UNIVERSIDADE DE SÃO PAULO
ESCOLA DE ENGENHARIA DE SÃO CARLOS
FACULDADE DE MEDICINA DE RIBEIRÃO PRETO
INSTITUTO DE QUÍMICA DE SÃO CARLOS

ALEXANDRE HENRIQUE MACRIS

**Estudo de Modelos para Predição de Reinternação em Saúde Mental no
Sistema Único de Saúde (SUS)**

São Carlos

2020

ALEXANDRE HENRIQUE MACRIS

**Estudo de Modelos para Predição de Reinternação em Saúde Mental no
Sistema Único de Saúde (SUS)**

VERSÃO CORRIGIDA

Dissertação apresentada ao Programa de Pós-Graduação Interunidades em Bioengenharia da Escola de Engenharia de São Carlos – Faculdade de Medicina de Ribeirão Preto e Instituto de Química de São Carlos da Universidade de São Paulo, como requisito para a obtenção do Título de Mestre em Ciências.

Área de concentração: Bioengenharia

Orientador: Prof. Dr. Domingos Alves

São Carlos

2020

AUTORIZO A REPRODUÇÃO E DIVULGAÇÃO TOTAL OU PARCIAL DESTES
TRABALHOS, POR QUALQUER MEIO CONVENCIONAL OU ELETRÔNICO,
PARA FINS DE ESTUDO E PESQUISA, DESDE QUE CITADA A FONTE.

Ficha catalográfica elaborada pela Biblioteca Prof. Dr. Sérgio Rodrigues
Fontes da EESC/USP

M174e	<p>Macris, Alexandre Henrique Estudo de modelos para predição de reinternação em saúde mental no sistema único de saúde (SUS) / Alexandre Henrique Macris ; orientador Domingos Alves. -- São Carlos, 2020.</p> <p>Dissertação (Mestrado) - Programa de Pós Graduação Interunidades em Bioengenharia e Área de Concentração em Bioengenharia - Escola de Engenharia de São Carlos, Faculdade de Medicina de Ribeirão Preto, Instituto de Química de São Carlos da Universidade de São Paulo, 2020.</p> <p>1. Reinternações hospitalares. 2. Saúde mental. 3. Modelos de predição. 4. Descoberta de conhecimento em bases de dados. I. Título.</p>
-------	---

FOLHA DE JULGAMENTO

Candidato(a): Alexandre Henrique Macris

Título: "Estudo de modelos para predição de reinternação em saúde mental no Sistema Único de Saúde (SUS)"

Data da defesa: 17/02/2020

Comissão Julgadora	Assinatura	Resultado
Prof(a). Dr(a). Domingos Alves FMRP/USP		Não Votante
Prof(a). Dr(a). Newton Shydeo Brandão Miyoshi FAEPA/HCFMRP		Aprovado
Prof(a). Dr(a). Rui Pedro Charters Lopes Rijo IPLeiria		Aprovado
Prof(a). Dr(a). João Mazzoncini de Azevedo Marques FMRP/USP		Aprovado

*Dedico este trabalho a meu pai
Jorge, que me ensinou que desafios
existem para serem enfrentados, a
minha mãe Dagmar, que sempre
apoiou minhas escolhas, e também a
minha esposa Ana Carolina, cujo
incentivo e compreensão me
ajudaram nesta jornada.*

AGRADECIMENTOS

Aos familiares e amigos, pela compreensão por minha ausência em vários dos momentos de confraternização.

Ao Professor Dr. Domingos Alves pelos ensinamentos, pela paciência, pela oportunidade e disponibilidade em me orientar, e por compreender minhas dificuldades no decorrer deste caminho.

Ao Professor Dr. Rui Pedro Charters Lopes Rijo, pelas conversas descontraídas e pelas dicas valiosas durante o caminho.

Aos colegas do Laboratório de Inteligência em Saúde, em especial ao André Vinci e Newton Miyoshi pelo compartilhamento de experiência em relação aos dados do SIHSUS e SISAM; também à colega Inácia Lima pela ajuda nos procedimentos de submissão de projeto na Plataforma Brasil.

Ao programa de Pós-graduação Interunidades em Bioengenharia e seus funcionários, em especial à secretária Márcia Geromini pelas dúvidas sanadas, conversas, apoio e incentivo.

Por fim, a todos aqueles que me ajudaram a atingir este objetivo de forma direta ou indireta.

RESUMO

MACRIS, A. H. **Estudo de modelos para predição de reinternação em saúde mental no sistema único de saúde (SUS)**. 2020. 68f. Dissertação (Mestrado) - Programa de Pós-Graduação Interunidades em Bioengenharia, Escola de Engenharia de São Carlos, Faculdade de Medicina de Ribeirão Preto e Instituto de Química de São Carlos da Universidade de São Paulo, São Carlos, 2020.

Introdução: O fenômeno da reinternação hospitalar não planejada é um problema crítico que afeta vários países, entre os efeitos diretos desta situação estão o descontrole da rotatividade de leitos hospitalares, o alto custo financeiro oriundo da duração de estadia e a diminuição da qualidade de vida dos pacientes. **Objetivo:** A busca por um modelo de predição de reinternação hospitalar eficiente torna-se relevante, considerando que seus resultados podem favorecer a tomada de decisões de forma antecipada dentro das instituições hospitalares, mitigando ou até evitando que um paciente seja internado novamente após a alta. **Métodos:** Dentre as atividades propostas neste trabalho estão a submissão de duas fontes de dados distintas à algumas etapas do processo de descoberta de conhecimento em bases de dados. **Resultados:** Inicialmente foi realizada uma busca nas bases de dados online a fim de avaliar a acurácia dos resultados obtidos na aplicação de modelos de predição de reinternação em pacientes de saúde mental, também comparar as variáveis mais utilizadas nestes modelos com as existentes nas fontes de dados do repositório. Em um segundo momento, as variáveis eleitas foram submetidas à análise e processamento sob a perspectiva da descoberta de conhecimento em bases de dados. **Conclusão:** A presença de determinadas variáveis nas fontes de dados que também foram utilizadas em estudos anteriores é um importante indicador da viabilidade de elaboração e aplicação de um modelo de predição de reinternações, a intenção é prosseguir nas etapas do processo de conhecimento em base de dados para elaboração deste modelo no futuro.

Palavras-chave: Reinternações hospitalares. Saúde mental. Modelos de predição. Descoberta de conhecimento em bases de dados.

ABSTRACT

MACRIS, A. H. **Study of prediction models for mental health rehospitization in the sistema único de saúde (SUS)**. 2020. 68f. Dissertação (Mestrado) - Programa de Pós-Graduação Interunidades em Bioengenharia, Escola de Engenharia de São Carlos, Faculdade de Medicina de Ribeirão Preto e Instituto de Química de São Carlos da Universidade de São Paulo, São Carlos, 2020.

Introduction: The unplanned hospital readmission phenomenon is a critical problem that affects many countries. Among the direct effects of this situation are the uncontrolled rotation of hospital beds, the high financial cost resulting from the length of stay and the decrease in the life quality of patients. **Objective:** The search for an efficient readmission prediction model becomes relevant, considering that its results may support early decision making within hospital institutions, mitigating or even preventing a patient for new hospitalization after discharge. **Methods:** Among the activities proposed in this paper are the submission of two distinct data sources to knowledge discovery process in databases stages. **Results:** Initially, an online database search was lead to evaluate the accuracy of the results obtained in the application of readmission prediction models in mental health patients, as well as to compare the variables most commonly used in these models with those in data sources repository. In a second moment, the elected variables were submitted to analysis and processing from knowledge discovery in databases perspective. **Conclusion:** The presence of certain variables in data sources that were also used in previous studies is an important indicator of the feasibility of elaboration and application for readmission prediction model, the intention is proceed in the steps of the database knowledge process to elaborate this model in the future.

Keywords: Hospital readmissions. Mental health. Prediction models. Knowledge discovery in databases.

LISTA DE ILUSTRAÇÕES

Figura 1 - Visão geral das etapas que compõem o processo <i>KDD</i>	29
Figura 2 - Processo de triagem.	33
Figura 3 - Proporção de preenchimento de variáveis do SIHSUS.	39
Figura 4 – Proporção de preenchimento de variáveis no SISAM.	40
Figura 5 - Distribuição de quantidade de variáveis com valores ausentes por número de registros no SIHSUS.....	41
Figura 6 - Distribuição de quantidade de variáveis com valores ausentes por número de registros no SISAM.	42
Figura 7 - Proporção das internações gerais no SISAM agrupadas por faixas de diagnóstico (CID).....	44
Figura 8 - Evolução semestral da quantidade de internações gerais no SISAM por faixas de diagnóstico (CID).	44
Figura 9 - Tendência semestral das reinternações no SISAM por intervalo de dias.	45

LISTA DE TABELAS

Tabela 1 - Modelos de predição de reinternação utilizados entre 2011 e 2015.....	34
Tabela 2 - Variáveis utilizadas em três estudos e variáveis disponíveis nas fontes SIHSUS e SISAM.	36
Tabela 3 - Total de internações de acordo com critérios de seleção de registros.....	37
Tabela 4 - Proporção de preenchimento de variáveis no SIHSUS.....	39
Tabela 5 - Proporção de preenchimento de variáveis no SISAM.	40
Tabela 6 - Grau de incompletude nos registros do SIHSUS.	41
Tabela 7 - Grau de incompletude nos registros do SISAM.....	42
Tabela 8 - Variáveis utilizadas nos estudos e fontes de dados do repositório.....	43
Tabela 9 - Internações e reinternações do SISAM agrupadas por faixa de diagnósticos.....	47
Tabela 10 - Características dos pacientes que não sofreram reinternação no SISAM agrupadas por faixa de diagnósticos (n=2.489).	48

