



ارائه شده توسط:

سایت ترجمه فا

مرجع جدیدترین مقالات ترجمه شده

از نشریات معتبر

روش ها و ابعاد ارزیابی کیفیت داده های پرونده های سلامت الکترونیک: امکان استفاده

مجدد از داده های پرونده های سلامت الکترونیک برای تحقیقات بالینی

چکیده :

هدف : بازنگری روش ها و ابعاد ارزیابی کیفیت داده ها در زمینه استفاده مجدد از داده های پرونده سلامت الکترونیک برای تحقیقات بالینی.

مواد و روش ها : مرور منابع تحقیقات بالینی در رابطه با روش شناسی ارزیابی کیفیت داده های پرونده های سلامت الکترونیک انجام شد. با استفاده از یک فرایند تکراری، ابعاد کیفیت داده های اندازه گیری شده خلاصه و دسته بندی شده و روش های ارزیابی مورد استفاده قرار گرفت.

نتایج : پنج بعد کیفیت داده ها شناسایی شدند که شامل فشردگی، صحت، تطابق، معقول بودن و رواج و هفت مقوله گسترده از روش های ارزیابی کیفیت داده ها: مقایسه با استاندارد های طلایی، همخوانی عناصر داده ها، همخوانی منابع داده ها، مقایسه توزیع، کنترل روایی، بازنگری ورود و حضور عناصر بودند.

بحث : بررسی روش هایی که از طریق آن محققان بالینی به ارزیابی کیفیت و مناسبت داده های پرونده های سلامت الکترونیک برای تحقیقات بالینی نشان می دهد که کیفیت داده ها دارای ویژگی های اساسی است و اندازه گیری این ویژگی ها و ابعاد آن ها سخت می باشد. محققان علاقه مند به استفاده مجدد از داده های پرونده سلامت الکترونیک برای تحقیقات بالینی برای در نظر گرفتن پذیرش طبقه بندی کیفیت داده های سلامت الکترونیک، آگاهی از تخصصی بودن و کیفیت داده ها، تلفیق کارها بر روی ارزیابی کیفیت داده ها از زمینه های دیگر و پذیرش روش های اماری، سیستماتیک و تجربی ارزیابی کیفیت داده ها می باشند.

نتیجه گیری : امروزه تعمیم پذیری بالقوه یا پیوستگی اندکی در روش های مورد استفاده برای ارزیابی کیفیت داده های پرونده های سلامت الکترونیک وجود دارد. در صورتی که استفاده مجدد از داده های پرونده سلامت الکترونیک برای تحقیقات بالینی پذیرفته شود، محققان بایستی از روش های معتبر و سیستماتیک برای ارزیابی کیفیت داده های پرونده های سلامت الکترونیک استفاده کنند.

از آنجا که پذیرش و استفاده از پرونده های الکترونیکی سلامت موجب دسترسی آسان تر و ترکیب ساده تر داده های بالینی شده اند، توجه روز افزونی به انجام تحقیقات با داده های جمع اوری شده در طی روند مراقبت های بالینی صورت گرفته است.(1-2). موسسه ملی سلامت، بر افزایش استفاده مجدد از پرونده های الکترونیک برای سلامت تاکید کرده است و انجمن تحقیقات بالینی به طور فعالانه روش هایی را برای استفاده ثانویه از داده های بالینی جست و جو کرده است(3). پرونده های الکترونیک سلامت عملکرد فراتری از رجیستری های موجود و انبار های داده ها از نظر حجم داشته و استفاده مجدد از این داده ها موجب کاهش هزینه ها و ناکارآمدی های مرتبط با تحقیقات بالینی می شود. همانند سایر اشکال تحقیقات گذشته نگرانه، مطالعاتی که از داده های پرونده های الکترونیک استفاده می کنند، نیازی به گزینش بیمار یا جمع اوری داده ها ندارند که هر دوی این فرایندها، پرهزینه و زمان بر هستند. داده های مربوط به پرونده های سلامت الکترونیک نیز زمینه را برای دسترسی به مراقبت های پزشکی، وضعیت و برایندهای یک جمعیت متنوع هموار ساخته اند، جمعیتی که معرف بیماران واقعی هستند. استفاده ثانویه از داده های جمع اوری شده در پرونده های الکترونیک سلامت، گامی مهم در کاهش هزینه های تحقیقاتی، افزایش تحقیقات مبتنی بر بیمار و تسریع سرعت کشفیات پزشکی جدید است.

علی رغم این مزیت ها، استفاده مجدد از داده های پرونده الکترونیک از سلامت با تعدادی از عوامل از جمله نگرانی های مربوط به کیفیت داده ها و مناسبت آن ها برای تحقیق، محدود شده است. به طور کلی پذیرفته شده است که، در نتیجه تفاوت ها در اولویت های بین شرایط تحقیقاتی و بالینی، داده های بالینی با دقت یکسان و مشابه با داده های پژوهشی(4) ثبت نمی شوند. به علاوه، بارمن(5) بیان کرده است که ظهور فناوری اطلاعات سلامت نظیر پرونده های سلامت الکترونیک منجر به بهبود کیفیت داده ها به جای ثبت مجدد مقادیر زیادی از داده های بد شده است. به دلیل این نگرانی ها در خصوص کیفیت داده ها، وان در لی(6) در مورد استفاده مجدد از داده های بالینی برای تحقیقات هشدار داده اند و اولین قانون انفورماتیک را پیشنهاد کرده است. داده ها بایستی تنها بر اساس هدفی که جمع اوری شده اند مورد استفاده قرار گیرند. اگرچه این نگرانی ها در خصوص کیفیت داده ها از زمان معرفی و ظهور پرونده های سلامت الکترونیک وجود داشته اند، با این حال هیچ گونه اجماعی در خصوص کیفیت داده های بالینی الکترونیک و یا مفهوم کیفیت داده ها در زمینه پرونده های سلامت الکترونیک وجود نداشته است. یکی از پذیرفته شده ترین مفاهیم کیفیت توسط جوران(7) ارائه شده است که بیان داشته

است که کیفیت به صورت " مناسب بودن برای استفاده " تعریف میشود. در زمینه کیفیت داده ها، این بدین معنی است که داده ها در زمان نیاز دارای کیفیت کافی باشند.

