



UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
FACULDADE DE MEDICINA
PROGRAMA DE PÓS-GRADUAÇÃO EM EPIDEMIOLOGIA

DISSERTAÇÃO DE MESTRADO

Modelos para automatização de análises de *desfechos clínico-assistenciais*

Bruna Stella Zanotto

Orientador: Prof. Dra. Carisi Anne Polanczyk

Co-orientador: Prof. Dra. Ana Paula da Silva Beck Etges

Porto Alegre, Brasil

2020



UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
FACULDADE DE MEDICINA
PROGRAMA DE PÓS-GRADUAÇÃO EM EPIDEMIOLOGIA

DISSERTAÇÃO DE MESTRADO

Modelos para automatização de análises de *desfechos clínico-assistenciais*

Bruna Stella Zanotto

Orientador: Prof. Dra. Carisi Anne Polanczyk

A apresentação desta dissertação é exigência do Programa de Pós-graduação em Epidemiologia, Universidade Federal do Rio Grande do Sul, para obtenção do título de Mestre.

Porto Alegre, Brasil.

2020

CIP - Catalogação na Publicação

Zanotto, Bruna Stella

Modelos para automatização de análises de desfechos clínico-assistenciais / Bruna Stella Zanotto. -- 2020. 160 f.

Orientadora: Carisi Anne Polanczyk.

Coorientadora: Ana Paula Etges.

Dissertação (Mestrado) -- Universidade Federal do Rio Grande do Sul, Faculdade de Medicina, Programa de Pós-Graduação em Epidemiologia, Porto Alegre, BR-RS, 2020.

1. Inteligência artificial. 2. Gestão em saúde. 3. prontuário eletrônico. 4. Valor em saúde. 5. Processamento de linguagem natural. I. Polanczyk, Carisi Anne, orient. II. Etges, Ana Paula, coorient. III. Título.

BANCA EXAMINADORA

Prof Dr Sotero Serrate Mengue, Programa de Pós-Graduação em Epidemiologia, Universidade Federal do Rio Grande do Sul

Prof Dr Antonio Luiz Pinho Ribeiro, Programa de Pós-Graduação em Ciências da Saúde, Universidade Federal de Minas Gerais.

Prof Dra Mariana Recamonde Mendonza, Programa de Pós Graduação em Computação, Universidade Federal do Rio Grande do Sul

DEDICATÓRIA ou MENSAGEM

Dedico este trabalho a minha avó Maria, que se manteve, em meio a chamas, forte, de cabeça erguida e, principalmente resiliente. Ela, que diante das viradas de avesso da vida, nos ensina que felicidade e a garra vêm de dentro pra fora, e que nunca nos falte esperança de dias melhores. És um exemplo de mulher e pilar de toda família.

“Brindemos por tudo de bom que realizamos neste ano, esperando que o próximo
concorra para a nossa paz e prosperidade”

Zulmiro Stella

AGRADECIMENTOS

A realização e o alcance dos objetivos desta pesquisa não teriam sido possíveis sem a contribuição direta de muitas pessoas, pessoas as quais costumo chamar de raras, raras pois compartilharam comigo toda forma de conhecimento, incentivo, e amizade neste ano. Aos meus raros eu dedico enorme agradecimento e reconhecimento da importância de cada um nesse processo.

Agradeço as minhas queridas professoras Carisi e Ana Paula. Quero que saibam que reconheço o privilégio de tê-las conhecido, de trabalhar e aprender com vocês. Carisi, obrigada por confiar na minha capacidade me dando total liberdade para conseguir conduzir esse trabalho. Sou grata por todos os ensinamentos compartilhados de forma admirável, por me guiar desde os primeiros passos no mundo da pesquisa, pós-graduação e ter despertado em mim a vontade de lecionar. Tens uma admirável e virtuosa capacidade de liderança e motivação dentro do nosso grupo de pesquisa.

Ana, obrigada por acreditar em mim e pelos tantos incentivos e desafios propostos de enxergar sempre além. Tenho certeza de que não chegaria aqui sem o seu apoio. Você foi e está sendo muito mais que coorientadora: para mim será sempre mestre e amiga.

Agradeço aos pesquisadores do IATS e NATS, colegas que se tornaram amigos, obrigada pela companhia, e torcida sempre. Aos novos colegas que encontrei nesta caminhada, de outras faculdades e estados do Brasil, sou eternamente grata pelos ensinamentos e pelas diferentes perspectivas que pude enxergar neste ano. Em especial ao professor Marcos, Eduardo e Avner, obrigada pela paciência. Ricardo, obrigada pelo “*team work*” e amizade que se fortaleceram este ano.

Agradeço aos meus amigos que acompanharam toda a trajetória até aqui. Em especial, Ingrid, Bárbara, Lari, Luisa, vocês são maravilhosas, sou muito abençoada em ter pessoas tão queridas que me querem tão bem. Obrigada família, de longe e de perto pelas boas vibrações e mensagens de incentivo. Em especial ao meu padrinho Steffan e Flávia, conversas com vocês sempre leves e repletas de aprendizado, vocês são exemplos de profissionais, pais e amigos.

Obrigada Gustavo pela cumplicidade, chegou tão de repente e de maneira tão sincera me mostrou a virtude que é ter alguém do lado, o significado de companheirismo e admiração. Muito feliz em compartilhar este momento contigo.

Meu maior agradecimento a meus pais, Cesar e Maristela, que me viram crescer e

deixaram com que eu tomasse decisões importantes. Obrigada por terem, com muito carinho e apoio, não medido esforços para que eu conquistasse mais esta etapa da minha vida.

Sou o resultado da força e confiança de cada um de vocês.

SUMÁRIO

1. INTRODUÇÃO	11
2. REVISÃO DE LITERATURA	14
2.1. CUIDADOS EM SAÚDE BASEADOS EM VALOR (VBHC).....	14
2.2. DESFECHOS EM SAÚDE	16
2.3. PRONTUÁRIO ELETRÔNICO / <i>ELETRONIC HEALTH RECORDS (EHR)</i>	20
2.4. A INTELIGÊNCIA ARTIFICIAL COMO CATALISADOR PARA MELHORAR O VALOR NOS CUIDADOS DE SAÚDE	23
2.4.1. <i>Aprendizado de máquina clássico / Machine Learning</i>	24
2.4.2. <i>Aprendizado Profundo / Deep Learning</i>	27
2.4.3. <i>Processamento de Linguagem Natural / Natural Language Processing</i>	28
2.5. O PROBLEMA DE CLASSIFICAÇÃO DE TEXTO	28
2.5.1. <i>Representação de texto</i>	29
2.5.2. <i>Classificadores do texto</i>	31
2.5.3. <i>Ontologias como modelo de classificação</i>	36
2.6. AVC.....	37
3. OBJETIVOS	40
4. REFERÊNCIAS BIBLIOGRÁFICAS	41
5. ARTIGO(S).....	49
5.1. ARTIGO 1.....	ERRO! INDICADOR NÃO DEFINIDO.
5.2. ARTIGO 2	ERRO! INDICADOR NÃO DEFINIDO.
6. CONCLUSÕES E CONSIDERAÇÕES FINAIS	50
7. LIMITAÇÕES DO ESTUDO	52
8. TRABALHOS FUTUROS	53
APÊNDICE A - DETALHES METODOLÓGICOS ADICIONAIS	54
ANEXO A. CÓPIA DO PARECER DE APROVAÇÃO DO PROJETO NO COMITÊ DE ÉTICA.....	60
ANEXO B. MANUAL DE TREINAMENTO PARA OS ALGORITMOS.....	64

ABREVIATURAS E SIGLAS

AVC- Acidente Vascular Cerebral

AVCi- Acidente Vascular Cerebral isquêmico

BERT - *Bidirectional Encoder Representations from Transformers*

BoW – Bag of Words

CNN- *Convolutional Neural Networks*

IA – Inteligência Artificial

KNN- *K-Nearest Neighbor*

MF- *Metafeatures*

NBC- *Naïve Bayes Classifier*

NIHSS- National Institute of Health Stroke Scale

PLN- Processamento de Linguagem Natural

PRO – *patient-reported outcome*

PROM- *patient-reported outcome measure*

SVM- *Support Vector Machine*

TDABC- time-driven activity-based costing

TF- *Term-frequency*

TF-IDF- *Term Frequency-Inverse Document Frequency*

TI – Tecnologia de Informação

UPI - Unidade Práticas Integradas

VBHC- Value-based healthcare

VMO- Value Management Office

APRESENTAÇÃO

Este trabalho consiste na dissertação de mestrado intitulada “Proposição de modelo para a automatização de análises de desfechos de saúde” apresentada ao Programa de Pós-Graduação em Epidemiologia da Universidade Federal do Rio Grande do Sul, em Porto Alegre, 11 de dezembro de 2020. O trabalho é apresentado em sete partes, na ordem que segue:

Introdução, Revisão da Literatura e Objetivos

Artigo(s)

Conclusões e Considerações Finais

Limitações do Estudo

Trabalhos Futuros

Apêndices

Anexos

1. INTRODUÇÃO

A fim de compreender e buscar alternativas ao atual cenário de recursos escassos e crescente demanda por qualidade nos serviços de saúde, Porter *et al.* (2010) definiram valor dos cuidados de saúde como resultados de saúde alcançado por cada dólar gasto para alcançar estes resultados (PORTER, 2010). Entre as premissas do modelo de gestão de cuidados de saúde orientados a valor (VBHC), a capacidade tecnológica de medição é fundamental para viabilizar informações que sustentem decisões voltadas à melhoria contínua dos cuidados em saúde (PORTER; LEE, 2013).

Em pesquisa feita pela Medtronic – The Economist, o Brasil posicionou-se atrás de seus vizinhos territoriais nos avanços em prol à adoção de uma gestão de saúde baseada em valor. O domínio mais crítico foi justamente de mensuração de desfechos reportados pelo paciente e de custos (ATEFA SHAH; BLASZCZYNSKI, 2016). Na maioria dos hospitais a mensuração de desfecho ainda é baseada na coleta e relato de desfechos clínicos como mortalidade e complicações intra-hospitalares (KATZ; FRANKEN; MAKDISSE, 2017), mas que não refletem outros aspectos de cuidado que importam para os pacientes.

Na tentativa de traduzir o conceito à prática médica, Lee *et al.* (2010) sugerem ser necessário que as organizações de saúde desenvolvam formas de capturar os dados que permitam avaliar a qualidade da entrega do serviço aos pacientes e os custos reais para prover esses resultados. Estes dados são essenciais para avaliar valor, mas requerem um avanço em capacidade tecnológica para que possam ser monitorados e controlados (TSAI; PORTER; ADAMS, 2018).

Alguns casos de sucesso de avanços tecnológicos para monitoramento de desfechos, como desenvolvido na Holanda (ATEFA SHAH; BLASZCZYNSKI, 2016), demonstram que é possível definir medidas robustas para resultados, coletar informações comparativas e ajustá-las a risco. Esses casos também demonstram que através de uma nova abordagem institucional melhorias substanciais foram alcançadas na capacidade de entrega de valor aos pacientes. Os esforços de medição de resultados existentes, embora imperfeitos, têm reduzido a mortalidade e melhorando outros resultados de saúde, beneficiando a tomada de decisões clínicas e motivando melhorias nos processos assistenciais e de gestão em saúde (PORTER; TEISBERG, 2006).