LISTA DE ABREVIATURAS E SIGLAS

n	-	Número de objetos em um conjunto
$\bar{x}$	-	Média aritmética
f	-	Frequência relativa
%	-	Porcentagem
AIH	-	Autorização de Internação Hospitalar
<i>AUC</i>	-	<i>Area Under the ROC Curve</i>
CEP	-	Comitê de Ética em Pesquisa
CID	-	Classificação Internacional de Doenças - 10ª Revisão
CNS	-	Conselho Nacional de Saúde
CSV	-	<i>Comma-Separated Values</i>
<i>DBC</i>	-	<i>dBase III Plus compressed</i>
<i>DBF</i>	-	<i>dBase III Plus</i>
DRS-XIII	-	Departamento Regional de Saúde de Ribeirão Preto
<i>KDD</i>	-	<i>Knowledge Discovery in Databases</i>
LIS	-	Laboratório de Inteligência em Saúde
<i>ROC</i>	-	<i>Receiver Operating Characteristic</i>
SGBD	-	Sistema de Gerenciamento de Banco de Dados
SIHSUS	-	Sistema de Informações Hospitalares do SUS
SISAM	-	Sistema de Informação em Saúde Mental
TCLE	-	Termo de Consentimento Livre e Esclarecido

SUMÁRIO

1. INTRODUÇÃO E CONTEXTUALIZAÇÃO	23
2. OBJETIVOS	25
2.1 Geral	25
2.2 Específicos.....	25
3. MATERIAL E MÉTODOS.....	27
3.1 Revisão Bibliográfica	27
3.2 Fontes de dados	27
3.2.1 SIHSUS	28
3.2.2 SISAM.....	28
3.3 Modelo de Estudo e Forma de Análise.....	28
3.4 O Processo <i>KDD</i>	28
3.4.1 Objetivos do processo.....	29
3.4.2 Criação de um conjunto de dados alvo	29
3.4.3 Limpeza de dados e pré-processamento	29
3.4.4 Redução de dados e projeção.....	30
3.4.5 Escolha do método de mineração de dados (<i>data mining</i>)	30
3.4.6 Escolha do(s) algoritmo(s) de <i>data mining</i>	30
3.4.7 <i>Data mining</i>	30
3.4.8 Interpretação de resultados	30
3.4.9 Ação sobre o conhecimento descoberto	30
3.5 Aspectos Éticos	30
4. RESULTADOS	33
4.1 Revisão sobre revisões sistemáticas em modelos de reinternação para todas as causas	33
4.2 Estudos sobre modelos de reinternação em saúde mental	35
4.3 Repositório de dados	36
4.4 Análise dos dados	37
4.5 Limpeza dos dados	38
4.6 Comparação entre variáveis	42
5. DISCUSSÃO.....	49
5.1 Modelos preditivos	49
5.2 Modelos de predição e o contexto da saúde mental no Brasil.....	50
5.3 O papel do SISAM	51

5.4 Limitações deste estudo.....	51
6. CONCLUSÕES	53
7. SUGESTÕES PARA TRABALHOS FUTUROS.....	55
REFERÊNCIAS	57
ANEXO I – PARECER DO COMITÊ DE ÉTICA EM PESQUISA.....	61
ANEXO II - DEFINIÇÃO DOS CAMPOS DA AIH REDUZIDA (SIH).....	63
ANEXO III - DEFINIÇÃO DOS CAMPOS DO SISAM (ADAPTADA)	67

1. INTRODUÇÃO E CONTEXTUALIZAÇÃO

Modelos matemáticos que têm condições de predizer o risco de nova internação hospitalar de um paciente tem sido alvo de constantes pesquisas (KANSAGARA *et al.*, 2011) e um fator importante na capacidade preditiva de um modelo é a escolha das melhores variáveis que podem incorporar o mesmo. A existência das variáveis nos estudos e similaridades nos resultados entre os mesmos embasam o fato de que as variáveis contidas nas fontes de dados do repositório local podem ter participação relevante na construção de um modelo para predição de reinternação hospitalar em saúde mental.

O fenômeno da reinternação hospitalar não planejada, ou "porta giratória", é um problema que afeta todos os países; além de causar uma interrupção na normalidade da vida do paciente e de seus familiares, também resulta em um alto custo financeiro para as instituições hospitalares (BOSCO *et al.*, 2014; JOYNT; JHA, 2012).

No âmbito da saúde mental, a situação nas redes de atenção não é diferente. Adicionalmente, elas são complexas pela heterogeneidade que apresentam e especificidade de seus processos. Justifica-se, assim, o estudo e a proposta de soluções para o suporte do pilar operacional nas redes de atenção em saúde mental de modo a garantir uma visão única da rede, dos seus processos e dos sistemas de informação de suporte. Neste cenário, torna-se importante a utilização de um modelo eficiente para predição de reinternação hospitalar, pressupondo que uma vez identificada a probabilidade de risco de internação, a instituição tenha condições de tomar medidas antecipadas a fim de diminuir ou até evitar que o paciente seja internado novamente após a alta. Uma das estratégias para tentar reduzir a taxa de reinternações hospitalares não planejadas é aplicação de modelos preditivos que tenham condições de identificar pacientes com alto risco de reinternação (KANSAGARA *et al.*, 2011). Estes modelos preditivos consideram, em geral, a exposição de registros eletrônicos de saúde (RES) a técnicas de análise de dados.

2. OBJETIVOS

2.1 Geral

Diante desse cenário, a presente dissertação de mestrado visa o estudo dos modelos de predição para reinternação de pacientes de saúde mental, cujos resultados podem ser incorporados como uma ferramenta de apoio à decisão ao Sistema de Informação em Saúde Mental (SISAM). Para fazer isso, será utilizada a abordagem *Knowledge Discovery in Databases* (*KDD*) para guiar o estudo sobre reinternação em pacientes de saúde mental.

2.2 Específicos

- Sintetizar os achados na literatura sobre revisões sistemáticas considerando os modelos de predição aplicados em reinternações de até 30 dias;
- Analisar os dados contidos nas fontes de dados do repositório, de acordo com o processo *KDD*;
- Descobrir as variáveis preditivas mais utilizadas nos modelos encontrados e comparar com as variáveis existentes nas fontes de dados selecionadas.

3. MATERIAL E MÉTODOS

3.1 Revisão Bibliográfica

Inicialmente foi realizada uma pesquisa sobre revisões de revisões sistemáticas a respeito de modelos de reinternação hospitalar em saúde pública. As seguintes palavras-chave foram consideradas relevantes e utilizadas na base de dados online Scopus:

- *public health*
- *hospital*
- *readmission*
- *review*

As entradas foram combinadas com o operador *AND*, obtendo como resultado a *string* de pesquisa: “*public health*” *AND* *hospital* *AND* *readmission* *AND* *review*. Os artigos retornados pela pesquisa foram selecionados em uma primeira etapa sem realizar distinções quanto ao país de origem, ou motivo de atendimento/internação, ou cenário (cirurgia/atendimento ambulatorial) ou ainda tipo de diagnóstico principal. Para realização da busca, a *string* de pesquisa foi aplicada em título e resumo no período de janeiro de 2010 até 2017.

Em uma pesquisa independente, também buscamos citações sobre modelos de predição de reinternação aplicados exclusivamente em saúde mental entre os anos de 2001 e 2015 para compor as análises. As entradas “*psychiatric*” *AND* *readmission* *AND* *estimate* *OR* *model* foram combinadas em uma *string* de pesquisa na base de dados Scopus.

Adicionalmente, e na intenção de aumentar o espectro de citações encontradas pela pesquisa inicial, refizemos a busca considerando outras bases de dados online (PubMed incluindo MEDLINE, Cochrane e IEEE Xplore) onde as entradas iniciais foram combinadas por disjunção com novas entradas, i.e., “*patient readmission*” *OR* “*inpatient re-admission*” *OR* “*revolving door*”. Embora um conjunto significativamente maior de citações tenha sido recuperado (n=311), o resultado não foi incorporado no processo de triagem (Figura 2) neste momento.

3.2 Fontes de dados

Para a realização deste trabalho, foram utilizados dados de duas fontes distintas: SIHSUS e SISAM. Por uma questão de comparabilidade, os dados trabalhados foram das internações que constam nessas bases de dados a partir de 2011, ano em que se iniciou a utilização do SISAM.

3.2.1 SIHSUS

O DATASUS (2009) é uma plataforma do Ministério da Saúde responsável por coletar, processar e disseminar informações sobre saúde que podem servir para subsidiar análises objetivas da situação sanitária, tomadas de decisão baseadas em evidências e elaboração de programas de ações de saúde. Um dos sistemas do DATASUS é o SIHSUS¹ (Sistema de Informações Hospitalares do SUS) cuja finalidade é registrar todos os atendimentos provenientes de internações hospitalares que foram financiadas pelo SUS. Os dados que servem como base para esse sistema são provenientes das Autorizações de Internação Hospitalar (AIHs), disponibilizados para comunidade pelo DATASUS e que serão utilizados neste trabalho.

3.2.2 SISAM

Como descrito anteriormente, o SISAM é o sistema de gerenciamento da saúde mental do DRS-XIII e seu banco de dados também servirá como fonte de dados para realização do presente trabalho. Nele há dados acerca do diagnóstico do paciente, data de internação e data da alta, motivo da alta, local de internação, local de residência, tempo de permanência, entre outros que poderão ser relevantes para execução do trabalho.

3.3 Modelo de Estudo e Forma de Análise

A pesquisa proposta tem por base uma perspectiva filosófica positivista, sendo do tipo *predictive research* (WALJEE; HIGGINS; SINGAL, 2014).