مطالعات گذشته در خصوص کیفیت داده های پرونده سلامت الکترونیک، نتایج بسیار متفاوتی را گزارش کرده اند. هوگان و وگنر(8) در مرور منابع خود در سال 1997، به این نتیجه رسیده اند که صحت داده ها بین 44 و 100 درصد و کامل بودن آن ها بین 1.1 و 100 درصد بسته به مفاهیم بالینی مطالعه شده متغیر است. به طور مشابه، تیرو و همکاران(9) در محاسبه حساسیت انواع مختلف داده های پرونده های سلامت الکترونیک در منابع، به این نتیجه رسیده اند که این مقادیر بین 0.26 و 1.00 متغیر بوده است. چان و همکاران(10) در سال 2010 به بررسی کیفیت مفاهیم بالینی در موسسات مختلف پرداخته و تغییرات زیادی را مشاهده کرده اند. برای مثال، کامل بودن اطلاعات فشار خون بین 0.1 و 51 درصد متغیر است. به دلیل تفاوت ها در سیستم های اندازه گیری، ثبت و اطلاعات و دیدگاه های بالینی، کیفیت داده های پرونده های سلامت الکترونیک بسیار متغیر است. از این روی، ارایه فرضیاتی در مورد یک دیتابیس مبتنی بر EHR بر اساس فرضیات دیگر غیر منطقی است. ما نیازمند روش های سیستماتیک برای ارزیابی کیفیت مجموعه داده های پرونده های سلامت الکترونیک برای یک کار تحقیقاتی معین می باشیم.

مقاله مروری ما از مقالات دیگر از نظر هدف تفاوت دارد. مقالات مروری قبلی به بررسی یافته های مربوط به کیفیت پرداخته اند، در حالی که این مقاله به بررسی روش هایی میپردازد که برای ارزیابی کیفیت داده ها استفاده شده اند. در حقیقت، مقالات قبلی صریحا محدود به مطالعاتی بوده اند که این مطالعات به استفاده از یک معیار مرجع متکی بوده است، در حالی که ما در عوض به بررسی طیف وسیعی از روش های ارزیابی کیفیت داده ها می پردازیم. این مرور منابع یک مدل مفهومی تجربی را از ابعاد کیفیت داده های پرونده سلامت الکترونیک توسط محققان بالینی و یک خلاصه و نقدی از روش های مورد استفاده برای ارزیابی کیفیت داده های پرونده سلامت الکترونیک به خصوص در زمینه استفاده مجدد از داده های بالینی برای تحقیقات، ارایه می کند. هدف ما توسعه یک دانش سیستمی از رویکرد هایی است که برای تعیین مناسبیت داده های پرونده الکترونیکی سیستم برای اهداف تحقیقاتی خاص استفاده می شود.

روش ها

ما با انجام یک جست و جو با استفاده از ابزار های کتاب شناسی الکترونیک استاندارد، مقالاتی را با انجام مرور منابع شناسایی کردیم. جست و جوی منابع توسط نویسنده اول در PubMed در فوریه 2012 انجام شد. بر طبق گفته هوگان و وگنر در مرور منابعی که انجام داده اند، هیچ اصطلاح زیر عنوان پزشکی (MeSH) در خصوص کیفیت داده ها وجود ندارد، از این روی یک مرور اکتشافی مختصر برای شناسایی کلمات کلیدی مربوطه انجام شد. فهرست نهایی شامل کیفیت داده، 'دقت و صحت داده'، 'پایایی داده ها'، 'روایی داده'، 'پیوستگی داده ها'، 'کامل بودن داده' و 'خطای داده' بودند. عنوان MeSH برای پرونده سلامت الکترونیک تا 2010 معرفی نشد، از این روی عنوان کلی تر و قدیمی تر MeSH "سیستم های پرونده پزشکی و کامپیوتری" در عوض استفاده شد. عبارت های "EHR"، پرونده پزشکی الکترونیک" و پرونده پزشکی کامپیوتری" نیز برای پوشش دادن مقالاتی که به طور صحیح علامت گذاری نشده بودند، در نظر گرفته شدند. ما مقالاتی را حداقل با یکی از اصطلاحات مربوط به کیفیت و یکی از اصطلاحات مربوط به EHR جست و جو کردیم. نتایج محدود به مقالات زبان انگلیسی بودند. جست و جوی کامل در زیر نشان داده شده است.

(کیفیت داده ها یا صحت داده ها یا پایایی داده ها یا روایی داده ها یا پیوستگی داده ها یا کامل بودن داده ها یا خطاهای داده ها یا خطای داده ها) و (EHR یا پرونده پزشکی الکترونیک یا پرونده پزشکی کامپیوتری یا سیستم های پرونده پزشکی کامپیوتری (mh) و انگلیسی (lang)).

این جست و جو منجر به پیدا شدن 230 مقاله شد که همه آن ها به طور دستی توسط نویسنده اول برای تعیین این که آن ها دارای معیار های انتخاب می باشند یا خیر، مرور و بازنگری شدند. به طور ویژه، مقالات حفظ شده برای مرور بیشتر: 1- شامل تحقیقات اولیه استفاده کننده از روش های ارزیابی کیفیت داده ها بودند 2- بر داده های حاصل از پرونده سلامت الکترونیک یا سیستم های مرتبط متمرکز بودند و 3- در یک مجله داوری همتا منتشر شده بودند. مقالات مربوط به داده های سیستم های مدیریتی (برای مثال دیتابیس های کلایم) در نظر گرفته نشدند. این معیار های ورود به مطالعه منجر به 44 مقاله شد. سپس، ما یک جست و جوی عمیق را انجام داده و منابع همه مقالات را در یک منبع اولیه جست و جو کردیم. به این ترتیب امکان شناسایی 51 مقاله دیگر نیز وجود داشت که جمعا 95 مقاله منطبق بر معیار های ورود ما بودند و سپس برای نتیجه گیری در این مطالعه استفاده شدند.