De acordo com algumas definições, qualquer programa de computador que exiba

um comportamento humano inteligente é uma forma de IA (MEULEN, 2019; RUSSELL, 1995). Alguns autores (FALCONER; NAND; LIOW *et al.*, 2014; GOPAL; SUTER-CRAZZOLARA; TOLDO *et al.*, 2018), alegam o uso de Inteligência Artificial (IA) como parte da entrega do cuidado em saúde, seja através de softwares para priorização de pacientes (VALUCK; BLAISDELL; DUGAN *et al.*, 2017), e para isso algumas técnicas tem se destacado. Alguns autores alegam que o uso de inteligência artificial também é uma forma de viabilizar a adoção da gestão baseada em valor, e para isso algumas técnicas têm se destacado. O processamento de linguagem natural faz uso de técnicas como aprendizado de máquina e aprendizado profundo para reconhecimento de estruturas a nível da semântica. Seu potencial na área de saúde está em conseguir identificar e fornecer um conjunto de informações confiáveis a partir de dados não estruturados, como os que compõem os registros clínicos eletrônicos dos pacientes (MESKÓ; GÖRÖG, 2020).

A agenda de valor apresentar por Porter é composta por seis componentes: unidade práticas integradas (UPI), medida de desfechos e custos para todos os pacientes, pagamento por *bundled* para o ciclo de cuidado, sistema de saúde integrado, alcance geográfico expandido, e desenvolvimento de capacidade tecnológica. Este último, é apresentado de forma transversal a todos os demais, por ser compreendido como essencial para que os outros cinco possam ser implementados. Sistemas de TI consistentes, robustos e integrados podem ajudar as partes de uma UPI a trabalharem juntas, permitir a medição e novas abordagens de reembolso, enlaçar as partes de uma entrega bem estruturada, e permitir alcance geográfico de assistência em pontos remotos (PORTER; LEE, 2013).

Avanços de capacidade computacional e modelos de IA para melhor analisar os dados de saúde têm sido observados como, também, uma forma de viabilizar a adoção de uma gestão baseada em valor para o sistema de saúde (ETGES; RUSCHEL; POLANCZYK *et al.*, 2020). Para isso, algumas técnicas têm se destacado. O processamento de linguagem natural (PLN) faz uso do aprendizado de máquina a fim de envolver a aplicação do conhecimento do léxico, sintático e semântico do idioma (VIEIRA; LIMA, 2001). Seu potencial na área de saúde está em conseguir identificar e fornecer um conjunto de informações confiáveis a partir de dados não estruturados, como os que compõem os registros clínicos eletrônicos dos pacientes (ROSENBLOOM; DENNY; XU *et al.*, 2011).

Por esse motivo, tem-se um investimento na aplicação de técnicas de PLN em

diferentes áreas da saúde, entre elas, saúde mental para avaliação de alterações de humor e risco de suicídio (PESTIAN; NASRALLAH; MATYKIEWICZ *et al.*, 2010; VELUPILLAI; SUOMINEN; LIAKATA *et al.*, 2018), avaliação de preditores e desfechos em grandes amostras da pesquisa clínica (WANG, S. Y.; PERSHING, S.; TRAN, E. *et al.*, 2020), seleção para grupo de pacientes para pesquisa clínica (SILVA, 2018), e identificação de determinantes e comorbidades a nível individual do paciente como uso de cigarro, circunstâncias sociais entre outros (BELL; KILIC; PRABAKARAN *et al.*, 2013; LYNCH; MOORE, 2016).

Na busca por migrar a gestão dos cuidados em saúde para valor, ferramentas que auxiliem na coleta, e análise de dados são avanços fundamentais para proporcionar o monitoramento assertivo e ágil de resultados (ETGES; RUSCHEL; POLANCZYK *et al.*, 2020). Já vem sendo descrito na literatura a lacuna de tecnologia da informação para automatizar a coleta de dados de resultados a partir de registros clínicos eletrônicos (ETGES; RUSCHEL; POLANCZYK *et al.*, 2020; LEE, 2010). A disponibilidade de uma sólida plataforma de tecnologia de informação, que inclui infra-estrutura de *big data*, pode acelerar a adoção de gestão de saúde baseada em valor e ajudar a convencer decisores clínicos e administrativos de que padronização de conjuntos de resultados pode ser alcançado rapidamente para uma crescente gama de condições médicas (KATZ; FRANKEN; MAKDISSE, 2017).

Considerando o cenário brasileiro e a necessidade de se propor novas maneiras de medir desfechos, o tema desta dissertação é o uso de ferramentas tecnológicas informacionais para suportar modelos de avaliação de valor em saúde.

2. REVISÃO DE LITERATURA

Para melhor contextualização dos conceitos abordados neste trabalho, o referencial teórico está dividido em 6 tópicos: cuidados em saúde baseados em valor, desfechos em saúde, prontuários eletrônicos, inteligência artificial como catalisador para melhorar o valor em cuidados em saúde, classificação de texto e, acidente vascular isquêmico (AVC).

2.1. CUIDADOS EM SAÚDE BASEADOS EM VALOR (VBHC)

Historicamente, os serviços e pagamentos de saúde nem sempre estão alinhados com a qualidade dos resultados observados no estado de saúde dos pacientes. Os modelos de pagamento por procedimento (*fee-for-service*) e por capitação limitam a concorrência do setor e a capacidade dos sistemas de saúde de realizar mudanças transformadoras (HERNANDEZ A.; KAPLAN R.S.; WITKOWSKI M.L. *et al.*, 2019).

No livro “*Redefining Health Care*” Porter e Teisberg (2007) debruçam-se sobre o tema da combinação de altos custos, qualidade insatisfatória, acesso limitado na assistência a saúde, e propõem que através do redesenho do sistema de forma orientada a valor, é possível alcançar melhores resultados para a saúde da população sem que os custos se elevem continuamente. A estratégia proposta para transformar o sistema de saúde foi fundamentada no conceito de valor para os pacientes. O valor na assistência à saúde foi definido pelos autores como a razão entre os resultados de saúde que importam para os pacientes e o custo para entregar estes resultados (PORTER, 2010). Merece um destaque importante que a unidade a ser analisado o valor é a condição do paciente: um conjunto inter-relacionado de circunstâncias médicas do paciente que são mais bem endereçados de uma forma integrada, podendo considerar complicações e doenças simultâneas que devem ser tratadas sob a perspectiva do paciente e seu estado de saúde. Os autores ainda ressaltam que a reestruturação da saúde requer a substituição do sistema fragmentado de hoje, baseado no volume e lucratividade (VAN LEIJEN-ZEELBERG; ELISSEN; GRUBE *et al.*, 2016) para um sistema centrado no paciente e organizado em torno de suas condições clínicas (PORTER; LARSSON; LEE, 2016).

Para orientar a implementação da definição conceitual de VBHC, seis elementos interdependentes foram sugeridos, e juntos constituem a agenda de valor proposta por

Porter *et al.* (2010) (PORTER, 2010). Desde então, líderes de sistemas de saúde e pesquisadores têm investido esforços em desenvolver ferramentas para implementar a agenda. Algumas organizações e instituições já estão nos estágios de implementar os vários dos componentes estratégicos para tal mudança, como Cleveland Clinic nos Estados Unidos (PORTER, 2009), Martini Klinik na Alemanha (PORTER, 2014) e *Value Management Office* (VMO) no Brasil (MAKDISSE; KATZ; RAMOS *et al.*, 2018). Os seis componentes estratégicos são:

Instituições de saúde devem se organizar em unidades práticas integradas (UPI), onde a equipe dedicada fornece o ciclo completo de cuidados para cada condição de saúde. Para isso, é sugerido o estabelecimento de uma entidade organizacional que concentra múltiplas especialidades e se organiza ao redor de uma condição clínica, com objetivo de entregar o máximo valor para os pacientes ao menor custo possível. O principal caso de sucesso conhecido de IPU é a Martini's Klinik (PORTER, 2014) na Alemanha, fundada em 2005 e concentrou-se unicamente no tratamento do câncer de próstata com o compromisso de melhorar a qualidade do serviço na Alemanha.

As instituições de saúde devem medir os resultados e os custos para todos os pacientes em condições clínicas específicas durante todo o ciclo de cuidados. Para qualquer condição de saúde os resultados são considerados em camadas com três níveis de hierarquia, sendo eles: variação do estado de saúde (nível 1); o processo de recuperação (nível 2), e sustentabilidade da saúde (nível 3). Assim como os desfechos, os custos devem ser medidos a nível de paciente para cada condição clínica e por isso, o método de microcusteio baseado em atividade e tempo (TDABC) tem sido recomendado, uma vez que é um método mais robusto no que diz respeito aos recursos que estão sendo usados para gerenciar a condição de saúde do paciente (ETGES; RUSCHEL; POLANCZYK *et al.*, 2020; PORTER; LARSSON; LEE, 2016; THAKER; PUGH; MAHMOOD *et al.*, 2016; TSAI; PORTER; ADAMS, 2018).

Instituições de saúde devem ser financiadas por meio de pagamento "*bundled*". Os pagamentos devem ser efetuados para uma condição de saúde através do ciclo completo de cuidados. O atual cenário, em que os pagamentos são feitos e incentivados a aumentar seu volume, é um fator limitante para a implementação de uma gestão em saúde centrada no paciente. Se há apenas um valor de pagamento para determinado procedimento, mas a quantidade de tempo, recursos humanos ou materiais necessários para fornecer o serviço variam de paciente para paciente, o prestador de serviços que atender pacientes complexos e de alto risco pode estar sendo penalizado

financeiramente. O pacote de pagamentos para todo o ciclo de atendimento obriga os provedores de saúde a aumentar os resultados e diminuir os custos em todo o ciclo de atendimento. Como tal, os prestadores de cuidados de saúde são responsáveis pelo valor de todo o percurso de atendimento, desde o diagnóstico à reabilitação (VAN VEGHEL; SCHULZ; VAN STRATEN *et al.*, 2018).

As instituições de saúde devem unificar a entrega de cuidados, a fim de eliminar a fragmentação e duplicação de atendimento, e otimizar os tipos de atendimento prestado em cada localização. Esforços que contribuam para que o paciente com sua devida complexidade chegue no local correto, pelo meio correto, no momento correto (PORTER, 2017; 2019).

As instituições de saúde devem expandir geograficamente. Os provedores de serviços devem concentrar-se em aumentar o seu alcance clínico, por meio do direto ou indireto envolvimento com outros provedores, assistência por telemedicina ou clínicas de acesso a atenção básica em saúde (PORTER, 2010).

Além da expansão geográfica, as instituições de saúde devem ter capacidade tecnológica informacional para sustentar a implementação dos demais componentes da agenda de valor. Um requisito importante desses componentes é a disponibilidade de uma plataforma de TI abrangente, que permite a integração dos dados do paciente e o rastreamento contínuo de métricas (MESKÓ; GÖRÖG, 2020).

Embora cada componente da estratégia de valor de Porter e Lee (2010) seja mutualmente importante, esta pesquisa se concentrou nos elementos de medir os resultados e os custos para cada paciente e de uso da tecnologia de informação.

2.2. DESFECHOS EM SAÚDE

Em relação ao numerador da equação de valor, enquanto na indústria e produção de serviços as medidas de desempenho referem-se a função do produto na cadeia de valor (COOPER, 2017), na área da saúde tais medidas são derivadas de diretrizes clínicas baseadas em evidências, que enfatizavam a medição do desfecho (LEUNG; VAN MERODE, 2019; PORTER; LEE, 2013). Tipos de medições de desfecho frequentemente inspiram-se na abordagem *donabediana*, descrita pela primeira vez em 1988, em que as medidas são categorizadas em três categorias: estrutura, processo e resultado (DONABEDIAN, 1988). As medidas estruturais referem-se a estruturas de suporte que permite o atendimento seguro e eficaz. As medidas de processo referem-se

a processos de tratamento, cuidado e manejo clínico (DONABEDIAN, 2002). E, finalmente, as medidas de resultado incluem o estado de saúde, medidas, experiência do paciente e qualidade de vida. É importante destacar que os resultados de saúde devem ser medidos considerando o ciclo completo de cuidados de uma condição médica (KATZ; FRANKEN; MAKDISSE, 2017).

