

CRITÉRIOS DE QUALIDADE DE DADOS EM SAÚDE: UMA ANÁLISE QUANTITATIVA

HEALTH DATA QUALITY CRITERIA: A QUANTITATIVE ANALYSIS

Fabício Amadeu Galdania^a
Fábio Henrique Alves^b
Késsia Rita da Costa Marchi^c
Leonardo Castro Botega^d

RESUMO

Objetivo: Este artigo propõe uma reflexão sobre como os critérios de qualidade de dados vêm sendo abordados em trabalhos que discutem dados provenientes da área da saúde, possibilitando reconhecer o panorama sobre estes critérios e identificar as lacunas que requerem maiores esforços. **Metodologia:** Revisão sistemática de literatura que buscou identificar, analisar e quantificar os critérios de qualidade de dados que são abordados na área da saúde. **Resultados:** Como resultados apresentam-se os critérios de qualidade de dados mapeados e categorizados, identificando a acurácia, consistência e completude como os critérios mais abordados e atualidade, temporalidade, confidencialidade e plausibilidade, sendo os menos citados. **Conclusões:** Há tanto uma sobrecarga como a falta de uso de determinados critérios, logo, os critérios que apresentaram esta falta geram a possibilidade de serem mais bem abordados e explorados em trabalhos futuros.

Descritores: Qualidade de dados em saúde. Critério de Qualidade de dados. Saúde.

^a Doutorando no Programa de Pós-Graduação em Ciência da Informação pela Universidade Estadual Paulista Júlio de Mesquita Filho (UNESP), campus de Marília.

^b Mestre no Programa de Pós-Graduação em Ciência da Informação pela Universidade Estadual Paulista Júlio de Mesquita Filho (UNESP), campus de Marília. Analista de Sistemas. E-mail: fh.alves@unesp.br.

^c Doutoranda no Programa de Pós-Graduação em Ciência da Informação pela Universidade Estadual Paulista Júlio de Mesquita Filho (UNESP), campus de Marília. Docente do Instituto Federal do Paraná (IFPR) – campus Paranaíba. E-mail: kessia.marchi@ifpr.edu.br.

^d Doutor em Ciência da Computação pela Universidade Federal de São Carlos (UFSCar). Docente do Programa de Pós-Graduação em Ciência da Informação da Universidade Estadual Paulista Júlio de Mesquita Filho (UNESP), campus de Marília. E-mail: leonardo.botega@unesp.br

1 INTRODUO

As unidades de sade atuam em um complexo sistema de processos, apoio tcnico, logstico e administrativo, com o objetivo de prestar a melhor assistncia possvel aos pacientes (MORAES, 2018). Para tanto, os sistemas de informao de sade consistem em ferramentas imprescindveis para alcanar esse fim, devido ao fato de que nestes sistemas podem estar inseridas informaes de sade que dizem respeito aos pacientes, em documentaes como os pronturios. Estas informaes podem ser consideradas crticas, sensveis e que exigem um alto grau de complexidade na sua utilizao, por conterem o histrico do paciente.

Diante desse cenrio, nasce a preocupao com a qualidade de dados e informaes internas, que caso apresentem um baixo desempenho, poder proporcional falhas, incertezas e erros quanto aos procedimentos executados. Portanto, a qualidade dos dados e das informaes preenchidas, tanto nos pronturios do paciente quanto nos sistemas de informao em sade no geral, compem-se de um elemento essencial para todos os protocolos que so executados no contexto da sade.

Alm de rigorosos procedimentos e anlises que podem ser inseridos para avaliar a qualidade de dados no campo da sade, possvel tambm realizar a fuso destes dados como uma atividade complementar para aperfeioar critrios voltados qualidade. Esta fuso de dados se torna necessria pois os dados em sade, em sua grande maioria, advm de fontes heterogneas em mltiplos formatos e em modelos no estruturados.

Ao tratar e combinar estes dados possvel obter concluses mais significativas que permitem reduzir a sua dimensionalidade, proporcionar um maior valor informao e gerar uma melhor produo de subsdios para a construo de conhecimentos que apoiam tomadas de deciso que resultam em poupar tempo, recursos financeiros e acrescentar contribuies para diagnsticos e tratamentos de pacientes. Logo, este artigo tem como objetivo apresentar uma reviso de literatura sobre como os critrios de qualidade de dados vm sendo abordados em trabalhos que discutem dados provenientes da

rea da sade, o que possibilitar reconhec o panorama sobre estes critrios e identificar as lacunas que requerem maiores esforos.

Os procedimentos metodolgicos adotados nesta pesquisa consistem, do ponto de vista da sua natureza em uma pesquisa bsica e ao considerar seus objetivos, trata-se de uma pesquisa exploratria. No que diz respeito aos procedimentos tcnicos, este trabalho compoe-se de uma pesquisa bibliogrfica que incluiu artigos cientficos publicados em peridicos, sendo que as principais reas que dialogam com a abordagem proposta foram as reas de cincia da informao, cincia da computao e cincias da sade.

Logo, o presente artigo pode ser configurado como uma reviso sistemtica de literatura, como considerada por Morandi e Camargo (2015) como "estudos secundrios utilizados para mapear, encontrar, avaliar criticamente, consolidar e agregar os resultados de estudos primrios relevantes, acerca de uma questo ou tpico de pesquisa especfico". Portanto, na reviso sistemtica de literatura realizada a leitura crtica e minuciosa de estudos cientficos, de forma sistemtica, seguindo um mtodo explcito, planejado e justificvel. Fontes-Pereira (2017) complementa dizendo que a reviso sistemtica de literatura permite realizar uma anlise de artigos que respondem a determinados critrios de incluso e excluso, o que possibilita encontrar melhores evidncias que fundamentam concluses.

A busca por artigos que compem esta reviso ocorreu no perodo de agosto de 2020 a abril de 2021 e foram inseridos os termos **Data Fusion** **Data Quality** **Electronic Health Records** em bases de dados como a SCIELO, ACM, ScienceDirect, PubMed, Archivaria e IBICT. Foram inseridos 25 artigos internacionais, publicados nos anos de 2003 at 2021.

Os critrios de incluso dos artigos compem-se de pesquisas que trabalharam diretamente com o desenvolvimento, aplicao, anlise, avaliao ou reviso de literatura a respeito dos processos de qualidade de dados e/ou fuso de dados e informaes direcionadas ao mbito da sade. Ressalta-se que, um exemplo de objeto informacional do qual est integrado nesses processos de qualidade e fuso de dados considerado, consiste no pronturio eletrnico do paciente.

A partir da coleta dos artigos procedeu-se a leitura classificatria, na qual foram construdos resumos que destacaram o objetivo, limitaes e concluses de cada artigo. Foram selecionados aqueles que abordaram, direta ou indiretamente, os critrios de qualidade e fuso de dados em sistemas de informao em sade e pronturios eletrnicos. Esses resumos sso apresentados na seqncia.