O modelo foi desenvolvido com base no processo *KDD*. A partir dos dados cadastrados no repositório que armazena informações de pacientes de saúde mental: SISAM e SIHSUS ao longo dos últimos 6 anos, onde será possível selecionar os doentes que sofreram reinternações psiquiátricas.

3.4 O Processo *KDD*

Geralmente, grande parte dos dados recuperados de fontes de dados representam valores atômicos, que em sua forma natural não favorecem a análise para que seja realizada uma tomada de decisão frente a um problema.

Para transformar os dados atômicos em conhecimento estratégico pode-se utilizar o processo *KDD*, que pode ser definido como “o processo, não trivial, de extração de informações implícitas, previamente desconhecidas e potencialmente úteis, a partir dos dados armazenados

¹ <http://www2.datasus.gov.br/DATASUS/index.php?area=060502>

em um banco de dados” (FAYYAD; PIATETSKY-SHAPIO; SMYTH, 1996), conforme ilustrado na Figura 1.

Figura 1 - Visão geral das etapas que compõem o processo *KDD*.

Fonte: Adaptado de Fayyad, Piatetsky-Shapiro e Smyth (1996)

O processo *KDD* pode ser desmembrado em uma sequência finita de etapas, sendo que o resultado de uma etapa depende do resultado das anteriores. As etapas são controladas pelo usuário e podem ser revisitadas a qualquer momento durante a execução do processo, representando um ciclo. Para este trabalho, serão aplicadas as etapas do processo *KDD*, sugeridas por Fayyad, Piatetsky-Shapiro e Smyth (1996), e descritas resumidamente a seguir:

3.4.1 Objetivos do processo

Desenvolver um entendimento do domínio da aplicação (e do conhecimento prévio) e identificar os objetivos do processo *KDD* sob o ponto de vista do cliente ou do usuário final

3.4.2 Criação de um conjunto de dados alvo

Selecionar um conjunto de dados (*dataset*), ou focar em um subconjunto de variáveis ou ainda amostra de dados, nos quais a descoberta de conhecimento será realizada

3.4.3 Limpeza de dados e pré-processamento

Operações básicas nesta etapa incluem remoção de ruídos, coleta de informações necessárias para modelar ou explicar os ruídos, estratégias para lidar com dados ausentes nos campos, contabilização de informações de tempo da sequência e também mudanças conhecidas

3.4.4 Redução de dados e projeção

Envolve encontrar recursos úteis para representação de dados, dependendo do objetivo da tarefa. Ao utilizar redução de dimensionalidade ou transformações, o número de variáveis relevantes pode ser reduzido, ou representações invariáveis de dados podem ser encontradas

3.4.5 Escolha do método de mineração de dados (*data mining*)

Combinar os objetivos do processo *KDD* (3.4.1) com um método específico de *data mining*, tais como: sumarização, classificação, regressão, clusterização, etc.

3.4.6 Escolha do(s) algoritmo(s) de *data mining*

Selecionar o(s) método(s) a ser(em) utilizado(s) na busca por padrões nos dados; isto inclui decidir quais modelos e parâmetros podem ser apropriados (e.g., modelos para dados qualitativos podem ser diferentes de modelos para dados quantitativos) e combinar um método específico de *data mining* (3.4.5) com os critérios gerais do processo *KDD* (e.g., usuário final pode estar mais interessado no entendimento do modelo do que em uma capacidade preditiva)

3.4.7 *Data mining*

Buscar por padrões de interesse em uma forma representacional específica ou um conjunto de tais representações, incluindo regras de classificação, regressão e agrupamento. O usuário final pode ajudar significativamente executando corretamente as etapas anteriores

3.4.8 Interpretação de resultados

Interpretar os padrões minerados, possivelmente retornando às etapas anteriores (3.4.1 a 3.4.7 para iterações adicionais. Esta etapa também pode envolver visualização dos padrões e modelos extraídos ou ainda visualização de dados fornecida pelos modelos extraídos

3.4.9 Ação sobre o conhecimento descoberto

Usando o conhecimento diretamente, incorporá-lo em outro sistema para ação extra, ou simplesmente documentá-lo e relatá-lo para as partes interessadas. Esta etapa também inclui verificar e resolver possíveis conflitos com conhecimento previamente suposto (ou extraído)

Este modelo será testado com base no repositório de dados real recolhido, i.e., com base nos dados do repositório recolhido para *KDD*. Será analisada sua eficácia e os pesos ajustados de modo a ser possível obter o máximo de acerto dos dados.

3.5 Aspectos Éticos

No desenvolvimento deste trabalho foram cumpridas todas as diretrizes e normas regulamentadoras descritas na resolução 466/12 do Conselho Nacional de Saúde (CNS) e suas

complementares. Vale ressaltar que a metodologia proposta não envolve experimentação em seres humanos, nem tem foco em grupos especiais, e que a utilização das informações foi feita de forma sigilosa e com a finalidade proposta neste trabalho.

Para tal, o projeto deste trabalho foi submetido ao Comitê de Ética em Pesquisa (CEP) (Anexo I) e obteve aprovação da dispensa do Termo de Consentimento Livre e Esclarecido (TCLE) em 03 de outubro de 2018/Parecer núm. 2.934.452. Visando a manutenção do sigilo das informações, nenhum paciente foi identificado no estudo.

4. RESULTADOS

4.1 Revisão sobre revisões sistemáticas em modelos de reinternação para todas as causas

Iniciamos toda a pesquisa a partir da revisão minuciosa de publicações sobre experiências no uso de indicadores. Isto foi feito de acordo com o guia de revisões sistemáticas proposto por Butler, Hall e Copnell (2016) e pelo *checklist* disponibilizado pelo PRISMA-*Preferred Reporting Items for Systematic Reviews and Meta-Analyses* (MOHER *et al.*, 2009).

Através do processo de triagem da pesquisa bibliográfica (Figura 2), 121 citações foram encontradas. Destas, 89 foram excluídas porque as informações contidas em título e resumo não apresentam uma relação direta com revisão de reinternação hospitalar. Sobre as citações excedentes (n=32), 31 foram excluídas por motivos distribuídos entre (a) utiliza amostras de diagnósticos específicos ou o objetivo do estudo não avalia taxa de risco de reinternação (n=26), (b) utiliza amostras de reinternações planejadas ou não faz uma referência direta à reinternações não-planejadas (n=4) e (c) não desenvolve ou valida um modelo de predição de reinternações hospitalares (n=1).

Figura 2 - Processo de triagem.

Fonte: O autor.

Assim, o resultado da pesquisa bibliográfica, conforme processo de triagem citado anteriormente, foi uma revisão sistemática que extraiu informações de vários modelos de predição desenvolvidos ou utilizados entre 2011 e 2015.

A extração dos modelos de predição encontrados na revisão sistemática (ZHOU *et al.*, 2016) foi filtrada considerando todas as causas de internação em períodos retroativos, onde o valor da medida de desempenho nos coortes de validação foi truncado em duas casas decimais para facilitar a visualização. O resultado da filtragem está sumarizado na Tabela 1, onde nota-se que as variáveis significativas mais utilizadas são: comorbidades (n=13), duração da estadia (n=13), número de reinternações anteriores (n=10), número de visitas em emergência (n=9), tipo de índice de internação (n=9), dados sociodemográficos (n=8) e diagnóstico de internação (n=8).

Nestes modelos, baseados em técnicas de análise estatística, as variáveis clínicas e sociodemográficas foram analisadas através da técnica de regressão logística, sendo que a habilidade discriminativa do modelo foi avaliada pela *Receiver Operating Characteristic (ROC)*.

Tabela 1 - Modelos de predição de reinternação utilizados entre 2011 e 2015.

Referência	Nome do modelo	Amostra (pacientes)	Taxa de reinternação (%)	Medida de desempenho ROC (validação)	Comorbidades	Duração da estadia	# de internações anteriores	# de visitas em emergência	Tipo de índice de internação	Dados demográficos/sociais	Diagnóstico de internação
(ESCOBAR <i>et al.</i> , 2015)	ED30	360.036 em 21 hospitais	12,5	0,73	✓	✓	✓				
	Discharge 30			0,75	✓	✓	✓				
	LACE (validação)			0,72	✓	✓		✓	✓		
(YU <i>et al.</i> , 2015)	(específico)	74.746 em 3 hospitais	18 a 23	0,7	✓	✓	✓	✓	✓	✓	✓
	LACE (validação)			0,55	✓	✓		✓	✓		
(BAILLIE <i>et al.</i> , 2013)	Prediction model	120.396 em 3 hospitais	14,4	0,62			✓				
(CHOUDHRY <i>et al.</i> , 2013)	ACC Admission Model	126.479 em 8 hospitais	7,25	0,76	✓	✓	✓			✓	

Continua

Conclusão

Referência	Nome do modelo	Amostra (pacientes)	Taxa de reinternação (%)	Medida de desempenho ROC (validação)	Comorbidades	Duração da estadia	# de internações anteriores	# de visitas em emergência	Tipo de índice de internação	Dados demográficos/sociais	Diagnóstico de internação
(GILDERSLEEVE; COOPER, 2013)	Risk of readmission score	8.700	14,8	0,7	✓	✓	✓	✓	✓	✓	
(WAKEFIELD; MEHR, 2013)	(sem nome)	463.351 em 91 hospitais	9,7	0,65		✓	✓				
(RICHMOND, 2013)	(sem nome)	4.717	14,4	0,6	✓		✓			✓	✓
(SHULAN; GAO; MOORE, 2013)	(sem nome)	8.718	16,2	0,7	✓	✓				✓	
(WALRAVEN <i>et al.</i> , 2013)	LACE+ (validação)	499.996	11,8	0,73	✓	✓	✓	✓	✓	✓	✓
(COTTER <i>et al.</i> , 2012)	LACE (validação)	507	17,8	0,55	✓	✓		✓	✓		
	Regression Model	502	14,8	0,57				✓			
(KHAN <i>et al.</i> , 2012)	Rehospitalisation risk score	227	15	22,30/92,33 (PPV/NPV)	✓		✓			✓	✓
(LEE, 2012)	(sem nome)	11.951	28,9	(não informado)		✓			✓		✓
(VAN WALRAVEN; WONG; FORSTER, 2012a)	CMG	200.000 em 3.277.033	6,8	0,63						✓	
	LACE (validação)			0,72							
	Combined CMG score and LACE			0,74	✓	✓		✓	✓		✓
(VAN WALRAVEN; WONG; FORSTER, 2012b)	LACE+ (validação)	500.000 em 2.277.033	14	0,74							
	LACE+ with CMG score			0,75	✓	✓	✓	✓	✓	✓	✓

Fonte: Adaptado de Zhou *et al.* (2016).