No entanto, mesmo em 2016, resultados não foram medidos com a frequência que deveriam ser. Em uma análise de 1.958 medidas da *US National Quality Measurement Clearinghouse* de um registro de medições de várias organizações de relatórios de qualidade, mostrou que apenas 7% as medidas foram realmente de resultados e menos de 2% foram resultados relatados pelo paciente (PORTER; LARSSON; LEE, 2016). Ainda, a medição de variáveis de processo tem certo efeito limitador sobre o valor, tal medidas recebem pouca atenção de pacientes, que estão interessados nos resultados (PORTER; LARSSON; LEE, 2016).

Dado que não se tem um mecanismo de padronização, não é incomum encontrarmos resultados inconsistentes de medidas e definições utilizadas pelas organizações provedoras, sociedades de especialidade, pagadores e países (DONABEDIAN, 1988; IORIO; CLAIR; INNEH *et al.*, 2016; NORIA; NEEDLEMAN; MIKAMI *et al.*, 2015). A definição de mensuração de desfechos em saúde frequentemente é sobreposta com a de medição de indicadores de qualidade assistencial. No entanto, a mensuração de desfechos visa superar as limitações impostas pelo uso de indicadores, uma vez que, neste contexto, um indicador de qualidade tem pouco valor se o desempenho que está sendo medido não pode ser melhorado (por exemplo, fornecendo feedback do profissional ou a oportunidade de reflexão) ou se nada é feito com o dado observado do indicador (DRIESSEN; VAN ZWET; HAAZEBROEK *et al.*, 2016).

Para entregar valor, os hospitais devem conseguir abranger todas as categorias de desempenho que são mais significativas para linha de cuidado escolhida e condizentes com sua realidade. O uso de medidas de resultado reportadas pelos pacientes (*patient-reported outcomes, PROMs*) deve se tornar cada vez mais frequente entre as formas de avaliação de desempenho adotadas pelos hospitais que almejam estabelecer uma gestão baseada em valor (PORTER; LARSSON; LEE, 2016).

Em meio a este contexto, o Consórcio Internacional de Medição de Desfechos em Saúde (*International Consortium of Health Outcomes Measurement, ICHOM*) convocou grupos de especialistas, juntamente com representantes dos pacientes, para

delinear conjuntos de métricas de resultados que importam para os pacientes e comorbidades que devem ser coletadas. Esse grupo de trabalho do ICHOM reforça que o papel de tal parametrização não é conceber novas métricas de resultados, mas concordar sobre qual indicador de medida de desfecho bem avaliado todos devem usar para uma dada condição médica. A natureza internacional do esforço permitiu aos participantes ver que os pacientes com uma condição específica têm a mesma ou semelhante necessidade em diferentes regiões do mundo (ICHOM; 2012).

Para competir em resultados, é preciso que estes sejam mensurados e amplamente difundidos. O desempenho do sistema de saúde só vai ser significativamente melhorado se os resultados forem mensurados e cada participante do sistema, responsabilizado pelos seus resultados (MICHAEL E. PORTER; KAPLAN, 2016). A possibilidade de medir os resultados e de usar a equidade no controle das circunstâncias iniciais do paciente é necessária para medir resultados e para melhorá-los (PORTER, 2010). Como exemplo de medição de desfechos sem o ajuste de risco, pode ser mencionado o caso do estado de Maine nos Estados Unidos: o sistema de tratamento de dependentes químicos em Maine (*The Maine Addiction Treatment System, MATS*) começou a exigir que as clínicas de reabilitação de drogas publicassem seus resultados. Estudos subsequentes mostraram que a melhoria nos resultados alcançados nos anos que se seguiram à legislação foi quase integralmente atribuível ao fato das clínicas afastarem pacientes considerados problemáticos e de alto risco para aumentar suas taxas de sucesso (SHEN, 2003). Também foram demonstradas melhorias no valor para o paciente quando os resultados são medidos e comparados (VAN DEN BERG; DIKSMAN; KEUS *et al.*, 2020; VAN VEGHEL; SCHULZ; VAN STRATEN *et al.*, 2018), como compreensão da variabilidade na taxa de cesárias feitas pelos prestadores (PBGH, 2014), eficiência no tempo de atendimento e alta do paciente (IORIO; CLAIR; INNEH *et al.*, 2016) e comparação desses indicadores com médias nacionais (PORTER, 2014).

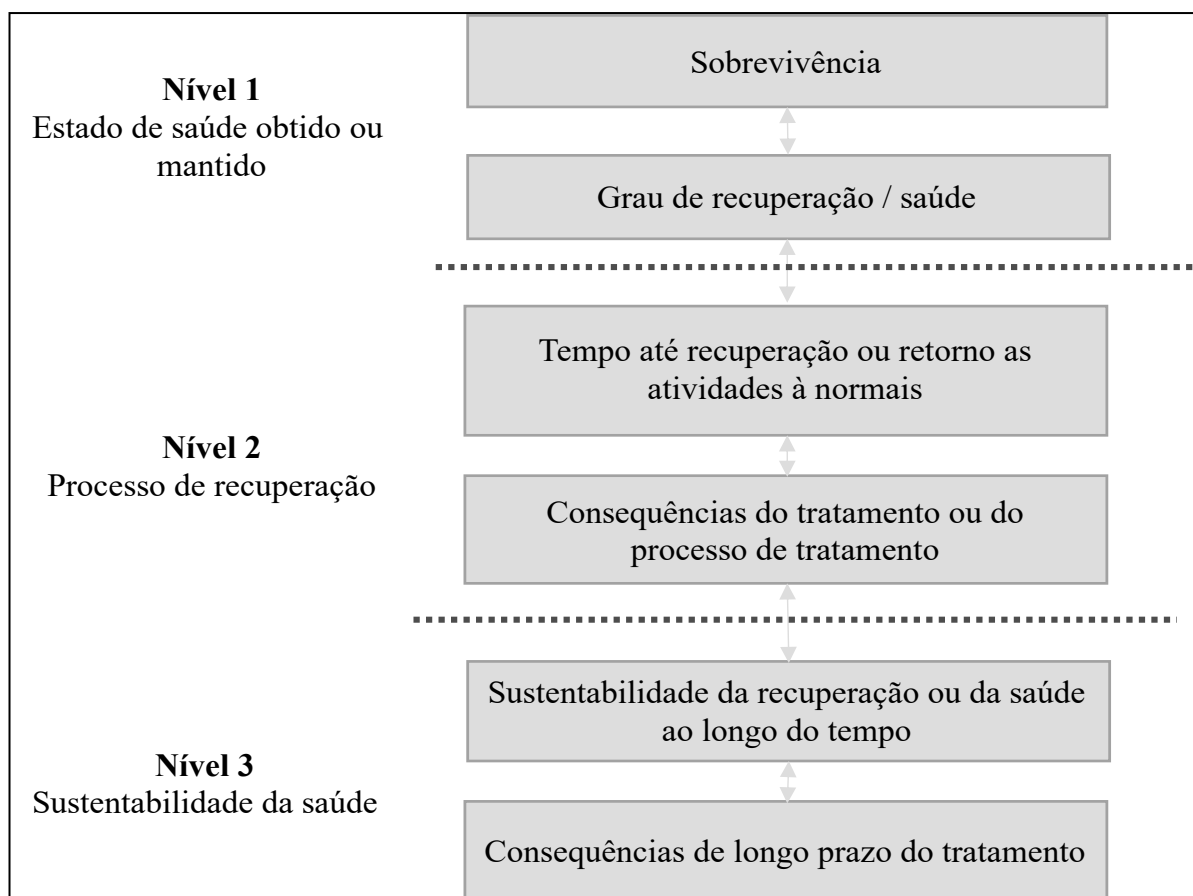
Influenciados pela iniciativa da Martini Klinik, esforços em caráter global de benchmarking em câncer de próstata hoje contém mais de 80 centros em 14 países usando medidas do ICHOM, e tem adicionado centros participantes anualmente (ATEFA SHAH; BLASZCZYNSKI, 2016; PORTER; LARSSON; LEE, 2016). Apesar de iniciativas internacionais, a medição de resultados tem sofrido resistência por parte dos prestadores, que temem vieses e comparações e são facilmente imergidos em inúmeras variáveis de medições que tornam esse trabalho de mensurar a qualidade

complexo (PORTER; TEISBERG, 2006). A obrigatoriedade de mensuração e a emissão de boletins de resultados talvez seja o passo mais importante de todos na reforma do sistema de saúde (PORTER; TEISBERG, 2006).

Medidas de resultado muitas vezes falham em capturar dimensões que são altamente importantes para os pacientes (NORIA; NEEDLEMAN; MIKAMI *et al.*, 2015; PORTER; LARSSON; LEE, 2016). O que geralmente importa para os pacientes são resultados que abrangem todo o ciclo de cuidado - incluindo estado de saúde alcançada (por exemplo, sobrevivência, status funcional, qualidade de vida); o tempo, complicações e sofrimento envolvido em obter cuidados; e a sustentabilidade dos benefícios alcançados (por exemplo, tempo até a recorrência) (PORTER; LARSSON; LEE, 2016; THAKER; PUGH; MAHMOOD *et al.*, 2016).

Diante disso, Porter et al. (2016) introduziu a hierarquia dos desfechos que consiste em resultados para qualquer condição de saúde que pode ser organizada em uma hierarquia de três níveis (**Figura 1**).

Figura 1-Hierarquia de Desfechos proposta por Porter.



Fonte: a autora, traduzida e adaptada de Porter (2010).

Construir modelos de monitoramento de desfechos que contemplem medidas dos três níveis requer gerenciamento da quantidade e qualidade de informações coletadas (IORIO; CLAIR; INNEH et al., 2016; NORIA; NEEDLEMAN; MIKAMI et al., 2015). A definição de formas de coleta e avaliação de desfechos na rotina de atendimento da equipe assistencial ainda é um obstáculo (MEULEN, 2019; PORTER; LARSSON; LEE, 2016; PORTER; TEISBERG, 2006).

A evolução para registros que informam e fornecem assistência centrada no paciente praticável pode ter dois caminhos principais: primeiro, a educação e disseminação de uma cultura de valor, que instantaneamente reforça a equipe para registrar resultados importantes sobre o paciente, seja por meio de questionários validados ou encontros multidisciplinares mais eficazes; e, segundo, a criação de uma plataforma de medição de resultados em tempo real integrada ao sistema de prontuários eletrônico do hospital (BLUMENTHAL; TAVENNER, 2010). Essa reflexão já foi colocada por diversos autores e nos leva a sugerir que o VBHC não é viável sem investimento em tecnologia da informação (BLUMENTHAL; TAVENNER, 2010; ETGES; RUSCHEL; POLANCZYK *et al.*, 2020; LEE, 2010; MEULEN, 2019).

2.3. PRONTUÁRIO ELETRÔNICO / *ELETRONIC HEALTH RECORDS (EHR)*

Ao descrever sobre registros eletrônicos em saúde, Pinto (2006) constatou que este possui diferentes denominações, sendo elas: prontuário do paciente, prontuário médico e registro do paciente.

A Resolução nº 1638, de 09 de agosto de 2002, no Artigo 1º do Conselho Federal de Medicina (CONSELHO FEDERAL DE MEDICINA, 2002), define o prontuário do paciente como “(...) documento único constituído por um conjunto de informações, sinais e imagens registradas, geradas a partir de fatos, acontecimentos e situações sobre a saúde do paciente e a assistência a ele prestada, de caráter legal, sigiloso e científico, utilizado para possibilitar a comunicação entre membros da equipe multiprofissional e a continuidade da assistência prestada ao indivíduo”.