2 APRESENTAO DOS ESTUDOS INCLUIDOS

Seguindo uma ordem cronolgica, o primeiro artigo considerado foi de Thiru, K et al. (2003), na qual os autores elaboraram uma reviso sistemtica sobre o escopo e a qualidade dos dados do pronturio eletrnico do paciente na ateno primria. Foram utilizadas pesquisas bibliogrficas de bancos de dados mdicos, bancos de dados especializados em informtica mdica, anais de conferncias e contatos institucionais como fontes de dados para a elaborao da pesquisa, que serviram como base para a extrao de referncias e medidas usadas para avaliar a qualidade dos dados, sendo que 37 artigos mensuram a qualidade dos dados, e 15 possuem a sua temtica voltada para o escopo de qualidade por meio do registro eletrnico do paciente. Com isso, foi possvel concluir, mediante a comparao de taxas, que as medidas de sensibilidade vinculadas qualidade eram altamente dependentes dos elementos de dados vinculados ao pronturio eletrnico do paciente em questo, incluindo dados prescritivos, que eram de melhor qualidade quando comparados aos dados de diagnstico ou estilo de vida.

No trabalho desenvolvido por Orfanidis, Bamidis e Eaglestone (2004), foram discutidos problemas relacionados garantia de qualidade de dados nos sistemas de registros eletrnicos da Grcia. A discusso ressaltou tecnologias com interfaces inteligentes, rvores de deciso e tecnologia mvel. A concluso apresentada que a Grcia possui limitaes no que diz respeito a um desenvolvimento rpido de um sistema de registros eletrnicos de sade nacional pela falta de infraestrutura tecnolgica e de uma populao com conhecimentos de informtica. No entanto, os autores reconhecem que a introduo contnua de sistemas de informao administrativos fornece uma

base interessante para uma implementao pragmtica gradual, na qual aspectos desses sistemas podem ser testados e estabelecidos de forma incremental.

Tang et al (2007) compararam as medidas calculadas de qualidade usadas a partir de dados administrativos provenientes de reclamaes com aquelas derivadas de dados clnicos pertencentes a um sistema de registro eletrnico de sade. Ao usar definies padres baseadas em dados administrativos, foi possvel observar que apenas 75% dos diabticos determinados pela reviso manual foram identificados. Em contraste, 97% dos diabticos foram constatados usando informaes codificadas no sistema de registros eletrnicos de sade. A concluso que h necessidade urgente de compreender melhor as implicaes do uso da regra de diagnstico baseado em reclamaes de dois encontros, antes de decidir sobre as medidas padronizadas nacionais nas quais basear os relatrios pblicos e polticas de reembolso.

A pesquisa conduzida por Ndira, Rosenberger e Wetter (2008) envolveu funcionrios do Hospital Distrital Tororo em Uganda. Com a implantao de sistemas de registros eletrnicos de sade, os autores utilizaram um estudo de interveno com mtodos qualitativos e quantitativos que permitiu observar e comparar melhorias de oportunidades, disponibilidade e preciso dos relatrios de sade de rotina, assim como os nveis de satisfao pessoal em relao ao atendimento prestado a determinados pacientes e da equipe multiprofissional de sade. Os resultados apontaram dificuldades em encontrar pacientes suficientes por meio de acompanhamento direto para indagar sobre a preciso das informaes registradas. Entrevistas com funcionrios mostraram que o sistema eletrnico foi muito apreciado e bem-aceito. Conclui-se que, embora a pontualidade e a disponibilidade dos dados tenham melhorado, o aperfeioamento da preciso no pode ser estabelecido. Logo, se faz necessrio conceber abordagens para verificar a preciso. Os autores reconhecem que para o sucesso da implementao de registros eletrnicos de sade, ainda h uma srie de desafios organizacionais, gerenciais e sociais que devem ser enfrentados alm dos aspectos tcnicos.

A anlise apresentada por Miettinen e Korhonen (2008) basea em

casos de problemas de qualidade de dados de sade, nos quais os dados de pacientes com diabetes s3o combinados a partir de diferentes sistemas de informao. Durante o desenvolvimento de vrios pronturios eletrnicos de sade, ficou claro que a integrao dos dados ainda 3 um desafio. Os dados recolhidos em vrios sistemas podem ter falhas de qualidade, como serem incoerentes e contraditrios, assim como os dados pretendidos estarem em falta. Na perspectiva dos autores, a falta de dados caracteriza-se como a falha mais grave, uma vez que os dados ausentes ou contraditrios podem ser vitais para o cuidado do paciente. A coerncia das informaes produzidas por diferentes organizaes de sade torna-se ainda mais crucial na introduo dos sistemas de dados uniformes nacionalmente. Conclui-se que a qualidade dos dados pode ser melhorada por meio de especificaes uniformes e da integrao padronizada de sistemas de informao, mas tambm 3 essencial que os profissionais de sade insiram as informaes com preciso e as utilizem corretamente.

Moulton, Chaczko e Karatovic (2009), exploraram questes relacionadas ao projeto de sistemas em que registros mdicos eletrnicos s3o atualizados com dados ou metadados de sistemas de sensores de monitoramento de sade internos ou corporais. Um prot3tipo foi desenhado, implementado e avaliado de forma a aprofundar o entendimento sobre a coleta e transferncia de dados de sensores para um pronturio eletrnico. Um dos principais questionamentos propostos pelos autores foi at3 que ponto os nveis de confiana nos dados podem ser afetados por fatores como o tipo desses dados, a qualidade, a instalao, manuteno e calibrao dos sensores. Os autores afirmam que questes relacionadas 3 adoo de tcnicas de agregao ou fuso de dados tambm foram consideradas, devido ao fato de que diferentes sensores podem ser adequados para melhorar o grau de confiana na extrao dos dados de qualquer sensor especfico. Conclui-se que ainda h3 muito a ser estudado e aplicado para se compreender qual 3 o funcionamento de prticas relacionadas 3 coleta e transferncia de dados de um sensor para um pronturio eletrnico, da mesma forma que a adoo de padres para os registros eletrnicos de sade ir3o trazer uma maior segurana e garantia no que diz respeito a

qualidade de dados de todos os tipos de sensores.

O relato apresentado por Duarte et al (2010) versa sobre como a qualidade dos dados pode ser avaliada a partir de registros eletrnicos de sade em um ambiente hospitalar. Sendo assim, os autores construram um mundo virtual dinmico de entidades ou agentes complexos, movidos pelo critrio nico da inteligncia para a prestao de cuidados da sade. Ao longo da pesquisa, os autores apresentaram o *Framework AIDA (Agency for the Integration, Diffusion and Archive of Medical Information)* abordando sua estrutura lgica, os mdulos que o compem e os modelos computacionais aplicados. Conclui-se que a avaliao da qualidade dos dados do modelo desenvolvido ainda requer uma comprovao formal.