از هر مقاله، ویژگی های مربوط به کیفیت داده های بررسی شده، روش های ارزیابی مورد استفاده و اطلاعات توصیفی اولیه در مورد مقاله و نوع داده های مورد مطالعه، استخراج و خلاصه شدند. از طریق مرور مکرر داده های خلاصه شده، ابعاد گسترده کیفیت داده ها و مقوله های کلی راهبرد های ارزیابی توصیف شده در منابع استخراج شد. در نهایت، 95 مقاله مجدداً بازنگری شده و هر مقاله بر اساس ابعاد و نیز راهبرد های ارزیابی مورد استفاده برای هر یک از سه بعد مورد ارزیابی شد.

قبل از شروع این تحلیل، ما به جست و جوی مدل های قبلی کیفیت داده های EHR پرداختیم با این حال هیچ مدلی پیدا نشد. ما به این نتیجه رسیدیم که مزایای بالقوه اقتباس یک مدل کیفیت داده از زمینه دیگر با ریسک های تحلیلی همراه است. به علاوه، استفاده از یک مدل موجود برای انجام تحلیل، یک رویکرد قیاسی است که موجب مبهم تر شدن اطلاعات موجود در داده ها می شود. با تحمیل مدل موجود از یک رشته متفاوت، ما به بررسی خطر مفقود شدن یافته های مهم نیز پرداختیم. این رویکرد نسبت به رویکرد قیاسی با امکان پوشش بهتر ابعاد و روش های ارزیابی کیفیت داده ها، مزیت دارد.

نتایج

اکثریت مقاله های مرور شده (73 درصد) تنها به بررسی داده های ساختاری و یا ترکیبی از داده های ساختاری و بدون ساختار (22 درصد) پرداختند. بر این اساس، انواع داده های بدون ساختار شامل متون آزاد بودند در حالی که انواع داده های دارای ساختار شامل داده های کد گذاری شده، مقادیر مربوط به فهرست های قبلی و یا داده های وارد شده به زمینه های (فیلد) نیازمند فرمت های الفبایی ویژه، بودند.

با نادیده گرفتن تغییرات ناشی از مقوله های لغوی و نفی، مقالات دارای 27 عبارت منحصر به فرد بودند که برای توصیف ابعاد کیفیت داده ها استفاده شدند. ویژگی های مربوط به کیفیت داده ها که توصیف شده بودند ولی ارزیابی نشده بودند، در تحلیل لحاظ نشده بودند. ما اصطلاحات را بر اساس تعاریف مشترک گروه بندی کردیم. تعداد کمی از ویژگی های داده های خوب توصیف شده در منابع از جمله دقت کافی و استفاده از استاندارد ها، در تحلیل ما در نظر گرفته نشده اند. این تصمیم به دلیل بحث محدود این ویژگی ها گرفته شد، حقیقت این است که آن ها به عنوان ویژگی های داده های خوب به جای کیفیت داده در نظر گرفته می شوند و این که

هیچ یک از روش های ارزیابی توصیف نشده اند. به طور کلی، ما به طور تجربی، پنج بعد متفاوت از کیفیت داده ها را از منابع استخراج کردیم. این ابعاد در زیر تعریف شده اند:

- کامل بودن: آیا این در مورد یک بیمار موجود در پرونده سلامت الکترونیک صدق می کند؟
- صحت: آیا این مولفه موجود در EHR صحیح است یا خیر؟
- انطباق: آیا یک هم خوانی بین مولفه ها در EHR و یا بین EHR و دیگر منابع داده وجود دارد؟
- توجیه پذیری: آیا یک مولفه در پرونده سلامت الکترونیک از حیث سایر اطلاعات و دانش در مورد این که این مولفه چه چیزی را اندازه گیری می کند، معنی می دهد؟
- رواج: آیا یک مولفه در پرونده سلامت الکترونیک یک معرف مناسب از وضعیت بیمار در یک نقطه زمانی معین است یا خیر.

فهرست عبارات کیفیت داده ها و طبقه بندی آن ها به پنج بعد فوق الذکر، در جدول 1 نشان داده شده است. عبارات انتخاب شده که نشان دهنده هر بعد می باشند، شفاف ترین و کم ابهام ترین عبارات بودند. حجم زیادی از تغییرات و هم پوشانی در عبارات مورد استفاده برای توصیف هر یک از این ابعاد وجود دارد. برای مثال، صحت گاهی اوقات به صورت مترادف با درستی و دقت به کار می رود، با این حال در دیگر مقالات هم به معنی صحت و هم به معنی کامل بودن است. ابعاد به خودی خود بر اساس تعاریف، خلاصه سازی شدند. هر مقاله شناسایی شده می تواند منطبق با یک یا چند بعد باشد.