Tal registro, portanto, é caracterizado pela variedade de autores que produzem seus registros (enfermeiros, médicos, fisioterapeutas, nutricionistas, fonoaudiólogos, psicólogos, assistentes sociais, entre outros) e pela variedade de informações nele registradas, as quais podem ser objetivas ou subjetivas (CONSELHO FEDERAL DE MEDICINA, 2002). Dentre esses dados estão aqueles em formato de texto livre,

também denominados de narrativas clínicas como, por exemplo, evoluções do estado clínico do paciente, laudos de procedimentos e exames, sumários de alta, anamnese, descrição dos sintomas, evolução do paciente ao longo de internação, entre outras (HYUN; JOHNSON; BAKKEN, 2009; RUIZ, 2016).

Os registros eletrônicos em saúde, portanto, podem conter campos não estruturados e estruturados. Campos estruturados são dados organizados de tal maneira a rapidamente fornecer a informação, possuem campos específicos de preenchimento e podem ser extraídos em formato de tabelas ou relatórios de indicadores. Geralmente este tipo de dados são os mais utilizados pela facilidade de identificação, uma vez que o conteúdo do campo estruturado representa exatamente a informação como, por exemplo, idade ou uma codificação referente à sexo (NI; WRIGHT; PERENTESIS *et al.*, 2015). Por outro lado, campos não estruturados são aqueles que apresentam a informação em formato de texto. Ao passo que possuem um maior detalhamento, é mais difícil extrair informação deste tipo de dado justamente por estar dependente ao tipo do relato ou escrita de quem o fez. Nesta ótica, os registros produzidos em formato de texto livre apresentam informações com maior riqueza de detalhes se comparados ao conteúdo dos dados estruturados (BEAM, ANDREW L.; KOHANE, ISAAC S., 2018; SHIVADE; HEBERT; LOPETEGUI *et al.*, 2015). Por esta razão, o tratamento adequado dessas informações representa um passo importante para apoiar o julgamento clínico e a tomada de decisões dos profissionais de saúde (NI; WRIGHT; PERENTESIS *et al.*, 2015).

Um exemplo dessa diferença é percebido quando se escolhe utilizar apenas dados estruturados para recuperação de informações. Essa escolha pode resultar na perda de informações clínicas relevantes (HYUN; JOHNSON; BAKKEN, 2009). Por exemplo, a lista de medicamentos pode ser facilmente capturada no formato estruturado, entretanto as nuances relacionadas à administração da medicação, como história dos esquemas anteriores e reações adversas durante o tratamento, são registradas e obtidas nas narrativas clínicas (HYUN; JOHNSON; BAKKEN, 2009). Isto também se aplica às informações colhidas durante a anamnese e transcritas no prontuário como, por exemplo, queixa principal, história da doença atual e pregressa, história familiar e social (GROSSMAN; CARDOSO, 2006). No **Quadro 1** são apresentados dois exemplos de narrativas clínicas.

Quadro 1-Exemplos de narrativas clínicas (evoluções)

Exemplo 1
<p>Nota de admissão: enfermagem emergência Paciente XXX, idade XXX, interna por AVC isquêmico. Paciente com queda da própria altura, e quadro de perda de força de membros inferiores. Cefaléia intensa. Desconhece alergias. #HAS; #AVC isquêmico; #Osteoporose. S- Paciente sem queixas no momento, inclusive algias. O- Paciente deitada no leito, LOC e tranquila. Ventila em ar ambiente. Em NPO. Diurese presente em comadre. Fralda para conforto. Cateter venoso periférico em MSE salinizado. Pele íntegra. Deambula com auxílio em domicílio. Sem familiar presente no momento. C- Observar sensorio. Observar integridade da pele. Observar sinais flogísticos em inserção de cateter venoso. Comunicar alterações. Cuidados de enfermagem. Transfiro paciente para leito de CTI adulto box 65.</p>
Exemplo 2
<p>Fisioterapia Neurológica - 09:50 Ambiente: Encontro paciente no leito, acompanhada, alerta, plégica em MSE, parética MIE. Condição Ventilatória: Ventilando em ar ambiente, mantém bom padrão ventilatório. AP: MV diminuído s/RA Funcionalidade: Alterada Origem da limitação: Neurológica Nível de Mobilidade: 8 Condutas: Mobilizações passivas de MSE e passivos de MIE. Ativos resistidos leves de MSD e MID. Sedação a beira do leito, reforço o treino de endireitamento de cervical, ruim desempenho. Pivô para poltrona e alongamentos de MsIs/MsSs. Plano: estimular movimentos ativos, treinar rolar, treinar tronco, tentar ortostase quando em condições. Observações: Ficar na poltrona o maior tempo possível. Fica na poltrona, estável, acompanhada.</p>

Fonte: a autora, 2020.

Nos Estados Unidos, embora 96% dos hospitais e 86% dos consultórios médicos tenham adotado prontuários eletrônicos, ainda não há uma maneira fácil e em tempo real que possa analisar as medidas de desfecho relevantes para sustentar uma gestão baseada em valor (BLUMENTHAL; TAVENNER, 2010). Os resultados do estado funcional (por exemplo, se um paciente com pós-AVC pode caminhar ou falar) estão ocultos em texto livre e não são capturados em forma analisável. Por isso, os relatórios de produtividade são mais voltados a número de atendimentos/visitas, mas não sobre capacidade do paciente de sustentar sua saúde pós-alta, ou consequências de médio-longo-prazo da terapia (GLASER, 2020).

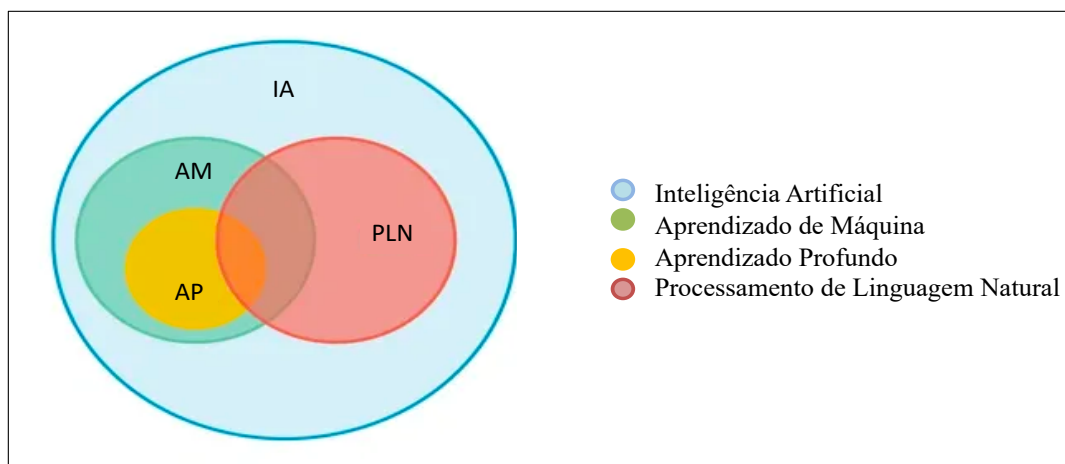
Existem duas fases típicas na avaliação da condição de um paciente a partir dos dados do prontuário eletrônico (KOWSARI; JAFARI; HEIDARYSAFA *et al.*, 2019; SHEIKHALISHAHI; MIOTTO; DUDLEY *et al.*, 2019; VELUPILLAI; SUOMINEN; LIAKATA *et al.*, 2018; WANG, S. Y.; PERSHING, S.; TRAN, E. *et al.*, 2020): extração de informações, na qual usa-se técnicas PLN para extrair palavras ou frases de

interesse do texto livre; e *classificação*, que aplica modelos de aprendizado de máquina para identificar resultados derivados dos termos extraídos.

2.4. A INTELIGÊNCIA ARTIFICIAL COMO CATALISADOR PARA MELHORAR O VALOR NOS CUIDADOS DE SAÚDE

Conceitualmente, IA pode ser definida como um conjunto de técnicas analíticas que alavancam grandes conjuntos de dados, estruturados e não estruturados, para oferecer previsões precisas ao encontrar padrões ou classificar dados de interesse (AGRAWAL; GOLDFARB; GANS, 2018; MESKÓ; GÖRÖG, 2020). Técnicas de IA podem ser categorizadas em três grandes grupos (**Figura 2**): i) Técnicas clássicas de aprendizado de máquina, ii) técnicas de aprendizado profundo mais recentes, e iii) métodos de processamento de linguagem natural (PLN) (AGRAWAL; GOLDFARB; GANS, 2018; BEAM, A. L.; KOHANE, I. S., 2018; LING; KURIAN; CASWELL-JIN *et al.*, 2019; MEULEN, 2019; SUN; CAI; LI *et al.*, 2018; XIAO, C.; CHOI, E.; SUN, J., 2018). As três subáreas serão abordadas, porém a ênfase maior será na última, visto seu papel no tema da dissertação.

Figura 2- Grande área da Inteligência artificial e como suas subáreas se conectam



Fonte: a autora, 2020.

Algoritmo é uma sequência finita de ações executáveis que visam obter uma solução para um determinado tipo de problema (MESKÓ; GÖRÖG, 2020). Os algoritmos são os blocos de construção da IA, e são divididos em 3 partes, conforme **Figura 3:**

Figura 3-Representação de um algoritmo.



Fonte: a autora, 2020

Um ponto importante é que algoritmos não aceitam dados de texto, então no caso de documentos e textos livres, o dado precisa assumir uma representação numérica para que possam ser submetidos a métodos formais de classificação (KOWSARI; JAFARI; HEIDARYSAFA *et al.*, 2019). A representação numérica é comumente chamada de *feature*. A *feature* compõe a entrada do dado no algoritmo e pode representar, por exemplo, idade, histórico médico e sexo (MEULEN, 2019).

2.4.1. Aprendizado de máquina clássico / Machine Learning

Os algoritmos de aprendizado de máquina aplicam métodos como estatísticas Bayesianas e regressão linear a conjuntos de dados com objetivo de reconhecimento de padrões (MESKÓ; GÖRÖG, 2020; MEULEN, 2019). Os modelos de aprendizado de máquina podem ser categorizados em três grandes grupos: modelos supervisionados, modelos não supervisionados e modelos de reforço (BARBA, 2019; MEULEN, 2019).

Em um modelo de aprendizado supervisionado, o algoritmo aprende com pares específicos de entrada-saída, fornecendo uma resposta-chave que o algoritmo pode usar para avaliar sua precisão nos dados de treinamento (KOWSARI; JAFARI; HEIDARYSAFA *et al.*, 2019; MESKÓ; GÖRÖG, 2020; MEULEN, 2019). Modelos mais avançados precisam de dados anotados para garantir que eles possam aprender a tarefa a qual foram desenhados, de modo que a contribuição dedicada de anotadores de dados familiarizados com formato do banco de dados é de importância crucial para o benefício de implementação de IA no ambiente de saúde (MESKÓ; RADÓ, 2019). A

Figura 4 ilustra o desenvolvimento de modelos supervisionados.

Figura 4 – Ilustração do aprendizado supervisionado.



Traduzido e adaptado pela autora, 2020. Fonte: Ronald van Loon, Big Data Made Simple.

A marcação e a anotação de texto desempenham um papel importante na integração de dados estruturados e não estruturados (LOPES; VIEIRA; FINATTO et al., 2009). Por sua natureza, os dados não estruturados dificultam a extração direta informações significativas e por isso a marcação e a anotação de texto desempenham um papel importante na integração desses dados (KNOX; RITA; EID et al., 2005).

