, fever, SPO2, and respiratory rate.

**Results:** The transmission of the COVID-19 virus depends on several factors, including human interaction. The evidence of the collected data shows that people with COVID-19 have low lymphocyte count and it is very consistent with the results of recent studies. Due to the lack of a dataset, a comparative study was conducted and a dataset was collected.

**Conclusion:** This study, leveraging machine learning algorithms, identified a clear seasonal correlation in the spread of COVID-19. Considering geographical and seasonal variations among patients, distinct symptoms were observed in each season corresponding to the prevalent strain during that period.

**Keywords:** covid-19, disease, epidemiology, machine learning, strains.

Copyright © 2024 Javanmard et al. Published by Tehran University of Medical Sciences.



This work is licensed under a Creative Commons Attribution-Non-Commercial 4.0 International license (<https://creativecommons.org/licenses/by-nc/4.0/>). Non-commercial uses of the work are permitted, provided the original work is properly cited.