جدول 1: عبارات و اصطلاحات مورد استفاده در منابع برای توصیف ابعاد رایج و متداول کیفیت داده ها

رواج	توجیه پذیری (معقول بودن)	انطباق	دقت	تمامیت
تازگی به موقع بودن	صحت قابل باور بودن قابل اعتماد بودن روایی	همخوانی پیوستگی پایایی تغییر	صحت تصحیحات انجام شده خطاها گمراه کننده مثبت	قابلیت دسترسی صحت فراهمی مفقود شدن حذف حضور

			ارزش پیش بین کیفیت روایی	کیفیت نرخ ثبت حساسیت روایی
--	--	--	--------------------------------	-------------------------------------

- یک فرایند مشابه برای شناسایی رایج ترین روش های ارزیابی کیفیت داده ها استفاده شد. راهبرد های مورد استفاده برای ارزیابی ابعاد کیفیت داده ها به هفت دسته از مقوله از روش ها طبقه بندی شد که بسیاری از آن ها برای ارزیابی چند بعد استفاده شده است. این روش های عمومی به صورت زیر فهرست بندی و تعریف شده اند:
- استاندارد طلایی: یک مجموعه داده بر گرفته از منبع دیگر یا چند منبع، با و بدون اطلاعات از پرونده سلامت الکترونیک، به عنوان استاندارد طلایی استفاده می شود.
 - هم خوانی عناصر داده ها: دو یا چند عنصر درون EHR مقایسه می شوند تا مشخص شود که آیا آن ها اطلاعات یکسان یا سازگار گزارش می کنند یا خیر.
 - حضور عنصر: حضور نشان دهنده این که آیا عناصر داده های مطلوب یا مورد انتظار حضور دارند یا خیر.
 - هم خوانی منبع داده ها: داده های مربوط به پرونده سلامت الکترونیک با داده های منبع دیگر برای تعیین هم خوانی آن ها مقایسه می شوند.
 - مقایسه توزیع: توزیعات یا آماره های خلاصه داده های ترکیبی از EHR، با توزیعات مورد انتظار برای مفاهیم بالینی مورد نظر مقایسه می شوند.
 - کنترل روایی: داده های پرونده سلامت الکترونیک با استفاده از روش های مختلف برای تعیین معنی داری مقادیر، ارزیابی می شوند.
 - کنترل ورود: اطلاعات مربوط به عملیات ثبت داده های واقعی (برای مثال، تاریخ، زمان و ویرایش) بررسی می شود.
- خلاصه روش های مورد استفاده برای ارزیابی نوع ابعاد در جدول 2 نشان داده شده است. نمودار در شکل 1 قدرت روابط بین ابعاد و روش ها را نشان می دهد. برخی از روش ها برای تنها ارزیابی ابعاد خاص کیفیت داده استفاده شدند، در حالی که روش های دیگر به طور گسترده استفاده شدند. برای مثال، حضور عنصر برای ارزیابی کامل

بودن و نه ابعاد دیگر استفاده شد. با این حال، همخوانی عنصر داده ها و منبع داده ها به طور گسترده استفاده شد. بیشتر ابعاد با استفاده از چند روش استفاده شد با این حال، رواج و به موقع بودن تنها با استفاده از یک روش اندازه گیری می شد.

تمامیت

تمامیت رایج ترین بعد ارزیابی کیفیت داده بوده و در 61مقاله(64 درصد) مقالات یک زمینه مورد مطالعه محسوب می شد. به طور کلی، تمامیت اشاره به این دارد که آیا واقعیت بیمار در پرونده سلامت الکترونیک وجود دارد یا خیر. بیشتر مقالات از اصطلاح تمامیت برای توصیف این بعد استفاده کردند با این حال برخی اشاره به دسترسی داده ها یا مفقود شدن داده دارند. در سایرین، تمامیت به چندین مفهوم کلی تر نظیر صحت یا کیفیت تقسیم شد. برخی از مقالات، به تعریف آماری تمامیت پیشنهاد شده توسط هوگان و وگنر اشاره کرده اند که در آن ها، تمامیت با حساسیت معادل است.

بسیاری از مقالات، تمامیت داده های پرونده سلامت الکترونیک با استفاده از منبع دیگر داده ها به عنوان استاندارد طلایی را ارزیابی کردند. استاندارد های طلایی مورد استفاده شامل حفظ هم زمان اطلاعات مقالات(12-18)، اطلاعات عرضه شده توسط بیمار(19-21)، مرور داده ها توسط بیماران(22-25) و مواجهه بالینی بیمار با داده ها(26-28) و اطلاعات ارایه شده توسط بیماران استاندارد، اطلاعات درخواستی از پزشک(31) و منابع داده های جایگزین که از آن ها عناصر پرونده سلامت الکترونیک گرفته شده بود، بود. یک رویکرد مشابه شامل مثلث بندی داده ها از منابع مختلف در پرونده سلامت الکترونیک برای ایجاد یک استاندارد طلایی(34-35) بودند.

سایر محققان به بررسی حضور یا غیاب عناصر در پرونده سلامت الکترونیک پرداختند. در برخی از موارد، این ها عناصری بودند که انتظار میرفت حضور داشته باشند حتی اگر آن ها برای یک کار خاص مورد نیاز نبودند(36-38، 57-58). در شرایط دیگر، عناصر بررسی شده به کار های موجود بستگی داشتند این بدین معنی است که محققان این موضوع را تعیین کرده اند که آیا داده های پرونده سلامت الکترونیک برای اهداف خاص به اندازه کافی کامل هستند یا خیر(39-49-59-82-86). سایر روش های ارزیابی تمامیت شامل بررسی هم خوانی بین عناصر مربوط به یک منبع، هم خوانی بین پرونده سلامت الکترونیک و اطلاعات مقالات(79-87-88)، هم خوانی

بین پرونده سلامت الکترونیک و منبع الکترونیک داده‌ها (78-79) و مقایسه توزیع وقوع عناصر خاص بین شیوه‌ها (97) و یا با نرخ‌های ثبت شده ملی (98-100) می‌باشد.

استاندارد طلایی	
تمامیت	هم خوانی عناصر داده‌ها
صحت	حضور عنصر
تطابق	هم خوانی منابع داده‌ها
توجیه پذیری	مقایسه توزیع
رواج	کنترل روایی
	ثبت ورود

شکل 1: نقشه بین روش‌های ارزیابی کیفیت داده‌ها و ابعاد کیفیت داده‌ها. ابعاد سمت چپ و روش‌های ارزیابی در سمت راست هر دو به ترتیب نزولی فراوانی از بالا به پایین نشان داده شده‌اند. وزن لبه متصل به بعد و روش نشان دهنده فراوانی نسبی آن ترکیب است.