Pompilio e Ermertice (2011) criaram indicadores sobre a usabilidade do pronturio eletrnico, gerando resultados que influenciaram diretamente na qualidade do atendimento prestado ao paciente. O objetivo principal da pesquisa foi demonstrar o fluxo de uso do pronturio eletrnico e suas aes proativas no processo de busca das informaes vinculadas aos pacientes para melhor atend-los. Os dados utilizados foram extraídos de uma base de logs de um sistema especfico, obtendo informaes referentes sala de espera, prescriao, solicitao de exames e outras funcionalidades vinculadas desde o incio at o encerramento da consulta. Conclui-se que o pronturio eletrnico foi impactante de maneira positiva no atendimento dos pacientes, e contribui com a localizao rpida de informaes no histrico do paciente, solicitao de exames *on-line* e eliminao do papel como objeto de cadastro de informaes, reduzindo assim a burocracia operacional.

Na pesquisa desenvolvida por Dixon, McGowan e Grannis (2011) foi analisada a qualidade dos dados de registros eletrnicos de sade, dando nfase no atributo de integridade dos dados no contexto de relatrios eletrnicos de laboratrio de informaes de doenas notificveis. Os autores examinaram a integridade dos dados de sistemas de informao clnica e compararam os dados brutos enviados diretamente destes sistemas com dados aprimorados por um intercmbio de informaes de sade antes da transmisso a uma agncia de sade pblica. Tambm foram examinados os impactos dos mtodos de

aprimoramento de intercmbio de informaes que tentam melhorar a integridade. Conclui-se que existem tanto dados de baixa qualidade como faltam elementos de dados necessrios em sistemas de informao clnica, resultando em desafios para os interessados em usos secundrios de dados. Por fim, os autores sugerem a construo de uma infraestrutura de informaes capaz de suportar usos secundrios de dados clnicos eletrnicos e, assim, melhorar os processos de vigilncia em sade.

Kahn et al (2012) entendem que a avaliao da qualidade dos dados uma tarefa particularmente importante em estudos “*multi-site*” para distinguir variaes reais no atendimento de problemas voltados para essa qualidade de dados. Os autores propuseram um modelo conceitual “*fit-for-use*” para avaliao de qualidade de dados e um modelo de processo para planejamento e realizao de avaliaes dessa qualidade em nico e em vrios locais. Conclui-se que para atingir o potencial da revoluo eletrnica na rea de sade essencial possuir uma abordagem baseada em conceitos sistematicamente executveis para a avaliao da qualidade dos dados.

O estudo desenvolvido por Weiskopf, N et al (2013) aborda a potencialidade das dimenses da avaliao de qualidade de dados dos registros eletrnicos de sade e procura avaliar a possibilidade de reutiliz-los para pesquisas clnicas. Na busca por esse objetivo, foi realizada uma reviso de literatura voltada para trabalhos que estudaram essa temtica. Atravs de um processo interativo, os fatores de qualidade de dados foram extraidos e categorizados, bem como os mtodos de avaliao utilizados. A pesquisa resultou em cinco dimenses de qualidade de dados, sendo a completude, exatido, concordncia, plausibilidade e atualidade. Alm das dimenses de qualidade, os autores tambm citam sete amplas categorias de mtodos de avaliao da qualidade dos dados, sendo: Comparao com outros padres, concordncia com elementos e fontes dados, comparao de distribuio, verificaes de validade, reviso de dados registrados e a presena de elementos.

Almutiry, Wills e Crowder (2013) desenvolveram uma estrutura que se ajusta aos atributos de qualidade dos dados. Esta estrutura reflete os principais

elementos dos sistemas de sade, suas funcionalidades e o seu desenvolvimento de vida, resultando na filtragem das dimenses de qualidade de dados existentes em muitas pesquisas e da verificao de sua adequao a natureza dos sistemas de sade. Conclui-se que no mbito da sade, apesar dos registros eletrnicos serem considerados como uma soluo promissora para os problemas no gerenciamento das informaes, ainda h ameaas durante o armazenamento e transmisso dos dados coletados no cotidiano do ambiente hospitalar, assim como o aumento da quantidade de dados de baixa qualidade. Por fim, no h consenso sobre um conjunto rigorosamente definido de dimenses de qualidade de dados, isso levantaria a necessidade de automatizar o mecanismo de medio de qualidade de dados e interoperabilidade semntica.

Com base no entendimento apresentado por Bowman (2013) sobre a implementao de sistemas de registros eletrnicos de sade defeituosos, assim como a sua utilizao inadequada, a autora realizou uma reviso de literatura que examinou o impacto das consequncias no intencionais do uso de sistemas de registros eletrnicos de sade na qualidade do atendimento e buscou compreender quais so as solues propostas para resolver os erros relacionados a esses sistemas. Ela conclui que ao observar os estudos produzidos na ltima dcada, o debate sobre os riscos relacionados a integridade da informao e os impactos reais e potenciais subsequentes na qualidade do atendimento e segurana ao paciente se fazem bastante presentes, no entanto, pouco foi feito para medir e analisar sistematicamente esses riscos, identificar as suas causas, implementar estratgias universais, adotar modificaes no design do sistema e a insero de princpios de usabilidade para reduzir os riscos. Propoe-se como soluo uma combinao de superviso do governo federal e aes da prpria instituio para evitar consequncias no intencionais do uso desses sistemas. Por fim, os prestadores de cuidados de sade devem implementar polticas e procedimentos que abordem o uso e treinamento de registro eletrnico de sade adequados.

Os exames referentes a preciso da aplicabilidade de uma ontologia para a criao de um registro especfico relacionado com a doena diabetes

desenvolvido por Liaw et al (2014) tiveram o intuito de oferecer suporte para um atendimento integrado. Visando esse objetivo, a ontologia desenvolvida foi validada e contextualizada por profissionais da rea da sade, especificamente mdicos, que formalizaram sua idealizao atravs de procedimentos especficos adotados pelos profissionais envolvidos. Conclui-se que as ontologias contriburam significativamente em um processo clnico, sendo utilizada por algoritmos de localizao que resultaram em um ndice de quase 100% de identificao dos registros especficos denominados DM2 (Diabetes Mellitus tipo 2), alm das contribuies no que se refere a preciso, volume e qualidade dos dados tratados no processo.