Um modelo não supervisionado, ao contrário, fornece dados não rotulados que o algoritmo tenta entender extraindo recursos e padrões por conta própria. Problemas de classificação e problemas de regressão são duas grandes áreas onde a aprendizagem supervisionada é útil; enquanto que, quando se tem uma coleção de exemplos sem um resultado desejado específico ou resposta correta, a abordagem não supervisionada pode ser uma alternativa, uma vez que o algoritmo tenta encontrar automaticamente a estrutura nos dados, extraindo recursos úteis e analisando sua estrutura, como por exemplo, separando a base de dados em grupos de características semelhantes (KOWSARI; JAFARI; HEIDARYSAFA *et al.*, 2019; MESKÓ; GÖRÖG, 2020; MEULEN, 2019). A **Figura 5** ilustra o desenvolvimento de modelos não supervisionados.

Figura 5- Ilustração do aprendizado não supervisionado



Traduzido e adaptado pela autora, 2020. Fonte: Ronald van Loon, Big Data Made Simple.

Por isso, a aprendizagem supervisionada é mais adequada para problemas em que existe um conjunto de pontos de referência disponíveis ou uma verdade fundamental com a qual treinar o algoritmo (KOWSARI; JAFARI; HEIDARYSAFA *et al.*, 2019; MEULEN, 2019; SUN; CAI; LI *et al.*, 2018). Dado que isso nem sempre está disponível, uma abordagem frequente que tem sido adotada é o uso de métodos mistos, denominados modelos semi-supervisionados. A aprendizagem semi-supervisionada consiste no conjunto de dados de treinamento com dados rotulados e não rotulados (MESKÓ; GÖRÖG, 2020; MEULEN, 2019).

2.4.2. *Aprendizado Profundo / Deep Learning*

O aprendizado profundo é um método relativamente novo de aprendizado de máquina que usa “camadas de neurônios” em uma rede para encontrar padrões ocultos em conjuntos de dados complicados, como imagens (MESKÓ; GÖRÖG, 2020). O aprendizado profundo é visto como a próxima etapa do aprendizado de máquina e tem a promessa de criar algoritmos que resolvam alguns dos desafios mais difíceis em IA, como visão computacional e reconhecimento de fala, permitindo o treinamento de algoritmos que funcionam em tempo real com altíssima precisão. Algoritmos de aprendizado profundo usam as chamadas "Redes Neurais Artificiais" para aprender

padrões, inspirados na relação e sinapses dos neurônios do nosso cérebro (XIAO, C.; CHOI, E.; SUN, J., 2018).

2.4.3. Processamento de Linguagem Natural / Natural Language Processing

Em paralelo, estudos de PLN desenvolvidos com o propósito de extração de informação vêm progredindo com o passar dos anos e revelando um caminho que pode melhorar a precisão da combinação paciente e desfecho (NI; WRIGHT; PERENTESIS *et al.*, 2015). Informações clínicas importantes geralmente são registradas em texto livre não estruturado e convertê-lo em um formato estruturado pode ser uma tarefa demorada que pode não capturar com sucesso todas as facetas das informações (KREIMEYER; FOSTER; PANDEY *et al.*, 2017).

As técnicas de PLN aproveitam a IA para analisar grandes volumes de dados textuais para interpretar a linguagem. No PLN, os métodos de análise sintática e semântica analisam o texto de acordo com as regras gramaticais e o significado real do texto, respectivamente. Algoritmos de PLN usam dados não estruturados, como notas clínicas ou textos de livros ou artigos de pesquisa (**Machine Learning for Natural Language Processing**, 2019; LING; KURIAN; CASWELL-JIN *et al.*, 2019; XIAO, CAO; CHOI, EDWARD; SUN, JIMENG, 2018).

2.5. O PROBLEMA DE CLASSIFICAÇÃO DE TEXTO

Estamos diante de uma esmagadora onda de crescimento de dados em todos os setores. Em pesquisa feita pela Ernst Young registrou que 90% dos dados no mundo foram criados nos últimos 2 anos, e que destes 80% eram dados não estruturados (EY, 2019). Constatações como essa evidenciam um paradigma sobre avançar em tecnologia que permita estruturar e entender esses dados, ou se seguir com trabalho manual que permite o alcance de análises menos representativas e escaláveis (XIAO, CAO; CHOI, EDWARD; SUN, JIMENG, 2018).

Infelizmente os avanços de uso de PLN na saúde ainda andam em ritmo mais lento que outros mercados e são pouco explorados na literatura de saúde (XIAO, CAO; CHOI, EDWARD; SUN, JIMENG, 2018). Wang *et al.* (2018) em sua revisão sistemática aponta que a falta de padronização na linguagem clínica e a dificuldade de acesso aos dados de pacientes ainda é a principal barreira para o desencontro de estudos

que interoperem entre PLN e EHR. Por meio dos esforços de colaboração existentes ou da construção e aproveitamento de uma infraestrutura de computação que preserve a privacidade, será mais comum usar dados de EHR para estruturar narrativas clínicas e apoiar a extração de informações clínicas para apoiar tomada de decisão (WANG, SOPHIA Y.; PERSHING, SUZANN; TRAN, ELAINE et al., 2020).

Além do banco de dados bruto - caracterizado por um documento com vários textos em sequência – é necessário um conjunto estruturado para o propósito de treinamento (KOWSARI; JAFARI; HEIDARYSAFA et al., 2019). Em geral, o sistema de classificação de texto contém quatro níveis diferentes de escopo que podem ser aplicados para construção do vetor, do documento, de parágrafo, de frase, e de sub-sentença (KOWSARI; JAFARI; HEIDARYSAFA et al., 2019), sendo a definição do mais apropriado dependente do objetivo da análise. No nível do documento, o algoritmo obtém as categorias relevantes de um completo documento. No nível de parágrafo, o algoritmo obtém as categorias relevantes de um único parágrafo (uma parte de um documento). No nível de frase, obtém as categorias relevantes de uma única frase (uma parte de um parágrafo). Por fim, no nível de sub-sentença, o algoritmo obtém as categorias relevantes de sub-expressões dentro de uma frase.

A extração e o pré-processamento de atributos são etapas cruciais para aplicações de classificação de texto. Neste sentido, não é incomum que métodos para limpar conjuntos de dados de texto, removendo assim o ruído implícito e permitindo caracterização, sejam a nível de sentença (KOWSARI; JAFARI; HEIDARYSAFA et al., 2019). No entanto, a desvantagem em trabalhar com nível de sentença é que dependendo do método que a quebra dos dados é feita, o contexto formado entre as sentenças pode fornecer informações importantes para a classificação, e pode ser ignorado pelo fato da marcação estar ocorrendo neste nível e não no nível da frase (KOWSARI; JAFARI; HEIDARYSAFA et al., 2019). Detalhes sobre o pré-processamento utilizado no experimento do artigo 2 é apresentado no APÊNDICE A.

2.5.1. Representação de texto

Em geral, textos e documentos são conjuntos de dados não estruturados. Contudo, essas sequências de texto não estruturadas devem ser convertidas em um espaço de atributo estruturado para só assim ser associado a um classificador. Primeiro, os dados precisam ser limpos para omitir caracteres e palavras. Com o texto de saída do

pré-processamento, são utilizados como entrada nos métodos que criam as representações. As técnicas comuns de representação de texto são *Term Frequency-Inverse Document Frequency* (TF-IDF)(ANDRADE; GONÇALVES, 2020), frequência de termo (do inglês, *term-frequency* (TF)) (SALTON; BUCKLEY, 1988), Word2Vec (GOLDBERG; LEVY, 2014), vetores globais para palavras representação (GloVe)(PENNINGTON; SOCHER; MANNING *et al.*, 2014) e *bag-of-words* (BoW).

O TF-IDF é constituído por duas partes, o TF quantifica a frequência dos termos, e o IDF quantifica a frequência inversa de documento, que atribui maior peso as palavras que ocorrem com menos frequência nos documentos (ANDRADE; GONÇALVES, 2020; CUNHA; MANGARAVITE; GOMES *et al.*, 2020). Dessa forma, o score final de um termo t é obtido pelo $TF * IDF$, que potencializa palavras raras considerando todos os documentos e frequências no documento específico (ANDRADE; GONÇALVES, 2020; CUNHA; CANUTO; VIEGAS *et al.*, 2020; CUNHA; MANGARAVITE; GOMES *et al.*, 2020). Definimos que o TF-IDF seja gerado com *unigrams* e *bigrams*, em que o *bigram* é a combinação do termo atual com o termo posterior. Ainda, podem existir variações do TF-IDF, as mais comuns neste objetivo são Word TF-IDF que corresponde a analisar os termos presentes no documento, e Char TF-IDF que analisa uma sequência de caracteres definida aqui entre o intervalo de 2 a 6 caracteres. Em relação ao Word TF-IDF e Char TF-IDF, o Char consegue capturar melhor termos com erros de digitação porque está analisando parte de um termo definido no intervalo, por exemplo, se estamos analisando o termo ‘tabagista’ e ‘tabaisgta’, ele considera que ambos os termos são similares para um intervalo definido dos quatro primeiros termos. É comum fazer uso da concatenação dos vetores de ambas as representações, Word TF-IDF e Char TF-IDF, em uma única representação e este vetor é utilizado como entrada para um classificador, que serão apresentados na próxima seção. O objetivo desta combinação é potencializar os benefícios de ambas as representações(CUNHA; CANUTO; VIEGAS *et al.*, 2020).

Meta-atributos (do inglês, *metafeatures* ((MFs)) correspondem a representações de documentos derivadas de representações mais básicas, como TFIDF (CANUTO; MARTINS; COUTO *et al.*, 2015; S.; X.; A. *et al.*, 2018). As pontuações dos MFs são obtidas por funções que relacionam dados textuais com categorias (YANG; GOPAL, 2012). Neste trabalho, MFs foram explorados com base nas funções de distância entre exemplos de texto que recentemente obtiveram os resultados mais efetivos em diversos problemas de categorização de texto (CANUTO; MARTINS; COUTO *et al.*, 2015;

CUNHA; CANUTO; VIEGAS *et al.*, 2020; CUNHA; MANGARAVITE; GOMES *et al.*, 2020; S.; X.; A. *et al.*, 2018). A racional por trás desses meta-atributos consiste na suposição de que se as distâncias entre um exemplo de texto e os vizinhos mais próximos pertencentes à categoria “c” (e seu centróide correspondente) forem pequenas, então o exemplo provavelmente pertence a “c”. Tais pontuações de distância são evidências adicionais, que quando concatenadas à representação original, têm potencial de auxiliar na construção de modelos de classificação em coleções com uma grande quantidade de exemplos rotulados (CANUTO; SALLES; GONÇALVES *et al.*, 2014).

O word2vec é um método de *word embedding*, esta técnica de aprendizagem de atributos em que cada palavra ou frase do vocabulário é mapeada para um vetor de dimensão N de números reais. É o usual ponto de partida para as tarefas mais profundas de PLN, pois permite que o aprendizado de máquina seja eficaz em conjuntos de dados menores, pois geralmente é a primeira entrada de um modelo e a maneira mais popular de transferir o aprendizado no PLN (KOWSARI; JAFARI; HEIDARYSAFA *et al.*, 2019).