جدول 2: ابعاد کیفیت داده‌ها و روش‌های ارزیابی کیفیت داده‌ها

کل	رواج	توجیه پذیری	انطباق	صحت	تمامیت	بعد
37	36 46 52 84 4	38 73	51 59 68-72	12-23 25 26 28-48 49 50 52-54 56-67	12-35 49-55 56	روش
26		50 101-103	91-96	90	36-39 49 57-59 68 69 74-86	استاندارد
23		73 89	31 57 97	32 62 104-106	79 87-89 97-100	طلایی
11		7	7	73	61	هم خوانی
10				57		عناصر داده
7						ها
5						حضور عنصر

						هم خوانی منبع داده ها مقایسه توزیع کنترل روایی ثبت ورود کل
--	--	--	--	--	--	--

صحت

دومین بعد رایج کیفیت داده ها شامل صحت بود که در 57 (60 درصد) مقاله قرار داشت. داده های پرونده سلامت الکترونیک زمانی صحیح در نظر گرفته شد که اطلاعات موجود صحیح بودند. سایر عباراتی که برای توصیف این مفهوم استفاده شدند شامل صحت، خطا و کیفیت می باشند. معمولاً، صحت شامل تمامیت می باشد و این بدین دلیل است که برخی از محققان داده های مفقود را به صورت عدم صحت در نظر می گیرند (خطای حذف). تعریف صحت توسط هوگان و وگنر بیان شده است که صحت داده ها نسبتی از عناصر داده های صحیح است که معادل با مقادیر پیش بین مثبت است. مقایسه داده های پرونده سلامت الکترونیک با یک استاندارد طلایی رایج ترین روش برای ارزیابی صحت بود. این استاندارد طلایی شامل موارد زیر هستند: اطلاعات مقالات (12-18، 38، 40، 41)، اطلاعات ارایه شده توسط بیماران از طریق مصاحبه (19-36، 20-42)، پرسش نامه (21-43)، مرور داده ها (22-23-25-44) و ثبت مستقیم داده ها، مواجهه بالینی با بیماران (26، 28، 45)، اطلاعات ارایه شده توسط بیماران (29-30)، داده های ثبت شده خودکار، تماس با پزشک معالج (31-39-47) و منابع جایگزین داده ها که از آن عناصر EHR اطبیب اطلاعات، خلاصه شدند (32-33). برخی از محققان، استاندارد های طلایی را با استخراج و مثلث بندی داده ها از پرونده سلامت الکترونیک (34-35-43-48)، دومین و رایج ترین رویکرد برای ارزیابی صحت، بررسی هم خوانی بین عناصر درون پرونده سلامت الکترونیک بود. معمولاً، این

شامل تایید یک تشخیص با بررسی عمل‌ها، دارو‌ها و یا مقادیر آزمایشگاهی است (49-50-52-54، 58، 60-61) بود.

به طور مشابه، برخی از مقالات، اطلاعاتی را در مورد همخوانی بین عناصر مربوطه (56-62) و خطاهای شناسایی شده از طریق بررسی استفاده از عملیات کپی و پیست (63-64) ارائه کردند. سایر محققان به طور ویژه به بررسی هم‌خوانی بین عناصر ساختاری و داده‌های بدون ساختار درون پرونده سلامت الکترونیک (56-65) پرداخته‌اند. یکی از رویکردهای رسمی تر توصیف شده برای ارزیابی صحت، بستر کیفیت داده‌ها می‌باشد که توسط براون و وارمینگتون (66-67) ارائه شده است و این کوئری ای است که وقتی در برابر دیتابیس EHR اجرا شود، تنها موارد با برخی ناهم‌خوانی‌های بین عناصر داده‌ها را باز می‌گرداند.

تعداد کمی از مقالات به توصیف استفاده از کنترل روایی برای ارزیابی صحت پرداختند. این موارد شامل مرور تغییرات داده‌های دنباله‌ای با گذشت زمان (105)، شناسایی مقادیر فشار خون (104-106) و مقایسه عناصر با بازه مورد انتظار آن‌ها (32-62) بودند. دو رویکرد دیگر شامل استفاده از اصلاحات در فایل‌های ورودی به عنوان پروکسی برای صحت و مقایسه داده‌ها بر روی بیماران از یک رجیستری و پرونده سلامت الکترونیک بودند (90).

انطباق:

شانزده (17 درصد) مقالات بازنگری شده، به ارزیابی انطباق پرداخته بودند. داده‌ها زمانی منطبق بر هم بودند که یک هم‌خوانی و سازگاری بین عناصر داده‌ها وجود داشت. این بدین کهمی است که دو عنصر ثبت اطلاعات یکسان برای یک بیمار دارای مقدار یکسانی است و یا این که عناصر ثبت اطلاعات مختلف دارای مقادیری هستند که وقتی به صورت ترکیبی در نظر گرفته شوند معنی می‌دهند (برای مثال، جنسیت بیولوژیکی به صورت ماده در نظر گرفته می‌شود و روش به صورت معاینه زنان ثبت می‌شود). اندازه‌گیری تطابق معمولاً بر اساس عناصر موجود در پرونده سلامت الکترونیک می‌باشد با این حال برخی محققان، اطلاعات سایر منابع داده‌ها را در نظر می‌گیرند. اصطلاحات رایج مورد استفاده در منابع برای توصیف تطابق داده‌ها شامل هم‌خوانی و پیوستگی می‌باشند.