Baron (2015) entende que so necessrias melhorias nas prticas mdicas voltadas para o preenchimento da documentao referente a prestao de cuidados de pacientes. Para o autor, o maior problema se encontra no fato de que no a mera transio de um sistema analgico para o digital que ir solucionar todos os problemas, ou seja, realizar essa mudana de mdias do suporte papel para o suporte digital no garante, automaticamente, uma melhoria no atendimento, assistncia e formulao do diagnstico voltado para o paciente. Por exemplo, ao usar como base um estudo de caso, a implementao de um sistema de registro eletrnico de uma determinada clnica, foi possvel observar elementos crticos para a plena realizao das necessidades dos profissionais de sade, o que inclui a avaliao mdica de dados estruturados, a necessidade de adoo generalizada de padres e uma reestruturaao da equipe principal com recursos adicionais.

Dziadkowiec et al (2016) analisaram a usabilidade de uma estrutura de qualidade de dados implementada para bancos de dados relacionais com o intuito de limpar os registros contidos em uma base extraída a partir de dois bancos de dados distintos. Ao buscar esse objetivo, dois conjuntos de dados de um departamento de emergncia foram utilizados como exemplos para examinar os conceitos de qualidade de dados com base em uma estrutura pr-projetada para servir como referncia a pesquisa. Os autores utilizaram conceitos de qualidade de dados que envolvem todo o fluxo iniciando-se pela extrao dos dados de uma base especfica, exportao para uma nova funcionalidade, que

fará o fluxo de instanciação de variáveis para contemplar as necessidades dos dados obtidos. Após isso, foi definida uma análise técnica para realizar melhorias nos dados, e, por fim, o desenvolvimento de uma solução para aplicar nos registros extraídos. Obteve-se como parte da conclusão uma sintaxe implementada especificamente sobre os conceitos de qualidade de dados.

Sholte et al (2016) testaram se os dados de pesquisa que descrevem casos de pacientes preenchidos por fisioterapeutas foram usados para calcular pontuações em oito indicadores de qualidade para medir a qualidade do atendimento fisioterapêutico. Os métodos de coleta de dados foram avaliados para comparabilidade sobre integridade e correção. Os resultados obtidos, indicaram que cinco dos oito indicadores de qualidade podem ser extraídos dos registros eletrônicos de saúde. Três foram omitidos do conjunto de indicadores, pois se mostraram muito difíceis de serem extraídos. Outro indicador se mostrou incomparável devido a erros no software de extração de alguns dos registros. Três em quatro indicadores comparáveis tiveram melhor desempenho em relação à integridade. Os autores chegaram à conclusão que se deve simplificar o uso de dados de registros eletrônicos de saúde para a qualidade da pesquisa de cuidados.

O próximo trabalho considerado, foi desenvolvido por Vimalachandran et al (2016). Os autores realizaram uma revisão bibliográfica que abordou o impacto da integridade de dados em sistemas de registros eletrônicos de saúde e foi desenvolvido um modelo para preservar a integridade dos dados nesses sistemas. Conclui-se que a técnica de pseudo minimização utilizada é adequada, no entanto, são necessárias mais pesquisas sobre a prevalência para cada tipo de risco e sobre o impacto na integridade desses registros, assim como na sua segurança e qualidade do atendimento. Por fim, os autores afirmam que há necessidade de mais pesquisas sobre as causas dos erros relacionados aos registros, como a definição de estratégias eficazes para a sua devida preservação e correção.

Doktorchik et al (2019) desenvolveram um estudo com o objetivo de avaliar qualitativamente os pontos fortes e as barreiras em relação à qualidade da codificação clínica na perspectiva dos gestores da informação em saúde. Os

resultados demonstraram que as barreiras comuns à qualidade da codificação clínica incluíram documentação de gráficos incompleta, desorganizada e a falta de comunicação entre os médicos. Além disso, a qualidade da codificação clínica apresentou recursos limitados, como a falta de profissionais e orçamento disponível para os departamentos de gestão de informações de saúde. Gerentes relataram unanimemente que melhorias na qualidade da codificação clínica podem ser feitas a partir de duas ações: Oferecer programas de treinamento intensivos e simplificar as fontes de informação dos gráficos. Conclui-se que embora a qualidade da codificação clínica seja um aspecto geralmente considerado em todo o Canadá, os gerentes dessa área perceberam que a qualidade era limitada por um gráfico incompleto e inconsistente. Finalmente, os autores afirmam que o seu estudo apresenta novas evidências para a melhoria da qualidade da codificação clínica em todo o Canadá.

Huang et al (2020) descreveram diferentes dados e técnicas de fusão que podem ser aplicadas para combinar imagens médicas com registros eletrônicos de saúde, da mesma forma que os autores realizaram uma revisão sistemática de literatura. Os resultados obtidos demonstraram que a fusão multimodal para a automação das tarefas de imagens médicas melhora amplamente o desempenho de modelos de modalidade única. Os autores concluíram que o campo da fusão multimodal para aprendizagem profunda em imagens médicas está se expandindo e espera-se que novos métodos de fusão sejam desenvolvidos, em trabalhos futuros há a pretensão de se concentrar em terminologias e métricas compartilhadas, incluindo avaliação direta de diferentes abordagens de fusão multimodal quando aplicável.

Zheng e Hu (2020) utilizaram uma abordagem de fusão de dados longitudinais para modelar a progressão de doenças crônicas e buscar, dessa forma, aprimorar os tratamentos vigentes. Diferente do método convencional, os autores usaram apenas dados clínicos iniciais ou estáticos para modelar a progressão da doença para previsão de tempo, projetaram um termo de regularização temporal para manter a sucessividade dos dados de diferentes pontos no tempo e, simultaneamente, analisar os dados da fonte e nível de recurso com base em uma abordagem de regressão de regularização esparsa.

Os resultados obtidos trouxeram uma srie de contribuies que na seqncia foram avaliadas com dados reais oriundos de registros eletrnicos de sade. Conclui-se que uma estrutura de fuso de dados longitudinais pode realmente oferecer melhores *insights* clnicos do que outros mtodos de modelagem de progresso. Os autores afirmam que h expectativa para trabalhos futuros de melhorar o desempenho do presente mtodo e analisar uma maior quantidade de dados referentes aos pacientes.

Ali et al (2020) desenvolveram um sistema de sade inteligente, que busca realizar a previso de doenas cardacas e utiliza aprendizado profundo em um conjunto de dados e abordagens de fuso de recursos. Os resultados demonstraram que o sistema proposto obteve uma preciso de 98,5%, superior aos sistemas existentes, o que permitiu que o sistema proposto seja mais eficaz para a previso de doenas cardacas em comparao com outros mtodos de ltima gerao. Em trabalhos futuros, os autores afirmam que o desempenho da fuso de recursos ser aprimorado com o uso de tcnicas de minerao de dados para produzir um conjunto de dados para diagnsticos de doenas cardacas. Alm disso, novos mtodos sero projetados para reduo de recursos para lidar com um grande volume de registros de sade. Finalmente, um mtodo mais sofisticado ser investigado para a remoo de recursos irrelevantes e gerenciamento de valores ausentes.