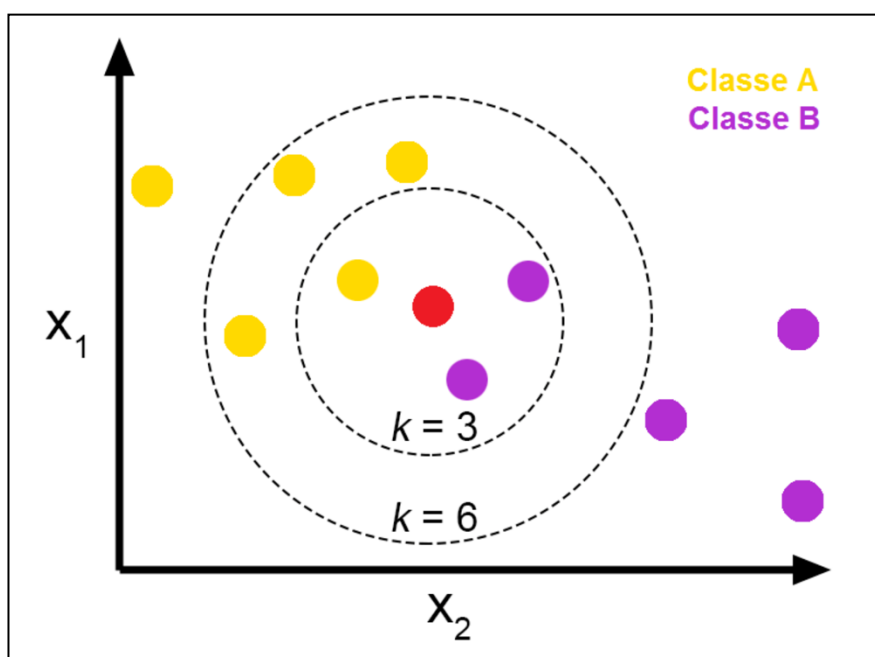
O modelo BoW, é uma representação reduzida e simplificada de um texto documento de partes selecionadas do texto, com base em critérios específicos, como frequência de palavras. A técnica BoW é usada em vários domínios, como visão computacional, PLN, bem como classificação de documentos, e recuperação de informações por aprendizado de máquina (KOWSARI; JAFARI; HEIDARYSAFA *et al.*, 2019).

2.5.2. Classificadores do texto

A etapa mais importante de classificação de texto é escolher o melhor classificador. Um dos algoritmos de classificação mais simples é a regressão logística (LR) que foi abordado na maioria dos domínios de mineração de dados (CHEN; XIE; WANG *et al.*, 2017; VELUPILLAI; SUOMINEN; LIAKATA *et al.*, 2018; WANG; WANG; RASTEGAR-MOJARAD *et al.*, 2018). Na história mais antiga da recuperação de informações como uma aplicação viável, o *Naïve Bayes Classifier* (NBC) foi muito popular, e ainda é bastante usado por ser computacionalmente barato e precisar de uma quantidade muito baixa de memória (LARSON, 2009), esta técnica é um modelo generativo, que é o método mais tradicional de categorização de texto (KOWSARI; JAFARI; HEIDARYSAFA *et al.*, 2019).

Técnicas não paramétricas foram estudadas e usadas como tarefas de classificação, como *K-Nearest Neighbor* (KNN) (LI; WEINBERG; DARDEN *et al.*, 2001). Dado um documento de teste x , o algoritmo KNN encontra os k vizinhos mais próximos de x entre todos os documentos no conjunto de treinamento e pontua os candidatos da categoria com base na classe de k vizinhos. A semelhança de x e o documento de cada vizinho pode ser a pontuação da categoria do vizinho documentos (**Figura 6**). Vários documentos KNN podem pertencer à mesma categoria; neste caso, o somatório de essas pontuações seriam a pontuação de similaridade da classe k em relação ao documento de teste x . Depois de classificar os valores de pontuação, o algoritmo atribui o candidato à classe com a maior pontuação do teste documento x (KOWSARI; JAFARI; HEIDARYSAFA *et al.*, 2019). KNN é um dos os algoritmos de aprendizado de máquina mais simples e tem bom desempenho no básico problemas de reconhecimento (LAYEGHIAN JAVAN; SEPEHRI; AGHAJANI, 2018).

Figura 6- Algoritmo *K-Nearest Neighbor* (KNN)

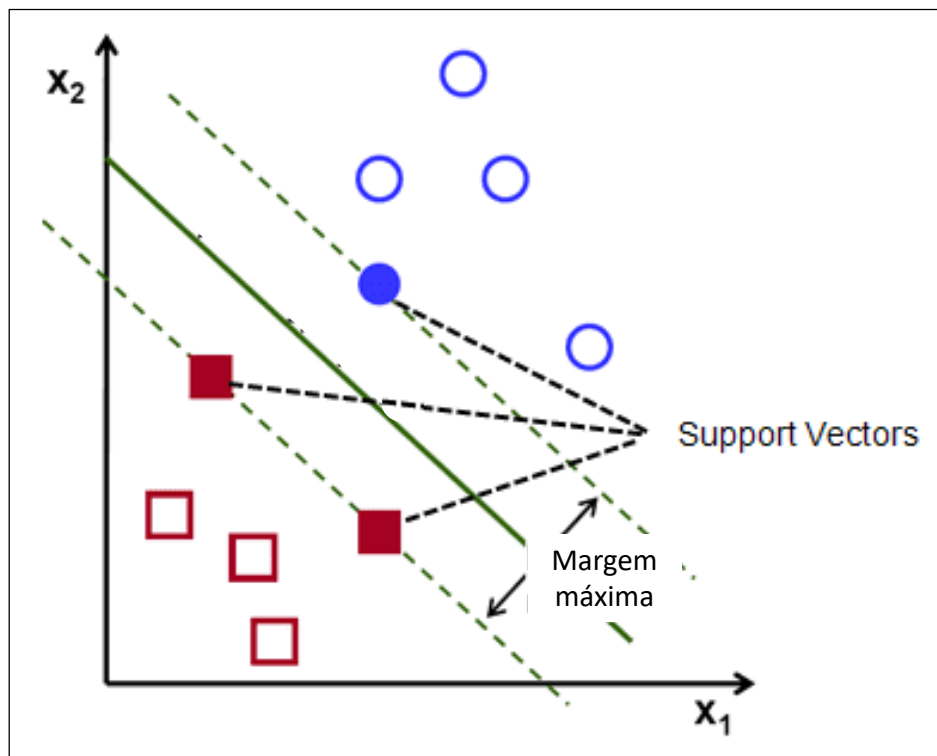


Traduzido e adaptado pela autora, 2020. Fonte: Pedregosa *et al.*, 2011.

Support Vector Machine (SVM) é outra técnica popular e com bons resultados que emprega um classificador discriminativo para categorização de documentos (CORTES; WOLOSZYN; BARONE, 2018; CRISTIANINI; SHAW-TAYLOR, 2000; CUNHA; MANGARAVITE; GOMES *et al.*, 2020; HAN; KARYPIS, 2000; MANEVITZ; YOUSEF, 2001), e esta entre os métodos mais empregados na área de saúde (LAYEGHIAN JAVAN; SEPEHRI; AGHAJANI, 2018; WANG; KIIK; PEEK *et al.*, 2020). Representa-se o texto através de um vetor de ocorrência de palavras,

desconsiderando gramática e ordem das palavras, mas mantendo a multiplicidade das mesmas (Figura 7) (YE, 2016). Esta técnica também pode ser usada em todos os domínios da mineração de dados, como bioinformática, imagem, vídeo, classificação da atividade humana, segurança e saúde. Este modelo também é usado como referência de comparação em muitas pesquisas para destacar novidades e contribuições (KOWSARI; JAFARI; HEIDARYSAFA *et al.*, 2019)./

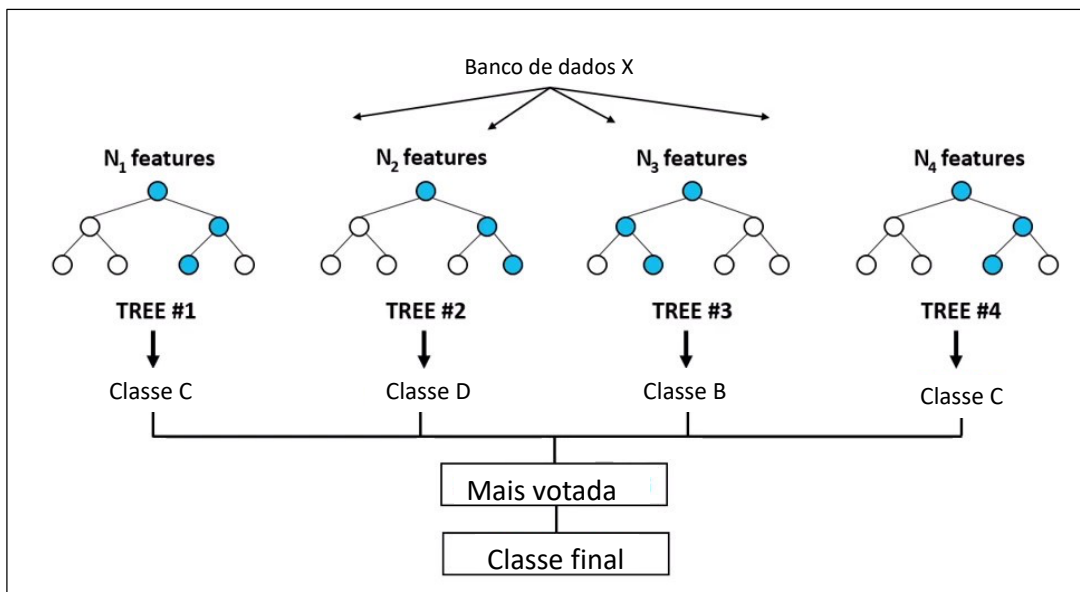
Figura 7- Algoritmo *Support Vector Machine* (SVM)



Traduzido e adaptado pela autora, 2020. Fonte: Pedregosa *et al.*, 2011.

Classificadores “baseados em árvore”, como árvore de decisão e *random forest* (BREIMAN, 2001) também foram estudados com no que diz respeito à categorização do documento (XU; YE; CHENG). O primeiro modelo consiste em um conjunto de decisões hierárquicas que chega a uma decisão final, enquanto a segunda abordagem tem como estratégia o uso de um conjunto de árvores de decisão (WANG; KIIK; PEEK *et al.*, 2020). Nos últimos anos, as classificações “baseadas em árvore” são usadas principalmente para sumarização de documentos (SHEN; SUN; LI *et al.*, 2007) e extração automática de palavras-chave (ZHANG, 2008).

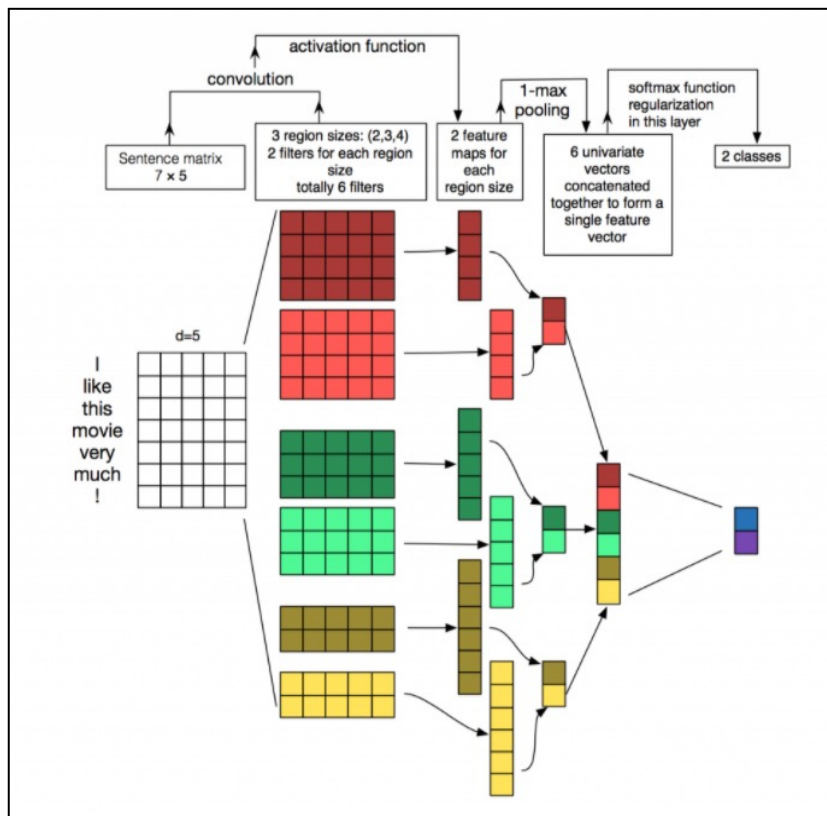
Figura 8- Algoritmo *Random Forest*



Traduzido e adaptado pela autora, 2020. Fonte: Pedregosa et al., 2011.