رایج ترین رویکرد برای ارزیابی تطابق، بررسی هم خوانی بین عناصر درون پرونده سلامت الکترونیک (59-68-70) به خصوص تشخیص و اطلاعات مرتبط نظیر دارو و روش های عمل (51-71-72)س می باشد. دومین روش رایج برای ارزیابی تطابق، بررسی هم خوانی داده های پرونده سلامت الکترونیک با داده های منابع دیگر بود. این منابع سایر شامل اطلاعات صدور صورت حساب (91)، داده های گزارش شده توسط بیمار و داده های گزارش شده توسط پزشک بود (96). رویکرد دیگر مقایسه توزیع داده ها درون پرونده های سلامت الکترونیک با توزیع اطلاعات یکسان از عملیات پزشکی یکسان (57-97) و یا با نرخ های ملی مشابه (31) بود.

توجیه پذیری (معقول بودن)

هفت (7درصد) مقالات، به ارزیابی معقول بودن داده های پرونده سلامت الکترونیک پرداخته بودند. در این زمینه ها داده ها در صورتی معقول بودند که با دانش پزشکی عمومی اطلاعات هم خوانی داشتند و از این روی عملی بودند. به عبارت دیگر، ارزیابی توجیه پذیری برای تعیین این که آیا داده ها قابل اعتماد هستند و یا دارای کیفیت نامطلوب می باشند، استفاده می شود. عبارات دیگر برای بحث و توصیف توجیه پذیری داده های پرونده سلامت الکترونیک شامل صحت و روایی داده ها استفاده شد.

رایج ترین رویکرد برای ارزیابی توجیه پذیری داده های پرونده سلامت الکترونیک، انجام آزمون روایی برای تعیین این است که آیا عناصر خاص درون پرونده سلامت الکترونیک، صحت دارند یا خیر. این شامل جست و جوی عناصری با مقادیری است که این مقادیر در خارج از بازه های بیولوژیکی قرار دارند و یا این که با گذشت زمان به طور غیر منطقی تغییر می کنند و یا دارای عناصر صفر هستند (73). محققان دیگر اقدام به مقایسه توزیعات مقادیر داده ها بین شیوه های (50-101) یا نرخ های ملی (102-103) پرداخته و هم خوانی بین عناصر مربوطه (38-73) را بررسی کردند.

رواج (به موقع بودن)

رواج داده های پرونده سلامت الکترونیک در چهار (4درصد) 95 مقاله ارزیابی شدند. رواج در منابع اشاره به به موقع بودن یا تازگی دارد. داده ها در صورتی به صورت داده های تازه در نظر گرفته می شدند که در پرونده در یک دوره زمانی معقول پس از اندازه گیری ثبت می شدند و یا این که معرف وضعیت بیمار در یک دوره زمانی مطلوب بودند. در همه چهار مقاله، رواج از طریق بازنگری ثبت داده ها ارزیابی شد. در سه مورد از چهار مورد،

محققان به مرور این پرداختند که آیا داده های مطلوب وارد پرونده در چارچوب زمانی مورد نظر وارد شده اند یا خیر(36، 46 و 52). در چهارمین مورد، محققان این موضوع را در نظر گرفتند که آیا هر نوع از داده ها به اندازه کافی از نظر پزشکی مرتبط و مناسب هستند یا خیر.

بحث

ما اقدام به شناسایی پنج بعد کیفیت داده ها و هفت مقوله از روش های ارزیابی کیفیت داده ها کرده ایم. بررسی انواع روش های مورد استفاده علاوه بر هم پوشانی روش های بین ابعاد، الگوهای معنی داری را نشان می دهد. در زیر ما به بررسی یافته های اصلی مرور منابع می پردازیم که به طور ویژه نشان دهنده زمینه هایی است که نیازمند توجه بیشتر بوده و پیشنهاداتی را برای تحقیقات آینده ارائه می کنند.

اصطلاحات و ابعاد کیفیت داده ها

یکی از بزرگ ترین مشکلات در انجام این مقاله، مربوط به اصطلاحات متناقض مورد استفاده برای بحث در مورد کیفیت داده ها بوده است. برای مثال، ما انتظار هم پوشانی عبارات را بین ابعاد و این که زبان درون یک مقاله گاهی نامتناقض است نداشتیم. انجمن تحقیقات بالینی قادر به توسعه یک طبقه بندی نامتناقض از کیفیت داده ها نبوده است.

با این حال، یک همپوشانی بین ابعاد کیفیت داده های شناسایی شده در این مقاله و سایر مقالات توصیف شده در طبقه بندی های قبلی و مدل های کیفیت داده ها وجود دارد. چارچوب مفهومی وانگ و استرانگ در خصوص کیفیت داده ها دارای 15 بعد است که به چهار مقوله طبقه بندی می شود: درونی، زمینه ای، بازنمودی و قابلیت دسترسی. مقاله مروری ما بر مسائل کیفیت داده های ذاتی و زمینه ای (وابسته به وظیفه) تاکید دارد. ابعاد شناسایی شده با دو ویژگی ذاتی (صحت و باورپذیری که معادل دقت و توجیه پذیری هستند) و دو ویژگی زمینه ای (به موقع بودن که معادل رواج است و کامل بودن) هم پوشانی داشتند. تنها بعد شناسایی شده که در چارچوب وانگ و استرانگ حضور نداشت، انطباق بود.

موسسه پزشکی، چهار ویژگی کیفیت داده متناسب با پرونده های بیماران را شناسایی کرده است: تمامیت، صحت، خوانایی و معنی (مربوط به قابلیت درک)(108). همان طور که این موسسه گزارش کرده است، پرونده های الکترونیک بسیاری از نگرانی های مربوط به خوانایی را رد می کنند، به این ترتیب سه ویژگی باقی می ماند که