Huang et al (2020) produziu um estudo que comparou diferentes arquiteturas de modelos de fuso multimodal que so capazes de utilizar dados de pixel de tomografia computadorizada volumtrica, angiografia pulmonar e dados clnicos de pacientes oriundos de registros eletrnicos de sade para classificar automaticamente os casos de Embolia Pulmonar. Os resultados demonstraram que o modelo de fuso aplicado foi significativamente melhor do que o baseado em pixels ou apenas em registros mdicos eletrnicos. Concluram que o modelo de fuso multimodal atinge uma separao clinicamente mais til entre casos positivos e negativos, assim como todos os casos centrais e segmentares positivos podem ser diagnosticados corretamente com um ponto operacional, atingindo um maior nvel de especificidade. Com base nesta anlise, o modelo de fuso pode ser mais indicado para integrao

na clnica devido a abordagem de ponta a ponta e alta sensibilidade, o que ajuda a reduzir falsos positivos de alerta clnico.

Colquhoun et al (2020) descrevem como o MPOG (Grupo de resultados perioperatrios multicntricos) abordaram as consideraes de dados estruturados, validao e acessibilidade para apoiar a integrao multicntrica de registros eletrnicos de sade. As concluses apresentadas priorizaram que para melhorar a qualidade dos dados s necessrios manter os dados disponveis em nvel local antes da transferncia para o Centro Coordenador, os dados devem ser rigorosamente validados em relao a mtricas padronizadas antes do uso, devem ser selecionados em fenotipos computveis que s facilmente acessveis e serem coletados para fins de pesquisa e melhoria da qualidade.

Lemma et al (2020) desenvolveram uma reviso de escopo que sintetizou resultados publicados de intervenes que visavam melhorar a qualidade e o uso dos dados em sistemas de informao de sade de rotina em pas de baixa e mdia renda. Os resultados obtidos demonstraram que as intervenes que aumentam a qualidade dos dados tm como alvo as unidades de sade e os seus gestores para uma melhor utilizao desses dados. Combinaes de aprimoramento de tecnologia com atividades de capacitao, avaliao de qualidade de dados e sistema de feedback, foram considerados teis para melhorar a qualidade dos dados. As intervenes que facilitam a disponibilidade de dados combinadas com o aprimoramento da tecnologia aumentaram a utilizao dos dados para um melhor planejamento, uma combinao de intervenes, abordando fatores comportamentais e tcnicos, o que permite melhorar tanto a qualidade como o uso dos dados.

Ao considerar os estudos apresentados, a prxima seo aborda uma discusso e interpretao dos resultados alcanados com base nos estudos anteriores, de modo a identificar quantitativamente o uso de cada critrio reconhecido nos artigos, subsidiando-os com conceitos cientficos retirados da literatura.

2.1 DISCUSSAO E INTERPRETAAO DOS RESULTADOS

Os critrios de qualidade de dados abordados nos artigos considerados foram mapeados e categorizados (Tabela 1), conforme suas caractersticas intrnsecas. Inicialmente, o critrio de qualidade mais abordado foi a acurcia, que, segundo Wang e Strong (1996), dada por quanto um determinado dado correto, confiavel e certificado. Outra definio apresentada por Redman (1996), onde a proximidade de um valor at outro valor parmetro para se considerar a confiabilidade dos dados. A partir destes conceitos, pode-se observar uma grande preocupao em relao a confiabilidade, que apresentaram resultados espelhados e utilizaram como base as caractersticas da dimenso acurcia. Ressalta-se tambm a importncia desse critrio, uma vez que avaliar a confiabilidade, bem como determinar se um dado correto e certificado de extrema importncia para os mais distintos campos de atuao.

Ao seguir a ordem da Tabela 1, o prximo critrio verificado foi relacionado a preciso e granularidade de dados. A preciso pode variar de acordo com suas medies sobre um determinado valor. No que se refere a granularidade de dados, seu significado se da diante da sua capacidade de distinguir objetos (FURBER, 2016). Dos trabalhos analisados, 9 trataram deste critrio, avaliaram a preciso de um determinado dado ou conjuntos de dados, que faziam parte de um processo especfico, bem como sua granularidade e apresentaram suas distines dos objetos do mundo real.

O terceiro critrio de qualidade analisado foi a completude, definido por Wand e Wang (1995) como sendo o grau no qual uma coleo de dados inclui dados diante da descrio de um conjunto de objetos que se relaciona com o mundo real. Este critrio de qualidade foi identificado em 11 trabalhos, no qual notou-se um comprometimento com descrio de dados, visando prover todas as informaes necessrias para atingir um objetivo. Em alguns casos, pode-se observar uma complexidade na fase em que se estabelece relaes com os dados tratados e o mundo real. Tal fato abordado por Jarke et al. (1995), que associa a definio de completude com o conceito de dado ausente em um conjunto especfico.

Dando sequncia, o prximo critrio de avaliao a consistncia.

Abordada em 12 pesquisas analisadas, notou-se sua relevncia no sentido de determinar conflitos entre os dados, identificando se a funo e a posio de um determinado dado so vlidas. Para isso, conexes que englobam todo o contexto em torno de um dado tambm foram verificadas, utilizando informaes prvias para tal ato. A consistncia no contexto de qualidade um critrio que pode impactar diretamente o resultado esperado, uma vez que, sua abordagem se baseia mediante regras semnticas sobre um determinado conjunto de dados.

O prximo critrio abordado a integridade. Para Vimalachandran et al (2016) integridade refere-se a preciso, qualidade interna e confiabilidade dos dados, ou seja, de acordo com o autor, a aplicao do critrio de integridade garante a qualidade dos dados. Este critrio foi identificado em 9 trabalhos analisados, o que se pode perceber que a integridade dos dados em pronturios eletrnicos envolve a preciso da documentao completa do registro de sade, o que abrange a governana da informao, a identificao do paciente, a preciso na etapa de coleta dos dados e a validao de registros.

O critrio de atualidade foi relatado no trabalho desenvolvido por Weiskopf, N et al (2013). Para o autor, os dados so considerados atuais se forem registrados em pronturios eletrnicos, dentro de um perodo razoavel ou se estes dados so representativos para o estado do paciente no momento de interesse desejado. Desta forma, o autor afirma que o critrio atualidade no redutvel e extremamente importante no que se refere a reutilizao de dados em pronturios eletrnicos.

O prximo critrio presente na Tabela 1 temporalidade. Wand e Wang (1996) definem o termo como o limite para o uso apropriado do dado. Posteriormente, Bovee et al (2001) abordam dois componentes para o critrio de temporalidade, primeiro, a idade, sendo o qunho velho um determinado dado e, segundo a volatilidade, abordando sobre a instabilidade e freqncia de mudana de um atributo pertencente a uma entidade. Seguindo as definies do termo, Liu e Chi (2002) abordam o critrio de forma a analisar se os dados esto atualizados suficientemente para realizar uma tarefa. Nos trabalhos analisados, pode-se observar o descarte de alguns conjuntos de dados por no estarem

totalmente alinhados e atualizados com a necessidade em questo, desta forma, faz-se presente o uso do critrio temporalidade em apenas 2 trabalhos.