As abordagens de aprendizagem profunda, *Convolutional Neural Networks* (CNN), alcançaram resultados superiores em comparação com os anteriores algoritmos de aprendizado de máquina (GEORGAKOPOULOS; TASOULIS; VRAHATIS *et al.*, 2018) em tarefas como classificação de imagens, processamento de linguagem natural, reconhecimento facial, entre outros (JOHNSON; ZHANG, 2017).

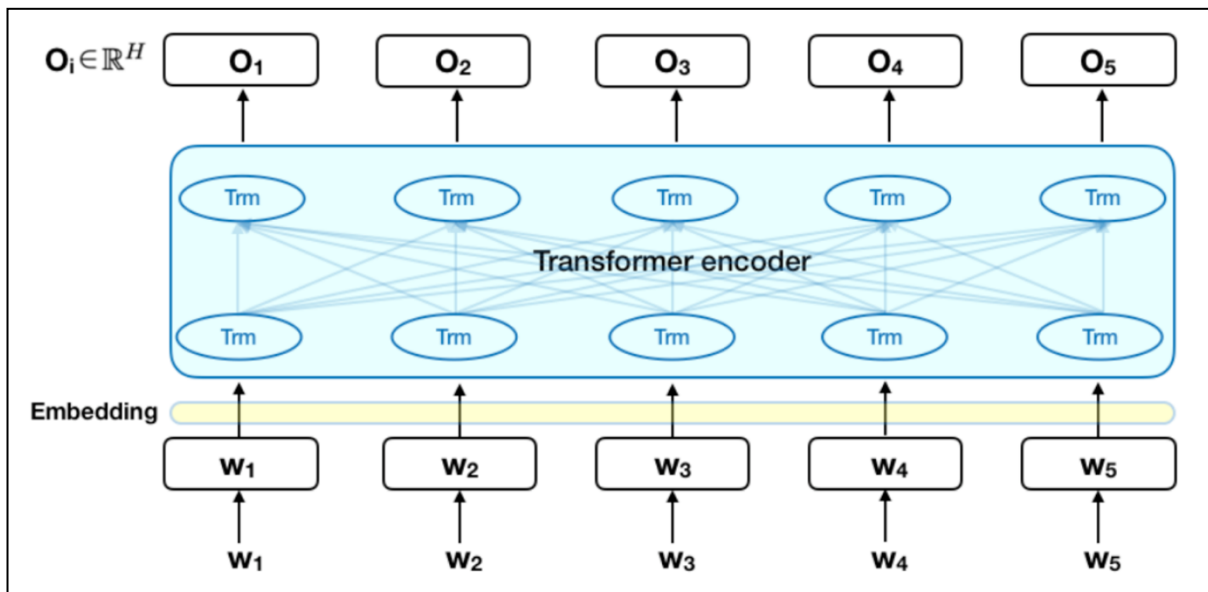
Figura 9- Figura clássica do algoritmo *Convolutional Neural Networks (CNN)*



Traduzido e adaptado pela autora, 2020. Fonte: Pedregosa et al., 2011.

Finalmente, o modelo BERT é método de aprendizado profundo de ponta-a-ponta pré-treinado com um corpus de 3,3 bilhões de palavras, incluindo *BooksCorpus* (800 milhões de palavras) e *Wikipedia* em inglês (2,5 bilhões de palavras). Dessa forma, ele prevê a falta de palavras em uma frase (**Figura 10**). BERT redefiniu o estado da arte para linguagem natural tarefa de processamento (CUNHA; MANGARAVITE; GOMES *et al.*, 2020). O sucesso desses algoritmos de aprendizado profundo depende de sua capacidade de modelar relacionamentos complexos e não lineares dentro dos dados (LECUN; BENGIO; HINTON, 2015).

Figura 10- Algoritmo BERT



Traduzido e adaptado pela autora, 2020. Fonte: Pedregosa et al., 2011.

2.5.3. Ontologias como modelo de classificação

Diferente dos processos descritos anteriormente que requerem a combinação de representadores de textos com algoritmos classificadores treinados, uma segunda opção para a classificação de textos livres é através da utilização de ontologias (ALMEIDA; BAX, 2003). Essa técnica tem mostrado eficácia de resultados equivalente aos modelos de classificação por aprendizado de máquina sem a necessidade de grande quantidade de dados e tempo de treinamento que comumente são requeridos por esses métodos (ALLAHYARI; KOCHUT; JANIK, 2014; CHI; LIN; HSIEH, 2014; GARLA; BRANDT, 2012; WANG; MCKAY; ABBASS *et al.*, 2003).

Ontologias são especificações, explícitas e formais, da conceitualização de um domínio de interesse, onde entende-se por formal uma especificação compreensível por um computador que com esse entendimento consiga realizar deduções sobre os conceitos (ALMEIDA; BAX, 2003; DAVIES; FENSEL; HARMELEN, 2003). Normalmente é incluso um vocabulário de termos e as especificações dos seus significados, e ainda as definições e as relações de como os conceitos estão inter-relacionados impondo uma estrutura sobre um domínio e limitando as possíveis interpretações dos termos (GOMEZ-PEREZ; CORCHO; FERNÁNDEZ-LÓPEZ, 2004).

Na ontologia, são explicitados conjuntos de termos que são relevantes para cada

variável e as regras de como estes se relacionam e se apresentam no texto de forma a categorizar determinada sentença como pertencente a uma determinada variável. Estes termos são elencados a partir do conhecimento de domínio e exemplos identificados nas evoluções. Esses termos são expandidos com uso de *word-embeddings* visando melhorar a abrangência desse conhecimento.

Com um algoritmo de busca textual é identificado os termos nas sentenças e definida uma classe na ontologia para essa sentença que contém todos os termos encontrados e as suas relações. Baseado nos termos e relações o processo de inferência é executado onde este deduz a qual variável a dada sentença pertence (LOPES; VIEIRA; FINATTO *et al.*, 2009).

2.6. AVC

As doenças cerebrovasculares estão entre as principais causas de morte no Brasil e no Mundo, apenas atrás das cardiovasculares e pesquisa sugerem que esse índice se mantenha até 2030 (MINISTÉRIO DA SAÚDE, 2018). O Acidente Vascular Cerebral (AVC) é uma das principais causas de mortalidade e invalidez de longo prazo em todo o mundo (JOHNSON; NGUYEN; ROTH *et al.*, 2019; OURIQUES MARTINS; SACKS; HACKE *et al.*, 2019). No Brasil, é um importante problema de saúde pública correspondendo a 107.658 mortes por ano, e um total de 80% dos pacientes com AVC são atendidos no sistema público de saúde (MINISTÉRIO DA SAÚDE, 2018). Ficando para trás apenas para doenças cardiovasculares, afecções do coração e circulação, que correspondem a mais de 1100 mortes por dia (MINISTÉRIO DA SAÚDE, 2018).

Uma iniciativa do Ministério de Saúde em 2008 em organizar serviços de urgência para melhorar a qualidade do atendimento a todos os pacientes fez com que a Rede Brasil AVC fosse criada (REDE BRASIL AVC, 2008). O projeto conta hoje com aproximadamente 160 hospitais de 19 Estados do Brasil. Através da Rede, profissionais e organizações são instruídos e apresentados ao que há de inovação e tecnologias aptas a serem incorporadas ao atendimento do AVC, permitindo tornar o atendimento nacional mais qualificado. Entre os avanços alcançados tem-se a implementação das Unidades de AVC que são áreas físicas preparadas e qualificadas para melhor atender aos pacientes com AVC. Elas também são caracterizadas pelo forte engajamento na educação continuada de pacientes, familiares e profissionais de saúde no âmbito da morbidade, e incentivo a pesquisa e desenvolvimento de tecnologias que possam

auxiliar na prevenção e/ou tratamento de doença (REDE BRASIL AVC, 2008).

Em 2013, pesquisadores brasileiros publicaram o manual de rotinas para a atenção do AVC com o objetivo de apresentar protocolos, escalas e orientações aos profissionais de saúde no manejo clínico ao paciente acometido por AVC (MINISTÉRIO DA SAÚDE, 2013a; b). Este manual tem sido adotado pelos hospitais da rede, permitindo o alcance de melhores resultados. Complementarmente a Rede AVC Brasil, o Programa *Angels* lançado em 2018, é uma iniciativa internacional da Boehringer Ingelheim que busca qualificar os centros de AVC já existentes e auxiliar na implementação de novos centros. O *Angels Award* classifica os hospitais como Prontos para o atendimento do AVC – quando tem a estrutura mínima e monitora os dados – e aqueles com melhores desempenhos são classificados em ouro, platina e diamante (REDE BRASIL AVC, 2008). Estes hospitais recebem certificados deste status e são reconhecidos nos grandes congressos nacionais e internacionais. Mais de 150 hospitais no Brasil participam e mais de 2800 hospitais no mundo.

Por ser uma doença tempo-dependente, esforços de redesenho da linha de cuidado AVC, que permitam uma avaliação mais precisa e rápida, são foco de pesquisa em várias partes do mundo (MICHAEL E. PORTER; JAMES S MOUNT FORD; KAMALINI RAMDAS *et al.*, 2012; OURIQUES MARTINS; SACKS; HACKE *et al.*, 2019). No âmbito de VBHC, a literatura mostra abordagens de atendimento mais personalizado para as necessidades de pacientes individuais, reduzindo custos e restrições de recursos (W. DAVID FREEMAN ; KEVIN M. BARRETT ; LISA NORDAN *et al.*, 2018). Recentemente, avanços em detectar características do AVC, como severidade pela escala do *National Institute of Health Stroke Scale* (NIHSS), com técnicas de aprendizado de máquina foram exploradas por Kogan *et al.* (2020) (KOGAN; TWYMAN; HEAP *et al.*, 2020), no entanto os mesmos autores reconhecem que a falta de avaliação desses dados é uma limitação de evidência de mundo real dos pacientes AVC, e incentivam as tecnologias que respondam o conhecimento estratificado de risco do AVC e outras condições médicas.

A capacidade de monitoramento e controle da adoção das orientações estabelecidas nos manuais e o alinhamento com modelos de reembolso são desafios ainda a serem explorados (VAN VEGHEL; SCHULZ; VAN STRATEN *et al.*, 2018). Uma recente revisão sistemática teve como objetivo identificar e avaliar criticamente o relato e o desenvolvimento de modelos de aprendizado de máquina para prever resultados após AVC (WANG; KIIK; PEEK *et al.*, 2020). Apesar de contemplar apenas

dados estruturados, e não de texto, os autores reportaram que a maioria dos estudos abordam desfechos duros, como mortalidade e rastreamento da capacidade funcional pelo Rankin (mRS), e que a presença de dados desbalanceados (isto é, em que a variável de destino tem mais observações em uma classe específica do que as outras) penaliza o treinamento de algoritmos mais complexos. Ainda, os mesmos autores apontam que existe grande variabilidade entre os estudos no quesito tamanho amostral, e que nenhum dos estudos discutiu a implementação do modelo na vida real em prática clínica, embora o objetivo final seja, presumivelmente, ajudar os médicos a fazer decisões de tratamento e estimativa de prognósticos (WANG; KIIK; PEEK *et al.*, 2020). No mesmo estudo, os modelos de aprendizado de máquina mais utilizados, conforme relatado pelos autores Wang *et al.* (2020) (WANG; KIIK; PEEK *et al.*, 2020) foram *random forest*, *support vector machine* (SVM), árvore de decisão e redes neurais.