دو مورد از آن‌ها در این مقاله بررسی شدند. معنی یک مفهوم انتزاعی تر بوده و اندازه گیری عینی آن سخت است و به همین دلیل، ما ارزیابی‌های این بعد را در منابع مشاهده نکرده ایم. اگرچه پنج بعد کیفیت داده‌های استخراج شده در طی این مقاله، صورت دو به دو ناسازگار در منابع در نظر گرفته شده اند، با این حال ما حس می‌کنیم که تنها سه مورد بسیار اساسی هستند: صحت، تمامیت و به موقع بودن. منظور این است که این ابعاد کاهش ناپذیر می‌باشند و مفاهیم اصلی مربوط به کیفیت داده‌ها در خصوص استفاده مجدد از داده‌های پرونده سلامت الکترونیک را نشان می‌دهند. از سوی دیگر، انطباق و توجیه پذیری، اگرچه به صورت ویژگی‌های مجزای کیفیت داده‌ها بحث شدند، بدیهی است که به عنوان شاخص‌هایی برای ابعاد اساسی استفاده می‌شوند که ارزیابی مستقیم آن‌ها غیر ممکن است. این فرض توسط هم پوشانی مشاهده شده در روش‌های مورد استفاده برای ارزیابی انطباق و توجیه پذیری با روش‌های مورد استفاده برای ارزیابی صحت و تمامیت تایید می‌شود. عدم انطباق بین دو منبع داده‌ها حاکی از وجود خطا در یکی از دو منبع است: یک خطای حذف، منجر به از بین رفتن تمامیت شده و یا خطای کمیسیون منجر به از بین رفتن صحت می‌شود. به طور مشابه، داده‌های غیر موجه ممکن است همانند اندازه گیری‌های ناقص همانند نرخ تشخیص ترکیبی درون شیوه‌هایی که منطبق با نرخ جمعیت مورد انتظار نمی‌باشند، غلط باشند. از این روی صحت، تمامیت و انطباق، رویکرد‌های روش شناسی برای ارزیابی کیفیت داده‌ها است. به علاوه، محققان، وقتی که بر این باورند مسائلی در خصوص صحت و تمامیت وجود دارند، به انطباق و توجیه پذیری اشاره می‌کنند، با این حال شیوه‌ای وجود ندارد که اطمینان حاصل شود که خطا وجود دارد و یا این که عناصر داده‌ها ممکن است غلط باشند.

روش ارزیابی کیفیت داده‌ها

ما تعدادی از الگوهای ضروری را در منابع از حیث انواع ارزیابی‌های کیفیت داده‌های مورد استفاده و شیوه بحث در خصوص ارزیابی‌های کیفیت داده‌ها، مشاهده کرده ایم. برای مثال، 37 مورد از 95 مقاله در این نمونه به استاندارد‌های طلایی برای ارزیابی کیفیت داده‌ها متکی هستند. در رابطه با این روش یک سری مشکلات وجود دارد. اولاً، منابع داده‌های مورد استفاده را می‌توان به عنوان استاندارد‌های طلایی واقعی در نظر گرفت. برای مثال، اطلاعات دفتری گاهی بیشتر از اطلاعات الکترونیکی قابل اعتماد هستند با این حال آن‌ها کامل و صحیح نیستند. از همه مهم‌تر، یک استاندارد طلایی برای داده‌های پرونده سلامت الکترونیک به آسانی در بیشتر

موارد قابل دسترس نیست. وقتی که استفاده از مجموعه داده های بدون هویت برای تحقیقات متداول تر باشد، این می تواند موجب بروز مشکل شود. یک رویکرد "تناسب با هدف" که نشان می دهد کیفیت هر مجموعه داده برای هر کار خاص بایستی ارزیابی شود، مستلزم پذیرش و استفاده از جایگزین هایی برای روش های مبتنی بر استاندارد طلایی شود.

علاوه بر اتکای بیش از حد به استاندارد های طلایی، اکثریت مطالعات شناسای شده به دانش شهودی کیفیت داده ها متکی بوده و از روش های ادهاک برای ارزیابی کیفیت داده ها استفاده می کنند. این روند را می توان در سایر زمینه ها نظیر مشاهده کرد (107-109).

بیشتر مطالعات در این مقاله مروری، یک سری روش های ارزیابی را مطرح کرده اند که با اصول و مبانی نظری و تجربی توسعه یافته اند. برخی محققان سعی کرده اند تا رویکرد های تعمیم پذیری را توسعه دهند که می توانند به عنوان گامی به سمت روش شناسی استاندارد استفاده شود. فالكونر و دیلوزینکان (52) از یک رویکرد آماری چند مرحله ای برای ارزیابی کیفیت داده ها استفاده کردند. هوگان و وگنر (8) شاخص های آماری خاص را در خصوص صحت و تمامیت عناصر EHR ارایه کرده اند که توسط سایر محققان پذیرفته شده اند. روش های خاص، از جمله مقایسه توزیع داده ها از EHR با توزیعات مورد انتظار یا جست و جوی هم خوانی بین عناصر درون پرونده سلامت الکترونیک، برای تعمیم از اهمیت زیادی برخوردار هستند. کاوشگر براون و وارمینگتون (66-67)، را می توان به عناصر داده های مختلف توسعه داد، اگرچه آن ها مستلزم دانش بالینی دقیقی هستند. برخی از محققان به بررسی کیفیت دیتابیس های تحقیقاتی از شیوه های عمومی در بریتانیا پرداخته اند (31-98-99-102). در بسیاری از موارد، با این حال، روش های ارزیابی ویژه توصیف شده در منابع، به سختی قابل تعمیم به سایر دیتاست ها یا سوالات تحقیق می باشند. در صورتی که استفاده مجدد از داده های EHR برای تحقیقات بالینی، عملی باشد، توسعه روش های ارزیابی کیفیت داده های سیستماتیک و استاندارد از اهمیت زیادی برخوردار است.