O critrio confidencialidade se fez presente em 4 trabalhos. O critrio de acurcia, que tambm trata de questes relacionadas a confiabilidade dos dados, tambm esteve presente nos mesmos trabalhos, de modo a maximizar o grau de confiana de um determinado dado, bem como certific-lo e garantir seu correto uso em um processo.

Por fim, o critrio de plausibilidade teve sua referncia em apenas 1 trabalho, que tambm abordou critrios como consistncia, granularidade, preciso e completude, visando garantir que um dado est consistente, preciso e pertinente para um uso correto, de acordo com conceitos abordados no critrio de completude, como a descrio de um dado mediante seu objeto correspondente no mundo real.

A tabela 1 apresenta os critrios de qualidade, seguindo a ordem de colunas: 1) acurcia, 2) preciso e granularidade, 3) completude, 4) consistncia, 5) integridade, 6) atualidade, 7) temporalidade, 8) confidencialidade e 9) plausibilidade.

Tabela 1 - Categorizao dos artigos de acordo com os critrios de qualidade de dados abordados

Estudos Considerados	Acur.	Prec. e gran.	Comp.	Consist.	Integ.	Atua.	Temp.	Conf.	Plaus.
Thiru, K et al. (2003)		✓	✓	✓	✓				
Orfanidis, Bamidis e Eaglestone (2004)	✓				✓			✓	
Tang et al (2007)	✓	✓		✓					
Ndira, Rosenberge r e Wetter (2008)	✓	✓							
Miettinen e Korhonen (2008)	✓	✓	✓	✓					
Moulton, Chaczko e Karatovic (2009)	✓			✓					
Duarte et al (2010)					✓				
Pompilio e Ermertice (2011)	✓			✓					
Dixon, MCGowan e Grannis (2011)			✓						

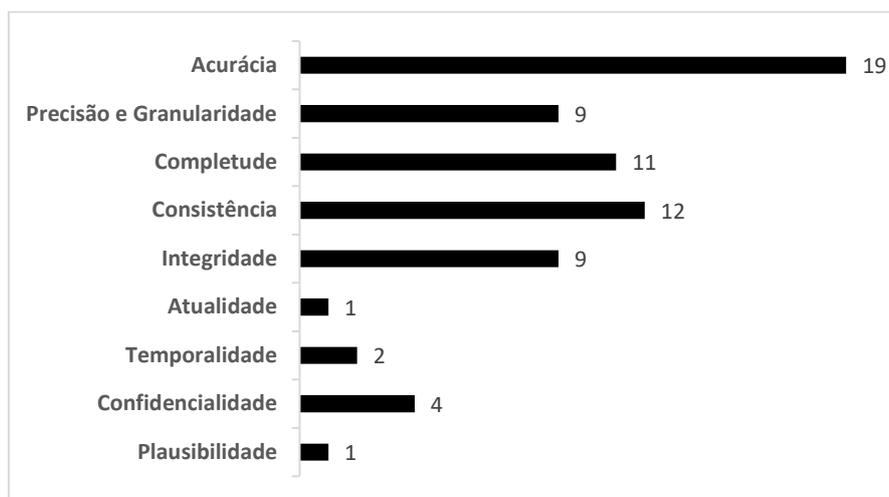
Kahn et al (2012)	✓				✓				
Weiskopf, N et al. (2013)		✓	✓	✓		✓			✓
Almutiry, Wills e Crowder (2013)	✓		✓	✓			✓		
Bowman (2013)	✓		✓	✓					
Liaw et al (2014)	✓	✓	✓						
Baron (2015)	✓		✓	✓					
Dziadkowiec et al (2016)	✓			✓			✓		
Sholte et al (2016)	✓				✓				
Vimalachandran et al (2018)	✓				✓			✓	
Doktorchik et al (2019)	✓		✓						
Huang et al (2020)		✓		✓	✓				
Zheng e Hu (2020)	✓	✓						✓	
Ali et al (2020)		✓			✓			✓	
Huang et al (2020)	✓			✓					
Colquhoun et al (2020)	✓		✓		✓				
Lemma et al (2020)	✓		✓						

Fonte: Elaborada pelos autores (2021)

Os trabalhos considerados no estudo no presente artigo no que diz respeito aos critrios de volatilidade, ambiguidade e correlao. O critrio presente em 76% dos trabalhos a acurcia, seguido pelo critrio consistncia, que foi mencionado em 48% dos artigos considerados. A completude foi abordada em 44% dos trabalhos. Os critrios preciso, granularidade e integridade estavam presentes em 36%. J os critrios menos citados foram a atualidade e plausibilidade, que participam de 4%, equivalente a apenas um trabalho, seguidos pelos critrios temporalidade e confidencialidade, que participaram de 8% e 16% respectivamente dos trabalhos. Esta representao pode ser vista no Grfico 1.

Diante dos resultados observados, o notrio que certos critrios foram mais explorados, por, isso no descarta a relevncia dos outros critrios, uma vez que sua aplicao depende de cenrios e situaes especficas.

Grfico 1 - Representao dos critrios de qualidade de dados.



Fonte: Elaborada pelos autores (2022)

3 CONSIDERAÇÕES FINAIS E TRABALHOS FUTUROS

A qualidade de dados é um tema propicio como objeto de estudo de diversos campos da cincia, tornando-se parte de um escopo interdisciplinar que abrange diversas áreas do conhecimento. O tema ganha ainda mais fora ao visar a quantidade de dados que é disponibilizada diariamente no mundo todo, muitos deles, podem ser mais bem disponibilizados e apresentados mediante utilizao de conceitos e tcnicas abordadas nesta pesquisa.

Diante da seo de discusso e resultados apresentada neste artigo, juntamente com as teorias sobre qualidade de dados e seus critrios de avaliao, observa-se tanto uma sobrecarga como uma carncia de uso em critrios especficos, sendo acurcia, consistncia e completude como os critrios mais abordados ao longo dos artigos e atualidade, temporalidade, confidencialidade e plausibilidade como aqueles que apresentaram uma carncia no sentido de serem mais bem abordados e explorados.

Com isso, conclui-se que há necessidade da compreenso de que trabalhar com a qualidade de dados envolve uma srie de critrios direcionados, inclusive, com o ciclo de vida dos dados (coleta, armazenamento, recuperao e descarte) que emerge de um campo informacional. Torna-se importante ressaltar que é possvel trabalhar, de maneira mais aprofundada, com critrio

especifico, no entanto, necessrio visualizar com clareza a amplitude relacionada s possibilidades de critrios de dados.