4.2 Estudos sobre modelos de reinternação em saúde mental

A Tabela 2 apresenta uma comparação entre algumas variáveis disponíveis nas fontes de dados do SIHSUS e SISAM, juntamente com algumas variáveis relevantes utilizadas em modelos de predição de reinternação em saúde mental encontradas em três estudos

(FIGUEROA; HARMAN; ENGBERG, 2004; HEGGESTAD, 2001; VIGOD *et al.*, 2015) dentro do conjunto de citações recuperados pela pesquisa independente (n=288).

Tabela 2 - Variáveis utilizadas em três estudos e variáveis disponíveis nas fontes SIHSUS e SISAM.

Variável	Autor/Sistema				
	(FIGUEROA; HARMAN; ENGBERG, 2004)	(HEGGESTAD, 2001)	(VIGOD <i>et al.</i> , 2015)	SIHSUS	SISAM
Idade	✓	✓	✓	✓	✓
Sexo	✓	✓	✓	✓	✓
Grupo de diagnóstico (CID)	✓	✓	✓	✓	✓
Estado civil		✓	✓		✓
Duração da estadia	✓		✓	✓	✓
Situação de emprego			✓	✓	✓
Área metropolitana	✓		✓		
Grau de escolaridade			✓	✓	
Situação de moradia			✓		
Tipo de plano	✓				
Tempo de hospitalização	✓				

Área metropolitana - moradia localizada em zona urbana ou zona rural

Tipo de plano - utilização de plano de saúde na modalidade gerenciado ou modalidade taxa por serviço

Fonte: O autor.

4.3 Repositório de dados

Com relação à fonte de dados do SIHSUS, os dados brutos são disponibilizados através de arquivos no formato *dbc* (*dBase III Plus compressed*), que não permitem uma importação em base local, assim estes arquivos foram expandidos para o formato *dbf* (*dBase III Plus*) (utilizando a ferramenta *dbf2dbc*, disponível no pacote de distribuição do tabulador denominado TabWin32 4.1.2²) e, por fim, convertidos no formato *csv* (*Comma-separated values*).

Os dados recuperados do SIHSUS agregam todas as informações disponibilizadas pelo DATASUS no período fiscal entre janeiro/2011 e dezembro/2017 dentro do estado de São Paulo, conforme definição dos campos adaptada (Anexo II) fornecida pelo DATASUS.

No caso da fonte de dados do SISAM, os dados brutos foram fornecidos pelo LIS sob a forma de um arquivo no formato *csv*, que compreende todas as informações das internações e

² TabWin32 - <http://datasus.saude.gov.br/download-do-tabwin>

reinternações realizadas nos postos da DRS-XIII no período de janeiro/2011 a dezembro/2017, conforme definição dos campos adaptada (Anexo III).

Desta forma, tanto os dados do SIHSUS, quanto do SISAM foram importados para uma base de dados local (repositório) para que todo o tratamento, desde a filtragem até os cálculos de incidências e probabilidades fossem reavaliados da forma mais eficiente e automática possível. Nesse caso, o Sistema de Gerenciamento de Banco de Dados (SGBD) adotado foi o PostgreSQL 10³. Por estar a tanto tempo ativo, sua arquitetura é considerada confiável e tem como característica manter a integralidade dos dados. Este SGBD também é compatível com a maioria dos sistemas operacionais conhecidos e possui suporte a chaves estrangeiras, junções, visões, gatilhos e procedimentos, características estas que garantem a qualidade deste software que ainda possui funcionalidades sofisticadas como: controle de concorrência multiversionado, recuperação em um ponto de tempo, replicação assíncrona, transações agrupadas, cópias de segurança, planejador de consultas e registrador de transações sequencial para tolerância a falhas.

4.4 Análise dos dados

Com relação aos dados do SIHSUS, iniciou-se uma primeira extração a partir da seleção por ano e mês fiscal, considerando o período janeiro/2011 a dezembro de 2017 (Tabela 3 - Mês/Ano fiscal). Entretanto, notou-se que parte da amostra de internações (n=541.796) correspondia a internações realizadas antes do período inicial.

Uma vez que foram encontradas internações fora do período proposto, realizou-se uma segunda extração, onde o critério de seleção foi composto por data da internação e data de alta dentro do período, considerando ainda internações que não são de longa duração. O resultado final (Tabela 3 - Data da internação e data da alta) compreende as internações com códigos de diagnóstico relacionado ao (ORGANIZAÇÃO MUNDIAL DA SAÚDE - OMS, 1997) (CID) de transtornos mentais e comportamentais (F00 - F99) (n=450.039).

Tabela 3 - Total de internações de acordo com critérios de seleção de registros.

Descrição	Ano/Mês fiscal	Data da internação e data da alta
Internações (todas)	17.430.387	16.474.096
Internações (F00-F99)	1.043.028	450.039

Fonte: O autor.

³ PostgreSQL - <https://www.postgresql.org>

Já com relação aos dados referentes às internações registradas no SISAM (n=8.760), primeiramente desconsiderou-se as internações anteriores ao segundo semestre de 2012, uma vez que estes dados poderiam sofrer certa deficiência em relação à uniformidade. Em um segundo momento, realizou-se uma extração de registros cujo critério de filtragem é composto por internações, cujo CID está relacionado à transtornos mentais e comportamentais (F00 - F99), realizadas entre julho/2012 a dezembro/2017 nos hospitais que fazem parte do DRS-XIII (n=6.898). Adicionalmente, o conjunto de dados resultante da última extração de registros do SISAM é composto por homens e mulheres acima dos 18 anos.

A incorporação dos conjuntos de dados resultantes (SIHSUS e SISAM) em um repositório de dados local, está em conformidade com a etapa de criação de um conjunto de dados alvo sugerido pelo processo *KDD*.

Seguindo para a próxima etapa do *KDD*, que trata da limpeza de dados e pré-processamento, os dados foram submetidos à algumas estratégias para padronização, tratamento de valores nulos ou ausentes e também identificação e remoção de ruídos, conforme descrito a seguir.

4.5 Limpeza dos dados

Uma das atividades mais importantes nesta etapa é a padronização dos dados, que tem por finalidade manter o mesmo formato das variáveis nas fontes de dados do repositório, propiciando assim, um futuro relacionamento entre elas (por exemplo: manter o mesmo formato do campo data e a mesma quantidade de caracteres), também tem por objetivo selecionar apenas as variáveis que serão mais relevantes para as próximas etapas do *KDD*.

No processo de padronização dos dados do SIHSUS e do SISAM, todos os campos foram convertidos para o tipo literal, devido ao fato de que este tipo é a forma mais simples de se padronizar os dados. Para o caso de campos cujos valores originais eram do tipo real, o sinal separador de decimal foi mantido, já quando o tipo de dados do campo era data, foi mantido o padrão *aaaammdd*.

Nas próximas etapas do processo *KDD*, serão analisados quais destes campos literais deverão ser convertidos para outros tipos de dados afim de participar dos algoritmos de *data mining*.

Uma vez que a padronização foi realizada nas fontes de dados do repositório, iniciou-se uma auditoria de dados, cuja intenção foi verificar a qualidade dos mesmos, procurando por discrepâncias, eventuais subnotificações e erros de preenchimento.

O resultado da auditoria apresenta-se através da incompletude, que corresponde ao grau de variáveis nas fontes de dados que apresentam valores ausentes. Neste sentido, e de acordo com a comparação com outros estudos (Tabela 2), as variáveis da fonte de dados do SIHSUS foram submetidas a uma verificação de preenchimento de valores considerando todos os registros do conjunto (n=450.039) e classificadas segundo uma escala de categorias (Tabela 4), os resultados demonstram que a maioria delas (n=4) está presente na categoria “Total”, ou seja, contém valores em todos os registros do conjunto. O restante (n=2) está na escala denominada “Inexistente”, o que corresponde à ausência de valores em todos os registros do conjunto. A Figura 3 apresenta as variáveis e também sua escala de preenchimento.

Tabela 4 - Proporção de preenchimento de variáveis no SIHSUS.

Categoria	Faixa (%)	Variáveis (n)
Total	100	4
Inexistente	0	2
Total		6

Fonte: O autor.

Figura 3 - Proporção de preenchimento de variáveis do SIHSUS.

Fonte: O autor.

As variáveis da fonte de dados do SISAM, de acordo com comparação com outros estudos (Tabela 2), também foram submetidas a uma verificação de preenchimento de valores considerando todos os registros do conjunto (n=6.898) e classificadas segundo uma escala de

categorias (Tabela 5), os resultados demonstram principalmente que metade delas (n=3) está presente na escala denominada como “Total”, ou seja, contém valores em todos os registros do conjunto; já a parte restante (n=3) está na escala denominada “Parcial”, o que corresponde ao preenchimento de valores entre 47% e 86% em todos os registros do conjunto. Adicionalmente, a Figura 4 demonstra as variáveis e sua escala de preenchimento.