به علاوه، در صورتی که هدف ما پذیرش مفهوم تناسب برای هدف باشد، در نظر گرفتن استفاده از داده های پرونده های سلامت الکترونیک برای زمینه های تحقیقاتی (در زمان تعیین کیفیت آن ها) مهم است (110). برخی از ابعاد وابسته به کار و یا ذهنی هستند و برخی دیگر مستقل از کار و عینی هستند (109). توسعه دانش و درک

کامل روابط درونی وظایف تحقیقاتی و ویژگی های داده ها از اهمیت زیادی برخوردار است زیرا آن ها به کیفیت داده ها مربوط می شوند. برای مثال، تمامیت یک مجموعه از عناصر داده های مورد نیاز توسط یک پروتکل تحقیقاتی، ممکن است از تمامیت مورد نیاز برای یک پروتکل متفاوت، متغیر باشد. عوامل بسیاری از جمله هدف بالینی، نیامند مناسب مجموعه داده برای وظایف تحقیقاتی خاص است(26).

مسیر های آینده

ما بر این باوریم که تلاش برای استفاده مجدد از داده های پرونده سلامت الکترونیک برای تحقیقات بالینی، از کار های انجام شده در برخی زمینه های خاص گرفته شده اند:: استفاده از طبقه بندی کیفیت داده های پرونده الکترونیک سلامت، افزایش آگاهی از وابستگی به وظیفه، یکپارچه سازی کار ها بر روی ارزیابی کیفیت داده ها از زمینه های دیگر و استفاده از روش های سیستمی و آماری ارزیابی کیفیت داده ها. طبقه بندی کیفیت داده ها، از روش های ارزیابی و گفتمان های ساختاری پشتیبانی می کند. یافته های مربوط به این مقاله در خصوص ابعاد کیفیت داده ها می تواند گامی در راستای این هدف باشد. وابستگی به وظیفه، یک مسئله رو به رشد با افزایش تلاش برای استفاده مجدد از داده های EHR می باشد زیرا ارزیابی کیفیت داده ها فاقد یک راه حل جامع است. یک رویکرد برای حل مسئله کیفیت و مناسبیت داده های پرونده سلامت الکترونیک برای استفاده مجدد، بررسی مطالعاتی است که در خارج از حیطه تحقیقات بالینی انجام شده اند، زیرا کیفیت داده ها یک زمینه مطالعاتی در زمینه های متغیر از امور مالی تا مهندسی صنعتی است. در نهایت، مهم است که محققان بالینی از رویکرد های موقت به سمت ارزیابی کیفیت داده ها حرکت کنند. روش های معتبر که می توان از آن ها برای پاسخ به سوالات تحقیق مختلف استفاده کرد، یک هدف ایده ال محسوب می شود.

محدودیت ها

این مقاله با محدودیت هایی همراه بوده است. اولاً، جست و جو محدود بود. به دلیل نبود اصطلاحات MeSH برای کیفیت داده ها و تغییرات در واژگان مورد استفاده برای بحث کیفیت داده ها، امکان دارد که جست جوی اولیه برخی از مقاله های مربوطه را از سرچ نکرده باشد. ما بر این باوریم که تصمیم ما برای بازنگری منابع هر مقاله، موجب بهبود اشباع نمونه شده است. هم چنین لازم به ذکر است که فرایند طبقه بندی ما کاملاً ذهنی بوده و

تنها توسط یکی از محققان انجام شد. محققان اصلی با تفاسیر ما مخالف بودند. ما از یک فرایند تکراری برای برچسب گذاری و طبقه بندی ابعاد کیفیت داده و روش های ارزیابی توصیف شده در مقاله برای توسعه یک طرح کد گذاری مطلوب استفاده کردیم.

در نهایت، احتمال دارد که کمبود منابع در خصوص کیفیت داده ها در استفاده مجدد از داده های EHR برای تحقیقات بالینی ناشی از گزارش کم تر از مقدار واقعی باشد. اولین روش رایج در تحلیل هر مجموعه داده، بررسی توزیعات، آماره ها و نمودار ها میباشد ولی این فرایند، به ندرت در مقالات مختلف منتشر می شوند. شفافیت بیشتر در خصوص پاک سازی داده ها و کنترل داده ها، قابل توصیه می باشد زیرا به تثبیت استاندارد های گزارش دهی قابل قبول برای استفاده مجدد از داده های پرونده سلامت الکترونیک در تحقیقات، کمک می کند.

نتیجه گیری

استفاده مجدد از داده های پرونده سلامت الکترونیک، یک زمینه پژوهشی در حال توسعه است. با این حال، مسائل مربوط به کیفیت داده های پرونده سلامت الکترونیک، بر لزوم استفاده از روش های ارزیابی کیفیت برای تعیین مناسبت این داده ها برای کار های تحقیقاتی معین تاکید می کنند. در مرور منابع، ارایه شده در این کتاب، ما ابعاد اصلی کیفیت داده ها که برای محققان جالب هستند و روش های ارزیابی عمومی را شناسایی کردیم. کیفیت داده ها یک مسئله ساده نیست و اگر استفاده مجدد از داده های EHR یک روش قابل قبول برای تحقیقات پزشکی باشد، انجمن تحقیقات بالینی بایستی روش های معتبر و سیستمی ارزیابی کیفیت داده های پرونده سلامت الکترونیک را توسعه دهند. ما محققان را تشویق می کنیم تا به بحث در مورد ابعاد کیفیت داده ها پرداخته و از رویکرد های سیستمی در اندازه گیری کیفیت داده استفاده کرده و بهترین شیوه های ارزیابی کیفیت داده های سلامت الکترونیک را در زمینه استفاده مجدد برای تحقیقات بالینی، توسعه داده و منتشر کنند.



این مقاله، از سری مقالات ترجمه شده رایگان سایت ترجمه فا میباشد که با فرمت PDF در اختیار شما عزیزان قرار گرفته است. در صورت تمایل میتوانید با کلیک بر روی دکمه های زیر از سایر مقالات نیز استفاده نمایید:

لیست مقالات ترجمه شده ✓

لیست مقالات ترجمه شده رایگان ✓

لیست جدیدترین مقالات انگلیسی ISI ✓

سایت ترجمه فا ؛ مرجع جدیدترین مقالات ترجمه شده از نشریات معتبر خارجی