Diante deste contexto e corroborando com Sant'Ana (2016), a Cincia da Informao pode e deve contribuir em ambientes que contam com a presena de acesso e uso intensivo de dados, ao buscar elementos que possibilitam a construo de estruturas de referncias que permitam identificar caractersticas em contextos especificos. Com isto, acredita-se que a presena da Cincia da Informao em estudos referentes qualidade de dados podem possibilitar a identificao, o reconhecimento de fatores e caractersticas que propiciem uma ampliao da harmonia entre os profissionais envolvidos no processo e na performance da utilizao desses dados, favorecendo assim, o desenvolvimento da rea.

Graas a natureza interdisciplinar e colaborativa da Cincia da Informao, foi possvel lidar com questes de qualidade e fuso de dados pertencentes tanto a sistemas gerais de informao, como aos registros eletrnicos de sade, isto , conforme afirma Sant'Ana (2019), esta rea lida diretamente com a camada informacional de determinado campo, sendo capaz de atender as especificidades que o acesso a dados acrescenta.

Ao considerar ainda a interdisciplinaridade da Cincia da Informao, entende-se que h possibilidade de desenvolver trabalhos que explorem, de forma mais aprofundada, os critrios que apresentaram dficit de pesquisas em torno da literatura, dessa forma, torna-se importante realizar estudos de caso peridicos em unidades de sade, de forma a avaliar os critrios j existentes nesses ambientes, tornando-se um cenrio factvel com anlises de dados, visando oferecer recomendaes de diretrizes para o escopo em questo. E, considerando a realidade vivenciada em ambientes hospitalares, especialmente no perodo da pandemia ocasionada pelo vrus SARS-CoV-2, abre-se questes de pesquisas que podem avaliar como a qualidade dos dados foram consideradas neste perodo pandmico e se os dados de baixa qualidade podem influenciar ou provocar erros.

REFERNCIAS

- ALI, Farman et al. A smart healthcare monitoring system for heart disease prediction based on ensemble deep learning and feature fusion. **Information Fusion**, v. 63, p. 208-222, 2020.
- ALMUTIRY, Omar; WILLS, Gary; CROWDER, Richard. Towards a framework for data quality in electronic health records. **IADIS International Conference, e-Society**, Lisbon, Portugal. 2013.
- BARON, Richard J. Quality improvement with an electronic health record: achievable, but not automatic. **Annals of internal medicine**, v. 147, n. 8, p. 549-552, 2007.
- BOVEE, Matthew; SRIVASTAVA, Rajendra P.; MAK, Brenda. A conceptual framework and belief-function approach to assessing overall information quality. **International journal of intelligent systems**, v. 18, n. 1, p. 51-74, 2003.
- BOWMAN, Sue. Impact of electronic health record systems on information integrity: quality and safety implications. **Perspectives in health information management**, v. 10, n. Fall, 2013.
- BROWN, Philip JB; WARMINGTON, Victoria. Data quality probes—exploiting and improving the quality of electronic patient record data and patient care. **International journal of medical informatics**, v. 68, n. 1-3, p. 91-98, 2002.
- BYRD, James Brian et al. Data quality of an electronic health record tool to support VA cardiac catheterization laboratory quality improvement: the VA Clinical Assessment, Reporting, and Tracking System for Cath Labs (CART) program. **American heart journal**, v. 165, n. 3, p. 434-440, 2013.
- CHAN, Kitty S.; FOWLES, Jinnat B.; WEINER, Jonathan P. Electronic health records and the reliability and validity of quality measures: a review of the literature. **Medical Care Research and Review**, v. 67, n. 5, p. 503-527, 2010.
- COLQUHOUN, Douglas A. et al. Considerations for integration of perioperative electronic health records across institutions for research and quality improvement: the approach taken by the Multicenter Perioperative Outcomes Group. **Anesthesia and analgesia**, v. 130, n. 5, p. 1133, 2020.
- CONEGLIAN, C. S.; DIEGER, R.; SEGUNDO, J. E. S.; CAPRETZ, M. A. M. O papel da web semntica nos processos da big data. **Encontros Bibli: Revista Eletrnica de Biblioteconomia e Cincia da Informao**, v. 23, n. 53, p. 137-146, 2018. Disponvel em: <http://www.brapci.inf.br/index.php/res/v/39601>. Acesso em: 02 setembro 2020.
- DIXON, Brian E.; MCGOWAN, Julie J.; GRANNIS, Shaun J. Electronic laboratory data quality and the value of a health information exchange to

support public health reporting processes. In: AMIA annual symposium proceedings. **American Medical Informatics Association**, 2011. p. 322.

DOKTORCHIK, Chelsea et al. A qualitative evaluation of clinically coded data quality from health information manager perspectives. **Health Information Management Journal**, v. 49, n. 1, p. 19-27, 2020.

DUARTE, Julio et al. Data quality evaluation of electronic health records in the hospital admission process. In: 2010 **IEEE/ACIS 9th International Conference on Computer and Information Science**. IEEE, 2010. p. 201-206.

DZIADKOWIEC, Oliwier et al. Using a data quality framework to clean data extracted from the electronic health record: a case study. **eGEMs**, v. 4, n. 1, 2016.

FONTES-PEREIRA, A. Reviso Sistemtica da Literatura: Como Escrever um Artigo Cientfico em 72 Horas. Rio de Janeiro: **Edio do Kindle.**, 2017.

FURBER, Christian. Data Quality in the Semantic Web. In: **Data Quality Management with Semantic Technologies**. Springer Gabler, Wiesbaden, 2016.

HUANG, Shih-Cheng et al. Fusion of medical imaging and electronic health records using deep learning: a systematic review and implementation guidelines. **NPJ digital medicine**, v. 3, n. 1, p. 1-9, 2020.

HUANG, Shih-Cheng et al. Multimodal fusion with deep neural networks for leveraging CT imaging and electronic health record: a case-study in pulmonary embolism detection. **Scientific reports**, v. 10, n. 1, p. 1-9, 2020.

JARKE, Matthias et al. ConceptBase—a deductive object base for meta data management. **Journal of Intelligent Information Systems**, v. 4, n. 2, p. 167-192, 1995.

KAHN, Michael G. et al. A pragmatic framework for single-site and multisite data quality assessment in electronic health record-based clinical research. **Medical care**, v. 50, 2012.

LEMMA, Seblewengel et al. Improving quality and use of routine health information system data in low-and middle-income countries: A scoping review. **PLoS one**, v. 15, n. 10, p. e0239683, 2020.

LIAW S-T et al. **Integrating electronic health record information to support integrated care: Practical application of ontologies to improve the accuracy of diabetes disease registers**. J Biomed Inform, 2014.

LIU, Liping; CHI, Lauren. Evolutional Data Quality: A Theory-Specific View. In: **ICIQ**. 2002. p. 292-304.