Tabela 5 - Proporção de preenchimento de variáveis no SISAM.

Categoria	Faixa (%)	Variáveis (n)
Total	100	3
Parcial	47 - 86	3
Inexistente	0	0
Total		6

Fonte: O autor.

Figura 4 – Proporção de preenchimento de variáveis no SISAM.

Fonte: O autor.

A verificação de inconsistência de dados também pode ser utilizada na auditoria de dados e corresponde ao grau de registros que apresentam valores inválidos, incorretos, redundantes ou ausentes. Para o caso do SIHSUS, verificamos a inconsistência através do grau de incompletude de registros, representado por categorias relacionadas à quantidade de variáveis significativas (Figura 3) com valores ausentes que cada registro possui (Tabela 6, Figura 5), os resultados demonstram que todos os registros contêm duas variáveis com valores ausentes.

Tabela 6 - Grau de incompletude nos registros do SIHSUS.

Categoria	Variáveis (n)	Registros (n)
Inexistente	0	0
Baixo	1 a 2	450.039
Médio	3 a 4	0
Alto	Acima de 4	0
Total		450.039

Fonte: O autor.

Figura 5 - Distribuição de quantidade de variáveis com valores ausentes por número de registros no SIHSUS.

Fonte: O autor.

Para a fonte de dados do SISAM utilizou-se a mesma técnica para verificação de inconsistência, ou seja, representou-se o grau de incompletude de registros através de categorias relacionadas à quantidade de variáveis significativas (Figura 4) com valores ausentes que cada registro possui (Tabela 7), os resultados demonstram principalmente que em cerca de 54% dos registros ($n=3.790$), existem até duas variáveis com valores ausentes, a distribuição da quantidade de variáveis com valores ausentes entre os registros também pode ser analisada pela Figura 6.

Tabela 7 - Grau de incompletude nos registros do SISAM.

Categoria	Variáveis (n)	Registros (n)
Inexistente	0	2.715
Baixo	1 a 2	3.790
Médio	3 a 4	393
Alto	Acima de 4	0
Total		6.898

Fonte: O autor.

Figura 6 - Distribuição de quantidade de variáveis com valores ausentes por número de registros no SISAM.

Fonte: O autor.

4.6 Comparação entre variáveis

Através da Tabela 8 percebe-se que as mesmas variáveis sociodemográficas que estão presentes em pelo menos três dos modelos de predição de reinternações encontrados na Tabela 1, também estão presentes nas duas fontes de dados do repositório construído neste trabalho (SIHSUS e SISAM).

Tabela 8 - Variáveis utilizadas nos estudos e fontes de dados do repositório.

Variável	Autor/Modelo			SIHSUS	SISAM
	(YU <i>et al.</i> , 2015)/(Especí- fico)	(CHOUDHR Y <i>et al.</i> , 2013)/ACC Admission Model	(SHULAN; GAO; MOORE, 2013)/(sem nome)		
Idade	✓	✓	✓	✓	✓
Sexo	✓	✓	✓	✓	✓
Raça	✓	✓	✓	✓	✓
Estado civil	✓	✓	✓		✓
Situação de emprego		✓		✓	✓
Nr. de dependentes			✓	✓	

Fonte: O autor.

A Tabela 9 apresenta a distribuição das internações da fonte de dados do SISAM considerando agrupamentos de faixas de diagnóstico principal. Nota-se que cerca de 85% do total geral de internações é referente às doenças relacionadas ao uso de substância psicoativa ($\approx 42\%$), esquizofrenia ($\approx 22\%$) e transtornos do humor ($\approx 22\%$) representados pelas faixas de CID F10-F19, F20-F29 e F30-F39 respectivamente. A distribuição geral destas internações ($n=6.898$) também pode ser vista na Figura 7, que agrupa o restante das faixas de CID na categoria denominada Outros ($\approx 14,5\%$).

A Figura 8 mostra a evolução semestral da quantidade de internações gerais referentes às faixas de CID identificadas pela Figura 7. Através da Tabela 9, destaca-se ainda, que as internações que não sofreram reinternação após a alta ($n=2.489$) representam aproximadamente 36% do total da fonte de dados, enquanto as reinternações em até 30 dias representam aproximadamente 10% ($n=723$), seguida pelas reinternações em 60 dias, representando 3% ($n=244$).

A Figura 9 mostra a tendência semestral dos tipos de reinternação distribuída em intervalo de dias a partir da data da alta anterior (de 0 a 30, de 31 a 60, de 61 a 90 e também de 91 a 120).

Figura 7 - Proporção das internações gerais no SISAM agrupadas por faixas de diagnóstico (CID).

Fonte: O autor.

Figura 8 - Evolução semestral da quantidade de internações gerais no SISAM por faixas de diagnóstico (CID).

Fonte: O autor.

Figura 9 - Tendência semestral das reinternações no SISAM por intervalo de dias.

Fonte: O autor.

Por fim, foram observadas características referentes ao total de pacientes que não sofreram reinternações no SISAM (n=2.489) (Tabela 10), a saber:

- as internações cujos pacientes encontram-se em meia-idade (40 aos 55 anos) estão concentradas na faixa cujos diagnósticos referem-se à transtornos mentais orgânicos (CID F00-F09) com média de 54 anos, seguido por transtornos neuróticos (CID F40-F48) com média de 45 anos e finalmente transtornos do humor e esquizofrenia (CID F30-F39 e F20-F29) ambos médias de 43 e 42 anos, respectivamente;
- os homens representam a maioria das internações (n=1.512, ≈60%), sendo internados com diagnósticos principais referentes ao uso de substância psicoativa (n=905, ≈60%), seguido por esquizofrenia (n=277, ≈18%), faixas de CID F10-F19 e F20-F29, respectivamente;
- das mulheres internadas (n=977), a maior parte (n=327, ≈33%) é caracterizada com o diagnóstico principal referente à transtornos do humor, seguido pelo uso de substância psicoativa (n=255, 26%), grupos de diagnósticos entre as faixas de CID F30-F39 e F10-F19 respectivamente;
- em cerca de 14% das internações (n=360) há registro que o paciente possui uma ocupação (ofício ou profissão);

- e) em boa parte das internações ($\approx 31\%$), os pacientes são declarados como solteiros;
- f) a duração média da estadia é de 35 dias no grupo que contém a maior frequência de internações (transtornos referentes ao uso de substância psicoativa – CID F10-F19).

Os achados, considerando as características apresentadas sobre pacientes que não sofreram reinternação (Tabela 9, Tabela 10), foram comparados com a literatura (Tabela 2):

- i. As internações cujo diagnóstico principal está relacionado à distúrbios psicóticos (faixa de CID F20-F29) tem 58 dias de duração média de estadia, sendo a maior encontrada entre os grupos, esta informação está em consonância com o estudo de Figueroa, Harman e Engberg (2004);
- ii. Heggstad (2001) identificou que a idade média nas internações era de 41 anos, sendo 55% mulheres e 33% casados, estas informações são parcialmente semelhantes ao que se verifica no SISAM. Uma média de idade aproximada (42 anos) foi encontrada nas internações cujo diagnóstico principal está relacionado à esquizofrenia (faixa de CID F20-F29). Entretanto, identificamos que a maioria dos pacientes são homens ($\approx 60\%$) e que o estado civil predominante é solteiro ($\approx 31\%$);
- iii. O estudo de Vigod *et al.* (2015) identificou que mais da metade das suas internações representam diagnósticos de transtorno depressivo e psicótico (presentes nas faixas de CID F20-F29 e F30-F39, respectivamente), enquanto que no SISAM a mesma faixa representa aproximadamente 43% das internações; variáveis como gênero, profissão/ocupação e idade também diferem na comparação entre o estudo e os dados do sistema regional, pois enquanto o SISAM considera aproximadamente 60% de homens, 14% empregado e idade média de 36 anos, Vigod *et al.* (2015) identificou mais de 50% de mulheres, 26% empregado e idade média de 41 anos.

Tabela 9 - Internações e reinternações do SISAM agrupadas por faixa de diagnósticos.

Grupo de diagnóstico	Faixa de CID	Internações		Não reinternados		Reinternados 30d		Reinternados 60d		Outros	
		n	f%	n	f%	n	f%	n	f%	n	f%
Transtornos mentais orgânicos, inclusive os sintomáticos	F00-F09	200	2,9	81	3,3	34	4,7	7	2,3	78	0,0
Transtornos mentais e comportamentais devidos ao uso de substância psicoativa	F10-F19	2875	41,7	1160	46,6	179	24,8	89	42,0	1447	42,0
Esquizofrenia, transtornos esquizotípicos e transtornos delirantes	F20-F29	1492	21,6	465	18,7	204	28,2	56	22,3	767	22,3
Transtornos do humor (afetivos)	F30-F39	1530	22,2	516	20,7	212	29,3	52	21,8	750	21,8
Transtornos neuróticos, transtornos relacionados com o estresse e transtornos somatoformes	F40-F48	69	1,0	33	1,3	8	1,1	3	0,7	25	0,7
Síndromes comportamentais associadas a disfunções fisiológicas e a fatores físicos	F50-F59	4	0,1	1	0,0	0	0,0	0	0,1	3	0,1
Distorções da personalidade e do comportamento adulto	F60-F69	469	6,8	165	6,6	54	7,5	24	6,6	226	6,6
Retardo mental	F70-F79	239	3,5	62	2,5	29	4,0	12	4,0	136	4,0
Transtornos do desenvolvimento psicológico	F80-F89	7	0,1	2	0,1	1	0,1	1	0,1	3	0,1
Transtornos do comportamento e transtornos emocionais que aparecem habitualmente durante a infância ou a adolescência	F90-F98	7	0,1	3	0,1	1	0,1	0	0,1	3	0,1
Transtorno mental não especificado	F99	6	0,1	1	0,0	1	0,1	0	0,1	4	0,1
		6898		2489	36,1	723	10,5	244	3,5	3442	49,9

Fonte: O autor.

Tabela 10 - Características dos pacientes que não sofreram reinternação no SISAM agrupadas por faixa de diagnósticos (n=2.489).

Descrição do grupo	Faixa de CID	Idade	Homens		Mulheres		Solteiro		Casado		Empregado		Duração da estadia (dias)
		$\bar{x}$	n	f%	n	f%	n	f%	n	f%	n	f%	$\bar{x}$
Transtornos mentais orgânicos, inclusive os sintomáticos	F00 a F09	54	45	3,0	36	3,7	26	3,3	14	4,0	8	2,2	35
Transtornos mentais e comportamentais devidos ao uso de substância psicoativa	F10 a F19	36	905	59,9	255	26,1	378	48,0	126	35,7	180	50,0	35
Esquizofrenia, transtornos esquizotípicos e transtornos delirantes	F20 a F29	42	277	18,3	188	19,2	185	23,5	51	14,4	42	11,7	58
Transtornos do humor (afetivos)	F30 a F39	43	189	12,5	327	33,5	112	14,2	112	31,7	81	22,5	29
Transtornos neuróticos, transtornos relacionados com o estresse e transtornos somatoformes	F40 a F48	45	9	0,6	24	2,5	10	1,3	9	2,5	7	1,9	27
Síndromes comportamentais associadas a disfunções fisiológicas e a fatores físicos	F50 a F59	23	0	0,0	1	0,1	0	0,0	1	0,3	0	0,0	8
Distorções da personalidade e do comportamento adulto	F60 a F69	38	41	2,7	124	12,7	35	4,4	38	10,8	39	10,8	19
Retardo mental	F70 a F79	37	43	2,8	19	1,9	37	4,7	2	0,6	2	0,6	36
Transtornos do desenvolvimento psicológico	F80 a F89	25	1	0,1	1	0,1	2	0,3	0	0,0	1	0,3	50
Transtornos do comportamento e transtornos emocionais que aparecem habitualmente durante a infância ou a adolescência	F90 a F98	20	2	0,1	1	0,1	2	0,3	0	0,0	0	0,0	17
Transtorno mental não especificado	F99	36	0	0,0	1	0,1	0	0,0	0	0,0	0	0,0	13
		36	1512	60,7	977	39,3	787	31,6	353	14,2	360	14,5	29,6

Fonte: O autor.

5. DISCUSSÃO

5.1 Modelos preditivos

Um modelo preditivo (ou método de *data mining*, conforme processo *KDD*) pode ser definido como um algoritmo de aprendizagem de máquina do tipo supervisionado, cuja intenção é construir um estimador (ou função matemática) que tem a capacidade de prever um valor de saída de um novo objeto considerando um conjunto de outros objetos de treinamento previamente rotulados.

Quando o valor de saída deste estimador é do tipo contínuo, como é o objetivo deste estudo ao identificar uma taxa de reinternação no futuro, o tipo de tarefa de aprendizagem é uma regressão; neste cenário, os algoritmos de aprendizagem mais comuns são os estatísticos, tais como Regressão linear, Regressão logística, Regressão de Cox, entre outros.

Também é necessário avaliar a capacidade discriminativa do modelo, algumas das métricas utilizadas para preditores são Matriz de confusão, Medida-f, *Cross-validation*, *ROC*, Área sob a curva *ROC* (*AUC*), Raiz Quadrada do Erro Quadrático Médio (*RMSE*), entre outros.

Um modelo preditivo tem capacidade discriminativa regular nos casos onde o valor da medida de avaliação de desempenho é maior que 0,7 (HOSMER, 2000).

Abaixo, um breve resumo dos três estudos sobre reinternações em saúde mental (FIGUEROA; HARMAN; ENGBERG, 2004; HEGGESTAD, 2001; VIGOD *et al.*, 2015) recuperados pela pesquisa independente com foco em seus objetivos, populações e algoritmos de aprendizagem:

Em 2004, Figueroa, Harman e Engberg (2004) analisaram dados demográficos de 5.735 pacientes de saúde mental em instituições hospitalares nos EUA entre os anos 1997 e 1998 para tentar estimar o impacto da duração da estadia em relação à taxa de reinternação hospitalar. Variáveis como idade, sexo, grupo de diagnóstico principal, duração da estadia, tipo de espaço geográfico, tipo de plano e tempo de hospitalização foram submetidas a técnicas de quadrados mínimos e também de variáveis instrumentais, tendo encontrado como resultado que pacientes com menor tempo de estadia tem mais chances de serem reinternados em 30 dias.

Na mesma frente, Heggstad (2001) analisou dados demográficos de 5.520 pacientes de hospitais psiquiátricos noruegueses em 1996 para avaliar a associação entre a alta rotatividade dos pacientes e a probabilidade de uma reinternação em 30 dias. Algumas variáveis como idade, sexo, grupo de diagnóstico principal e estado civil foram submetidas à técnica de regressão de Cox, obtendo como resultado uma associação estatística entre as situações.

Em um estudo sobre modelos preditivos, Vigod *et al.* (2015) analisou dados de 65.789 pacientes de unidades psiquiátricas canadenses entre 2008 e 2011 para tentar encontrar um índice de risco clínico para quantificar a probabilidade de reinternação. Algumas variáveis como sexo, idade, estado civil, grau de escolaridade, situação de emprego, situação de moradia, tipo de espaço geográfico, grupos de diagnóstico principal e duração da estadia a foram submetidas à técnica de regressão logística, obtendo como resultado um modelo cuja capacidade discriminativa do índice foi considerada moderada.

Pelos estudos acima, podemos perceber que o tipo de algoritmo de aprendizagem utilizado nos modelos foi baseado em estatística, mesmo tipo empregado nos modelos de reinternação para todas as causas (Tabela 1).

5.2 Modelos de predição e o contexto da saúde mental no Brasil

Em 2011 foi publicada uma revisão sistemática (KANSAGARA *et al.*, 2011), onde os autores procuraram identificar modelos validados de previsão de risco de reinternação e examinaram seu desempenho. Eles procuraram em MEDLINE, CINAHL e Cochrane além de buscas em listas de referências recuperadas para modelos de previsão testados e validados em pacientes. 7843 citações foram identificadas, das quais 30 estudos (incluindo 26 modelos) corresponderam aos critérios de inclusão.

Os pesquisadores descobriram:

“O resultado mais comum utilizado foi uma reinternação de 30 dias. Apenas um modelo abordou especificamente reinternações evitáveis. Quatorze modelos que contavam com dados administrativos retrospectivos poderiam ser potencialmente utilizados para calcular taxas de reinternação de ajuste de risco para comparação entre hospitais. Destes, 9 foram testados em amplas populações dos Estados Unidos e resultaram em pequena habilidade descritiva. Sete modelos poderiam potencialmente ser usados para identificar pacientes em alto risco para intervenção precoce durante a hospitalização e 5 poderiam ser usados no momento da alta hospitalar. Seis estudos compararam modelos diferentes na mesma população e 2 desses concluíram que variáveis funcionais e sociais melhoraram a discriminação do modelo. Apesar de a maior parte dos modelos incorporar variáveis para comorbidades e uso de serviços médicos anteriores, poucos examinaram as variáveis associadas às condições de saúde gerais, gravidade da doença ou determinantes sociais de saúde”.

Os pesquisadores também concluíram:

“A maior parte dos modelos atuais de previsão de risco de reinternação designados para propósitos comparativos ou clínicos teve resultados ruins. Apesar de em certos ambientes

tais modelos terem sido úteis, esforços para melhorar seu desempenho são necessários conforme o uso se torna mais generalizado”.

No âmbito da Saúde Mental, Machado (2012) fez um estudo detalhado para analisar o fenômeno da reinternação psiquiátrica no contexto da Atenção Psicossocial, a partir da visão dos pacientes reinternados em um hospital psiquiátrico público. A ideia era compreender o processo de reinternação psiquiátrica ao qual o paciente está sujeito; investigar o cenário assistencial e sócio familiar e suas interferências no fenômeno da reinternação psiquiátrica; explorando, junto aos pacientes, a existência de perspectivas que vislumbram saídas às repetidas internações psiquiátricas. Particularmente a autora, com vistas a sintetizar e analisar a produção científica nacional e internacional acerca do fenômeno da reinternação psiquiátrica, no contexto da desinstitucionalização, realizou uma revisão integrativa da literatura publicada em fontes de pesquisa de impacto que detectou deficiência de estudos que investigassem variáveis psicossociais envolvidas na problemática, bem como ausência da perspectiva do paciente sobre o assunto.

5.3 O papel do SISAM

Diante deste cenário, desde 2011 o DRS-XIII vem se utilizando de um sistema de organização e gerenciamento do acesso aos serviços de saúde mental - o SISAM - que permite o monitoramento da rede de assistência dos 26 municípios que compõem a área de abrangência do DRS-XIII. O principal intuito deste sistema é agilizar e qualificar a avaliação de urgência inicial e tornar mais acessíveis as informações do paciente de saúde mental, para que possam ser facilmente compartilhadas nas diferentes entidades de saúde mental. A ideia do sistema é criar uma rede informatizada abrangendo todos os serviços de saúde mental, na qual seja possível monitorar a situação dos pacientes, em tempo real, desde atenção básica até os serviços especializados.