NDIRA, S. P.; ROSENBERGER, K. D.; WETTER, T. Assessment of data quality of and staff satisfaction with an electronic health record system in a developing country (Uganda). **Methods of information in medicine**, v. 47, n. 06, p. 489-498, 2008.

MIETTINEN, Merja; KORHONEN, Maritta. Information quality in healthcare: coherence of data compared between organization's electronic patient records. In: 2008 21st **IEEE International Symposium on Computer-Based Medical Systems**. IEEE, 2008. p. 488-493.

MORAES, M. F. Requisitos de qualidade e segurana para pronturrios do paciente. **Informao em Pauta**, v. 3, p. 141-160, 2018. Disponvel em: <http://www.brapci.inf.br/index.php/res/v/106556>. Acesso em: 02 setembro 2020.

MORANDI, M. I. W. M.; CAMARGO, L. F. R. Reviso Sistemtica de Literatura. In: Design Science Research: Mtodo de pesquisa para avano da cincia e tecnologia. Porto Alegre: **Bookman**, 2015. p. 181.

MOULTON, B. D.; CHACZKO, Z. C.; KARATOVIC, Mark. Data fusion and aggregation methods for pre-processing ambulatory monitoring and remote sensor data for upload to personal electronic health records. **International Journal of Digital Content Technology** a..., 2009.

ORFANIDIS, Leonidas; BAMIDIS, Panagiotis D.; EAGLESTONE, Barry. Data quality issues in electronic health records: an adaptation framework for the Greek health system. **Health informatics journal**, v. 10, n. 1, p. 23-36, 2004.

POMPILIO, Antonio Pompilio Junior; ERMETICE, Edson. Indicadores de uso do pronturio eletrnico do paciente. **Journal of Health Informatics**, v. 3, n. 1, 2011.

REDMAN, T. C. Data quality for the information age. Norwood, **Mass.: Artech House**. 1996.

SANT'ANA, Ricardo Csar Goncalves. Ciclo de vida dos dados: uma perspectiva a partir da cincia da informao. **Informao & Informao**, [S.l.], v. 21, n. 2, p. 116-142, dez. 2016. ISSN 1981-8920. Disponvel em: <<http://www.uel.br/revistas/uel/index.php/informacao/article/view/27940>>. Acesso em: 17 abr. 2021. doi:<http://dx.doi.org/10.5433/1981-8920.2016v21n2p116>.

SANT'ANA, R.C.G. Campo Informacional Resultante da Interao de Ciclos de Vida dos Dados. In: DIAS, G.; FREIRE, B. (org). **Dados Cientficos: perspectivas e desafios**. Editora UFPB - Joao Pessoa. 2019 p.5-19

SCHOLTE, Marijn et al. Data extraction from electronic health records (EHRs) for quality measurement of the physical therapy process: comparison between EHR data and survey data. **BMC medical informatics and decision making**, v. 16, n. 1, p. 141, 2016.

TANG, Paul C. et al. Comparison of methodologies for calculating quality measures based on administrative data versus clinical data from an electronic health record system: implications for performance measures. **Journal of the American Medical Informatics Association**, v. 14, n. 1, p. 10-15, 2007.

THIRU, Krish; HASSEY, Alan; SULLIVAN, Frank. Systematic review of scope and quality of electronic patient record data in primary care. **Bmj**, v. 326, n. 7398, p. 1070, 2003.

VIMALACHANDRAN, Pasupathy et al. Ensuring data integrity in electronic health records: a quality health care implication. In: 2016 **International Conference on Orange Technologies (ICOT)**. IEEE, 2016. p. 20-27.

WAND, Yair; WANG, Richard Y. Anchoring data quality dimensions in ontological foundations. **Communications of the ACM**, v. 39, n. 11, p. 86-95, 1996.

WANG, R. Y.; STRONG, D. M. Beyond accuracy: what data quality means to data consumers. **Journal of Management Information Systems**, v. 12, n. 4, p. 5-33, 1996.

WEISKOPF, Nicole Gray; WENG, Chunhua. Methods and dimensions of electronic health record data quality assessment: enabling reuse for clinical research. **Journal of the American Medical Informatics Association**, v. 20, n. 1, p. 144-151, 2013.

ZHENG, Yi; HU, Xiangpei. Healthcare predictive analytics for disease progression: a longitudinal data fusion approach. **Journal of Intelligent Information Systems**, v. 55, p. 351-369, 2020.

HEALTH DATA QUALITY CRITERIA: A QUANTITATIVE ANALYSIS

ABSTRACT

Objective: This article proposes a reflection on how data quality criteria have been approached in works that discuss data from the health area, making it possible to recognize the panorama on these criteria and identify the gaps that require greater efforts. **Methodology:** The methodological procedures consist of a systematic literature review that aimed to identify, analyze and quantify the criteria of quality of data that are addressed in health. **Results:** As results, the quality criteria of mapped and categorized data are presented, identifying the accuracy, consistency and completeness as the most addressed criteria and currentness, temporality, confidentiality and plausibility, being the least mentioned. **Conclusions:** It is concluded that there is both overload and the lack of use of certain criteria, therefore, the criteria that presented this lack generate the possibility of being better addressed and explored in future works and, also, the relevance of these criteria cannot be ruled out, considering that its application depends

on specific scenarios and situations.

Descriptors: Data quality. Data quality criteria. Health data quality. Health.

CRITERIOS DE CALIDAD DE LOS DATOS SANITARIOS: UN ANÁLISIS CUANTITATIVO

RESUMEN

Objetivo: Este artículo propone una reflexión sobre cómo los criterios de calidad de los datos han sido abordados en trabajos que discuten datos del área de la salud, posibilitando reconocer el panorama sobre esos criterios e identificar los vacíos que requieren mayores esfuerzos. **Metodología:** Los procedimientos metodológicos consisten en una revisión sistemática de la literatura que buscó identificar, analizar y cuantificar los criterios de calidad de los datos que se abordan en el área de la salud. **Resultados:** Como resultados se presentan los criterios de calidad de los datos mapeados y categorizados, identificando la exactitud, consistencia e integridad como los criterios más abordados y la oportunidad, la temporalidad, confidencialidad y plausibilidad, siendo los menos mencionados. **Conclusiones:** Se concluye que existe tanto la sobrecarga como la falta de uso de ciertos criterios, por tanto, los criterios que presentaron esta carencia generan la posibilidad de ser mejor abordados y explorados en trabajos futuros y, además, no se puede descartar la relevancia de estos. fuera de criterios, considerando que su aplicación depende de escenarios y situaciones específicas.

Descriptores: Calidad de los datos. Criterios de calidad de los datos. Calidad de los datos sanitarios. Salud.

Recebido em: 28.05.2021

Aceito em: 19.12.2022