5.4 Limitações deste estudo

Diante do exposto, nosso estudo apresentou limitações referentes aos dados analisados do SISAM: a amostra contém somente pacientes cuja idade é superior a 18 anos e a estatística descritiva não foi realizada em pacientes que sofreram uma reinternação após a alta. Entretanto, os achados deste trabalho ao analisar e comparar as variáveis do SIHSUS e do SISAM com a literatura selecionada, mostram que suas informações podem contribuir para a construção de um modelo para predição de reinternações hospitalares psiquiátricas.

6. CONCLUSÕES

Neste trabalho nós (a) estudamos a utilização de alguns modelos de predição de reinternação de acordo literatura, (b) analisamos e processamos os dados contidos nas duas fontes de dados (SIHSUS e SISAM) em um repositório de acordo com algumas das etapas do processo *KDD*, (c) descobrimos as variáveis mais utilizadas nos estudos e comparamos as mesmas com as fontes de dados no repositório.

Com base nos resultados obtidos, identificamos que existe uma tendência positiva com relação à viabilidade da elaboração de um modelo de predição de reinternações de pacientes de saúde mental utilizando os dados do repositório local, uma vez que as variáveis utilizadas em alguns dos modelos estudados também estão presentes nas fontes de dados do repositório.

7. SUGESTÕES PARA TRABALHOS FUTUROS

Existem algumas limitações neste trabalho que necessitam de ações na intenção de melhorar os resultados obtidos até aqui.

Um ponto seria continuar a revisão bibliográfica a partir dos resultados obtidos com a inclusão de outras bases de dados online e retificação da *string* de pesquisa.

Além disso, seria importante agregar outras fontes de dados provenientes dos hospitais na região de atendimento do DRS-XIII de forma a enriquecer o repositório de dados local, e assim, prosseguir nas próximas etapas do processo *KDD*.

REFERÊNCIAS

- BAILLIE, C. A. et al. The Readmission risk flag: Using the electronic health record to automatically identify patients at risk for 30-day readmission. **Journal of Hospital Medicine**, v. 8, n. 12, p. 689–695, 2013.
- BOSCO, J. A. et al. Cost burden of 30-day readmissions following medicare total hip and knee arthroplasty. **Journal of Arthroplasty**, v. 29, n. 5, p. 903–905, 2014.
- BUTLER, A.; HALL, H.; COPNELL, B. A Guide to writing a qualitative systematic review protocol to enhance evidence-based practice in nursing and health care. **Worldviews on Evidence-Based Nursing**, v. 13, n. 3, p. 241–249, 2016.
- CHOUDHRY, S. A. et al. A Public-private partnership develops and externally validates a 30-day hospital readmission risk prediction model. **Online Journal of Public Health Informatics**, v. 5, n. 2, p. 1–17, 2013.
- COTTER, P. E. et al. Predicting readmissions: poor performance of the LACE index in an older UK population. **Age and Ageing**, v. 41, n. 6, p. 784–789, 2012.
- DATASUS. **DATASUS** - Ministério da Saúde. 2009. Disponível em: <<http://datasus.saude.gov.br/>>. Acesso em: 27 jan. 2020.
- ESCOBAR, G. J. et al. Nonelective rehospitalizations and postdischarge mortality: predictive models suitable for use in real time. **Medical Care**, v. 53, n. 11, p. 916, 2015.
- FAYYAD, U.; PIATETSKY-SHAPIO, G.; SMYTH, P. From data mining to knowledge discovery in databases. **AI Magazine**, v. 17, n. 3, p. 37, 1996.
- FIGUEROA, R.; HARMAN, J.; ENGBERG, J. Use of claims data to examine the impact of length of inpatient psychiatric stay on readmission rate. **Psychiatric Services**, v. 55, n. 5, p. 560–565, 2004.
- GILDERSLEEVE, R.; COOPER, P. Development of an automated, real time surveillance tool for predicting readmissions at a community hospital. **Applied Clinical Informatics**, v. 4, n. 2, p. 153–169, 2013.
- HEGGESTAD, T. Operating conditions of psychiatric hospitals and early readmission: effects of high patient turnover. **Acta Psychiatrica Scandinavica**, v. 103, n. 3, p. 196–202, Mar. 2001.
- HOSMER, D. W. Exact methods for logistic regression models. In: HOSMER, D. W.; LEMESHOW, S. **Applied logistic regression**. New York: Wiley, 2000.
- JOYNT, K. E.; JHA, A. K. Thirty-day readmissions — truth and consequences. **New England Journal of Medicine**, v. 366, n. 15, p. 1366–1369, 2012.
- KANSAGARA, D. et al. Risk prediction models for hospital readmission: a systematic review. **JAMA - Journal of the American Medical Association**, v. 306, n. 15, p. 1688–1698, 2011.

KHAN, A. et al. An Electronic medical record-derived real-time assessment scale for hospital readmission in the elderly. **Wisconsin Medical Journal**, v. 111, n. 3, p. 119–122, 2012.

LEE, E. W. Selecting the best prediction model for readmission. **Journal of Preventive Medicine and Public Health**, v. 45, n. 4, p. 259–266, 2012.

MACHADO, V. **Reinternação psiquiátrica no campo da atenção psicossocial: a perspectiva dos pacientes reinternantes**. 2012. Tese (Doutorado) - Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto, Ribeirão Preto, 2012.

MOHER, D. et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. **PLoS Medicine**, v. 6, n. 7, 2009. DOI: 10.1371/journal.pmed.1000097.

ORGANIZAÇÃO MUNDIAL DA SAÚDE. **CID-10** - critérios diagnósticos para pesquisas. Porto Alegre: Artes Médicas Sul, 1997.

RICHMOND, D. M. **Socioeconomic predictors of 30-day hospital readmission of elderly patients with initial discharge destination of home health care**. Birmingham: The University of Alabama at Birmingham, 2013.

SHULAN, M.; GAO, K.; MOORE, C. D. Predicting 30-day all-cause hospital readmissions. **Health Care Management Science**, v. 16, n. 2, p. 167–175, 2013.

VAN WALRAVEN, C. et al. Predicting post-discharge death or readmission: deterioration of model performance in population having multiple admissions per patient. **Journal of Evaluation in Clinical Practice**, v. 19, n. 6, p. 1012–1018, 2013.

VAN WALRAVEN, C.; WONG, J.; FORSTER, A. J. Derivation and validation of a diagnostic score based on case-mix groups to predict 30-day death or urgent readmission. **Open Medicine**, v. 6, n. 3, 2012a. DOI: PMC3654506.

VAN WALRAVEN, C.; WONG, J.; FORSTER, A. J. LACE+ index: extension of a validated index to predict early death or urgent readmission after hospital discharge using administrative data. **Open Medicine**, v. 6, n. 3, 2012b. DOI: PMC3659212.

VIGOD, S. N. et al. READMIT: a clinical risk index to predict 30-day readmission after discharge from acute psychiatric units. **Journal of Psychiatric Research**, v. 61, p. 205–213, 2015.

WAKEFIELD, D. S.; MEHR, D. R. Risk factors for all-cause hospital readmission within 30 days of hospital discharge. **JCOM**, v. 20, n. 5, p. 203-214, May 2013.

WALJEE, A. K.; HIGGINS, P. D. R.; SINGAL, A. G. A Primer on predictive models. **Clinical and Translational Gastroenterology**, v. 5, n. 12, 2014. DOI: 10.1038/ctg.2013.19.

YOSHIURA, V. T. et al. A Web-based information system for a regional public mental healthcare service network in Brazil. **International Journal of Mental Health Systems**, v. 11, n. 1, p. 1–10, 2017.

YU, S. et al. Predicting readmission risk with institution-specific prediction models. **Artificial Intelligence in Medicine**, v. 65, n. 2, p. 89–96, 2015.

ZHOU, H. et al. Utility of models to predict 28-day or 30-day unplanned hospital readmissions: an updated systematic review. **BMJ Open**, v. 6, n. 6, 2016. DOI: <http://dx.doi.org/https://bmjopen.bmj.com/content/6/6/e011060>.

ANEXO I – PARECER DO COMITÊ DE ÉTICA EM PESQUISA

**FACULDADE DE MEDICINA DE RIBEIRÃO PRETO
UNIVERSIDADE DE SÃO PAULO
COMITÊ DE ÉTICA EM PESQUISA**

Ofício nº 033/18-CEP/CSE-FMRP-USP

Ribeirão Preto 04 de Outubro de 2018.

Prezado Senhor,

Comunicamos que o projeto de pesquisa abaixo especificado foi analisado e **Aprovado "Ad-Referendum"** do Comitê de Ética em pesquisa do Centro de Saúde Escola da Faculdade de Medicina de Ribeirão Preto da Universidade de São Paulo, na data de 03 de Outubro de 2018.

CAAE: 96556518.6.0000.5414

Projeto de pesquisa: "Estudo de Modelos para Predição de Reinternação em Saúde Mental no Sistema Único de Saúde (SUS)".

Pesquisador: Alexandre Henrique Macris

Em atendimento à Resolução 466/12, deverá ser encaminhado a este CEP o relatório final da pesquisa e a publicação de seus resultados, para acompanhamento, bem como comunicada qualquer intercorrência ou a sua interrupção.

Atenciosamente,

Prof. Dr. Laércio Joel Franco
Coordenador do CEP/CSE-FMRP-USP.

Ilmo. Prof. Dr. Domingos Alves
Departamento de Medicina Social da
Faculdade de Medicina de Ribeirão Preto- SP

ANEXO II - DEFINIÇÃO DOS CAMPOS DA AIH REDUZIDA (SIH)

SEQ	Nome do campo	Descrição
1	uf_zi	Município Gestor
2	ano_cmpt	Ano de processamento da AIH, no formato aaaa
3	mes_cmpt	Mês de processamento da AIH, no formato mm
4	espec	Especialidade do Leito
5	cgc_hosp	CNPJ do Estabelecimento.
6	n_aih	Número da AIH
7	ident	Identificação do tipo da AIH.
8	cep	CEP do paciente
9	munic_res	Município de Residência do Paciente
10	nasc	Data de nascimento do paciente (aaaammdd).
11	sexo	Sexo do paciente
12	uti_mes_in	Zerado
13	uti_mes_an	Zerado
14	uti_mes_al	Zerado
15	uti_mes_to	Quantidade de dias de UTI no mês.
16	marca_uti	Indica qual o tipo de UTI utilizada pelo paciente.
17	uti_int_in	Zerado
18	uti_int_an	Zerado
19	uti_int_al	Zerado
20	uti_int_to	Quantidade de diárias em unidade intermediária.
21	diar_acom	Quantidade de diárias de acompanhante.
22	qt_diarias	Quantidade de diárias.
23	proc_solic	Procedimento solicitado.
24	proc_rea	Procedimento realizado.
25	val_sh	Valor de serviços hospitalares.
26	val_sp	Valor de serviços profissionais.
27	val_sadt	Zerado
28	val_rn	Zerado
29	val_acomp	Zerado
30	val_ortp	Zerado
31	val_sangue	Zerado
32	val_sadtsr	Zerado
33	val_transp	Zerado
34	val_obsang	Zerado
35	val_pedlac	Zerado
36	val_tot	Valor total da AIH.
37	val_uti	Valor de UTI.
38	us_tot	Valor total, em dólar.
39	dt_inter	Data de internação no formato aaammdd.
40	dt_saida	Data de saída, no formato aaaammdd.

41	diag_princ	Código do diagnóstico principal (CID10).
42	diag_secun	Código do diagnóstico secundário (CID10). Preenchido com zeros a partir de 201501.
43	cobranca	Motivo de Saída/Permanência
44	natureza	Natureza jurídica do hospital (com conteúdo até maio/12). Era utilizada a classificação de Regime e Natureza.
45	nat_jur	Natureza jurídica do Estabelecimento, conforme a Comissão Nacional de Classificação - CONCLA
46	gestao	Indica o tipo de gestão do hospital.
47	rubrica	Zerado
48	ind_vdrl	Indica exame VDRL
49	munic_mov	Município do Estabelecimento
50	cod_idade	Unidade de medida da idade.
51	idade	Idade.
52	dias_perm	Dias de Permanência.
53	morte	Indica Óbito
54	nacional	Código da nacionalidade do paciente.
55	num_proc	Zerado
56	car_int	Caráter da internação
57	tot_pt_sp	Zerado
58	cpf_aut	Zerado
59	homonimo	Indicador se o paciente da AIH é homônimo do paciente de outra AIH.
60	num_filhos	Número de filhos do paciente.
61	instru	Grau de instrução do paciente.
62	cid_notif	CID de Notificação
63	contracep1	Tipo de contraceptivo utilizado.
64	contracep2	Segundo tipo de contraceptivo utilizado
65	gestrisco	Indicador se é gestante de risco.
66	insc_pn	Número da gestante no pré-natal.
67	seq_aih5	Sequencial de longa permanência (AIH tipo 5)
68	cbor	Ocupação do paciente, segundo a Classificação Brasileira de Ocupações – CBO.
69	cnaer	Código de acidente de trabalho
70	vincprev	Vínculo com a Previdência.
71	gestor_cod	Motivo de autorização da AIH pelo Gestor.
72	gestor_tp	Tipo de gestor.
73	gestor_cpf	Número do CPF do Gestor.
74	gestor_dt	Data da autorização dada pelo Gestor (aaaammdd).
75	cnes	Código CNES do hospital.

76	cnpj_mant	CNPJ da mantenedora.
77	infehosp	Status de infecção hospitalar
78	cid_asso	CID causa
79	cid_morte	CID da morte.
80	complex	Complexidade.
81	financ	Tipo de financiamento.
82	faec_tp	Subtipo de financiamento FAEC.
83	regct	Regra contratual.
84	raca_cor	Raça/Cor do paciente.
85	etnia	Etnia do paciente, se raça cor for indígena.
86	sequencia	Sequencial da AIH na remessa
87	remessa	Número da remessa.
88	aud_just	Justificativa do auditor para aceitação da AIH sem o número do Cartão Nacional de Saúde
89	sis_just	Justificativa do estabelecimento para aceitação da AIH sem o número do Cartão Nacional de Saúde.
90	val_sh_fed	Valor do complemento federal de serviços hospitalares. Está incluído no valor total da AIH.
91	val_sp_fed	Valor do complemento federal de serviços profissionais. Está incluído no valor total da AIH.
92	val_sh_ges	Valor do complemento do gestor (estadual ou municipal) de
93	val_sp_ges	Valor do complemento do gestor (estadual ou municipal) de serviços profissionais. Está incluído no valor total da
94	val_uci	Valor de UCI.
95	marca_uci	Tipo de UCI utilizada pelo paciente.
96	diagsec1	Diagnóstico secundário 1
97	diagsec2	Diagnóstico secundário 2
98	diagsec3	Diagnóstico secundário 3
99	diagsec4	Diagnóstico secundário 4
100	diagsec5	Diagnóstico secundário 5
101	diagsec6	Diagnóstico secundário 6
102	diagsec7	Diagnóstico secundário 7
103	diagsec8	Diagnóstico secundário 8
104	diagsec9	Diagnóstico secundário 9
105	tpdisec1	Tipo de diagnóstico secundário 1
106	tpdisec2	Tipo de diagnóstico secundário 2
107	tpdisec3	Tipo de diagnóstico secundário 3
108	tpdisec4	Tipo de diagnóstico secundário 4
109	tpdisec5	Tipo de diagnóstico secundário 5
110	tpdisec6	Tipo de diagnóstico secundário 6
111	tpdisec7	Tipo de diagnóstico secundário 7
112	tpdisec8	Tipo de diagnóstico secundário 8
113	tpdisec9	Tipo de diagnóstico secundário 9

ANEXO III - DEFINIÇÃO DOS CAMPOS DO SISAM (ADAPTADA)

SEQ	Nome do campo	Descrição
1	data_internacao_real	Data de internação no formato aaaa-mm-dd hh:mm:ss
2	cod_pessoa	Código de paciente
3	tipo_pessoa	PA (fixo)
4	data_cadastro	Data da criação do cadastro formato dd-mm-aaaa hh:mm:ss
5	sexo	Sexo do paciente
6	data_nascimento	Data de nascimento do paciente no format aaaa-mm-dd
7	uf	Unidade federativa
8	municipio	Cidade
9	profissao_ocupacao	Profissão ou ocupação
10	cod_paciente	Código de paciente
11	unidade_cadastro	Unidade de cadastro
12	idade_aprox	Idade aproximada
13	estado_civil	Estado civil
14	etnia	Etnia
15	arranjo_domiciliar	Arranjo domiciliar
16	cod_regulacao	Código de regulação
17	cod_unidade_origem	Código da unidade de origem
18	cod_unidade_regulacao	Código da unidade de regulação
19	data_solicitacao	Data da Solicitação
20	cod_unidade_destino	Código da unidade de destino
21	tipo_regulacao	Tipo de regulação
22	status_reg	Status do registro
23	status_reg_compl	Complemento do status do registro
24	via_direta	Via direta
25	tipo_transferencia	Tipo de transferência
26	cid10_diagnostico	Diagnóstico principal
27	cid10_seg_diagnostico	Diagnóstico secundário
28	data_lista_reg	Data de aceite na lista da UE (Unidade de Emergência do HCFMRP)
29	data_vaga_reg	Data da liberação da vaga na UE (Unidade de Emergência do HCFMRP)
30	data_admissao_reg	Data de admissão na UE
31	data_lista_hospital	Data de aceite na lista do hospital
32	data_vaga_hospital	Data da vaga no hospital
33	data_internacao	Data da internação
34	data_cancel_solicitacao	Data de cancelamento da solicitação
35	data_transferencia	Data da transferência
36	traumatismo	Se possui traumatismo
37	prob_respiratorios	Se possui problemas respiratórios
38	avc	Se possui AVC
39	convulsao	Se possui convulsão
40	has	Se possui hipertensão arterial sistêmica

41	gravidez	Se possui gravidez
42	diabetes	Se possui diabetes
43	doenca_infecto	Se possui doenças infecto-contagiosas
44	tipo_internacao_de	Tipo de internação
45	data_expiracao	Data da expiração
46	cod_cidade_residencia	Código da cidade
47	cod_estado_residencia	Código do estado
48	cod_internacao	Código da internação
49	cod_transf_1	Código da unidade da primeira transferência
50	cod_transf_2	Código da unidade da segunda transferência
51	cod_transf_alta	Código da unidade de transferência na alta
52	status_internacao	Status da internação
53	data_alta	Data da alta
54	cid10_diagnostico_alta	Diagnóstico principal na alta
55	cid10_seg_diagnostico_alta	Diagnóstico secundário na alta
56	tipo_alta	Tipo de alta
57	tipo_internacao	Tipo de internação
58	nome_unidade_solicitacao	Nome da unidade de solicitação
59	nome_unidade_destino	Nome da unidade de destino
60	via_internacao	Via de internação
61	tempo_espera_internacao	Tempo de espera da internação
62	tempo_espera_aceite	Tempo de espera do aceite
63	idade	Idade do paciente
64	grupo_diagnostico_alta	Grupo de diagnóstico na alta
65	faixa_etaria	Faixa etária do paciente
66	tempo_internacao	Tempo de internação