Neste sentido, pela relevância do AVC na população brasileira e das evidências sobre condutas adequadas que devem ser adotadas na fase aguda, a linha de cuidado do AVC isquêmico é objeto no estudo piloto deste projeto no que diz respeito a etapa de desenvolvimento de algoritmos para extração de desfechos em dados de texto livre a partir de prontuários eletrônicos.

3. OBJETIVOS

- Identificar na literatura quais medidas de desfechos são considerados em estudos que seguiram a agenda de valor, e analisar a origem dos dados usados para relatar a iniciativa de cuidado em saúde baseada em valor.
- Explorar modelos e regras que possibilitem a automatização da coleta de dados em prontuário eletrônico e seu potencial uso para analisar indicadores de desfechos de saúde.

4. REFERÊNCIAS BIBLIOGRÁFICAS

- AGRAWAL, A.; GOLDFARB, A.; GANS, J. **Prediction Machines: The Simple Economics of Artificial Intelligence**. 2018.
- ALLAHYARI, M.; KOCHUT, K. J.; JANIK, M. Ontology-based Text Classification into Dynamically Defined Topics. *In: IEEE International Conference on Semantic Computing*, 2014, p. 273-278. DOI: 10.1109/ICSC.2014.51.
- ALMEIDA, M. B.; BAX, M. P. **Uma visão geral sobre ontologias: pesquisa sobre definições, tipos, aplicações, métodos de avaliação e de construção**. *Ciência da Informação*, 32, p. 7-20, 2003.
- ANDRADE, C. M. d.; GONÇALVES, M. A. **Combining Representations For Effective Citation Classification**. Association for Computational Linguistics, 2020.
- ATEFA SHAH; BLASZCZYNSKI, E. **Value-based healthcare: A global assessment**. 2016.
- BALYAN, R.; CROSSLEY, S. A.; BROWN, W., 3rd; KARTER, A. J. *et al.* **Using natural language processing and machine learning to classify health literacy from secure messages: The ECLIPPSE study**. *PLoS One*, 14, n. 2, p. e0212488, 2019.
- BARBA, P. Machine Learning for Natural Language Processing. Lexalytics. 2019. Disponível em: <https://www.lexalytics.com/lexablog/machine-learning-natural-language-processing>.
- BEAM, A. L.; KOHANE, I. S. **Big Data and Machine Learning in Health Care**. *JAMA*, 319, n. 13, p. 1317, 2018/04/03/ 2018.
- BEAM, A. L.; KOHANE, I. S. **Big Data and Machine Learning in Health Care**. *JAMA*, 319, n. 13, p. 1317-1318, Apr 3 2018.
- BELL, J.; KILIC, C.; PRABAKARAN, R.; WANG, Y. Y. *et al.* **Use of electronic health records in identifying drug and alcohol misuse among psychiatric in-patients**. *The Psychiatrist*, 37, n. 1, p. 15-20, 2013.
- BLUMENTHAL, D.; TAVENNER, M. **The “Meaningful Use” Regulation for Electronic Health Records**. *New England Journal of Medicine*, 363, n. 6, p. 501-504, 2010/08/05/ 2010.
- BOAVENTURA, E. M. **Metodologia da pesquisa: monografia, dissertação, tese**. São Paulo: Atlas, 2004.
- BREIMAN, L. **Random Forests**. *Machine Learning*, 45, n. 1, p. 5-32, 2001/10/01 2001.
- BUTTE, A. J.; WEINSTEIN, D. A.; KOHANE, I. S. Enrolling patients into clinical trials faster using real time recruiting. *In: AMIA Annual Symp Proceedings*, 2000, p. 111-115.
- CANUTO, S.; MARTINS, W. S.; COUTO, T.; GONÇALVES, M. A. Efficient and Scalable MetaFeature-based Document Classification using Massively Parallel Computing. *In: In Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2015, p. p. 333--342.
- CANUTO, S.; SALLES, T.; GONÇALVES, M. A.; ROCHA, L. *et al.* On Efficient Meta-Level Features for Effective Text Classification. *In: In Proceedings of the 23rd ACM International Conference on Information and Knowledge Management -- CIKM*, 2014, p. p. 1709--1718.
- CAROPRESO, M. F.; MATWIN, S. Beyond the bag of words: A text representation for sentence selection. *In: In Conference of the Canadian Society for Computational Studies of Intelligence*, 2006, Germany. p. 324–335.

CAWLEY, G. C.; TALBOT, N. L. C. **On Over-fitting in Model Selection and Subsequent Selection Bias in Performance Evaluation.** *Journal of Machine Learning Research*, p. 2079–2107, 2010.

CHEN, W.; XIE, X.; WANG, J.; PRADHAN, B. *et al.* **A comparative study of logistic model tree, random forest, and classification and regression tree models for spatial prediction of landslide susceptibility.** *Catena*, 151, p. 147-160, 2017.

CHI, N.-W.; LIN, K.-Y.; HSIEH, S.-H. **Using ontology-based text classification to assist Job Hazard Analysis.** *Advanced Engineering Informatics*, 28, n. 4, p. 381-394, 2014.

CONSELHO FEDERAL DE MEDICINA. Resolução nº 1638 de 09 de Agosto de 2002. Define prontuário médico e torna obrigatória a criação da Comissão de Revisão de Prontuários nas instituições de saúde. Resolução nº 1638 de 09 de Agosto de 2002. MEDICINA, C. F. D. Brasília, DF: Diário Oficial da União: 184-185 p. 2002.

COOPER, R. **Target costing and value engineering.** Routledge, 2017. 1351411764.

CORTES, E. G.; WOLOSZYN, V.; BARONE, D. A. C., 2018, Cham. **When, Where, Who, What or Why? A Hybrid Model to Question Answering Systems.** Springer International Publishing. 136-146.

CRISTIANINI, N.; SHAWE-TAYLOR, J. **An introduction to Support Vector Machines: and other kernel-based learning methods.** Cambridge University Press, 2000. 0521780195.

CUNHA, W.; CANUTO, S.; VIEGAS, F.; SALLES, T. *et al.* **Extended pre-processing pipeline for text classification: On the role of meta-feature representations, sparsification and selective sampling.** *Information Processing & Management*, 57, n. 4, 2020.

CUNHA, W.; MANGARAVITE, V.; GOMES, C.; CANUTO, S. *e. et al.* **On the Cost-Effectiveness of Neural and Non-Neural Approaches and Representations for Text Classification: A Comprehensive Comparative Study (no prelo).** *Information Processing & Management*, 2020.

DAVIES, J.; FENSEL, D.; HARMELEN, F. **Towards the Semantic Web: Ontology-Driven Knowledge Management.** 03/01 2003.

DONABEDIAN, A. The quality of care. How can it be assessed? *In: Jama.* 1988/09/23 ed., 1988. v. 260, p. 1743-1748.

DONABEDIAN, A. **An Introduction to Quality Assurance in Health Care.** 2002. (Oxford University Press. 9780195158090.

DRIESSEN, S. R.; VAN ZWET, E. W.; HAAZEBROEK, P.; SANDBERG, E. M. *et al.* **A dynamic quality assessment tool for laparoscopic hysterectomy to measure surgical outcomes.** *Am J Obstet Gynecol*, 215, n. 6, p. 754 e751-754 e758, Dec 2016.

ETGES, A. P. B. d. S.; RUSCHEL, K. B.; POLANCZYK, C. A.; URMAN, R. D. **Advances in Value-Based Healthcare by the Application of Time-Driven Activity-Based Costing for Inpatient Management: A Systematic Review.** *Value in Health*, 23, n. 6, p. 812-823, 2020/06// 2020.

EY. Dados Ernest Young. 2019.

FALCONER, N.; NAND, S.; LIOW, D.; JACKSON, A. *et al.* **Development of an electronic patient prioritization tool for clinical pharmacist interventions.** *Am J Health Syst Pharm*, 71, n. 4, p. 311-320, Feb 15 2014.

GARLA, V. N.; BRANDT, C. **Ontology-guided feature engineering for clinical text classification.** *J Biomed Inform*, 45, n. 5, p. 992-998, Oct 2012.

GEORGAKOPOULOS, S. V.; TASOULIS, S. K.; VRAHATIS, A. G.; PLAGIANAKOS, V. P. Convolutional Neural Networks for Toxic Comment Classification. *In: Proceedings of the 10th Hellenic*

Conference on Artificial Intelligence, 2018, Patras, Greece. Association for Computing Machinery, p. Article 35. DOI: 10.1145/3200947.3208069. Disponível em: <https://doi.org/10.1145/3200947.3208069>.

GLASER, J. It's Time for a New Kind of Electronic Health Record. Harvard Business Review. ANALYTICS: Harvard Business Print 2020.

GOLAS, S. B.; SHIBAHARA, T.; AGBOOLA, S.; OTAKI, H. *et al.* **A machine learning model to predict the risk of 30-day readmissions in patients with heart failure: a retrospective analysis of electronic medical records data.** BMC Medical Informatics and Decision Making, 18, n. 1, p. 44, 2018/06/22 2018.

GOLDBERG, Y.; LEVY, O. **Word2vec explained: Deriving mikolov et al.'s negative-sampling word-embedding method.** arXiv, 2014.

GOMEZ-PEREZ, A.; CORCHO, O.; FERNÁNDEZ-LÓPEZ, M. **Ontological engineering: with examples from the areas of knowledge management, e-Commerce and the Semantic Web. (advanced information and knowledge processing).** 01/01 2004.

GOPAL, G.; SUTER-CRAZZOLARA, C.; TOLDO, L.; EBERHARDT, W. **Digital transformation in healthcare - Architectures of present and future information technologies.** Clinical Chemistry and Laboratory Medicine (CCLM), 57, 12/11 2018.

GROSSMAN, E.; CARDOSO, M. H. C. d. A. **As narrativas em medicina: contribuições à prática clínica e ao ensino médico.** Rev. bras. educ. med., 30, n. 1, p. 6-14, 2006.

HAN, E. H. S.; KARYPIS, G. Centroid-based document classification: Analysis and experimental results. *In:* In European Conference on Principles of Data Mining and Knowledge Discovery; Springer: Berlin/Heidelberg,, 2000, Berlin/Heidelberg, Germany. p. 424-431.

HERNANDEZ A.; KAPLAN R.S.; WITKOWSKI M.L.; FORREST FAISON C. *et al.* **Navy Medicine Introduces value-based health care.** Health affairs (Project Hope), 8, 38(8), n. 8, p. 1393-1400, 2019.

HONNIBAL, M.; MONTANI, I. **spaCy.** Explosion AI, 2015. spaCy is an open-source software library for advanced natural language processing, written in the programming languages Python and Cython.

HYUN, S.; JOHNSON, S., B.; BAKKEN, S. **Exploring the ability of natural language processing to extract data from nursing narratives.** . Comput Inform Nurs, 27, n. 4, p. 215-225, 2009.

ICHOM; 2012. **ICHOM 2012 | Healthcare Improvement | Patient-Reported Outcomes.** Disponível em: <https://www.ichom.org/files/28502/www.ichom.org.html>.

IORIO, R.; CLAIR, A. J.; INNEH, I. A.; SLOVER, J. D. *et al.* **Early Results of Medicare's Bundled Payment Initiative for a 90-Day Total Joint Arthroplasty Episode of Care.** The Journal of Arthroplasty, 31, n. 2, p. 343-350, 2016/02// 2016.

JAGANNATHA, A.; LIU, F.; LIU, W.; YU, H. **Overview of the First Natural Language Processing Challenge for Extracting Medication, Indication, and Adverse Drug Events from Electronic Health Record Notes (MADE 1.0).** Drug Saf, 42, n. 1, p. 99-111, Jan 2019.

JAMES, G.; WITTEN, D.; TIBSHIRANI, T. H. R. **An Introduction to Statistical Learning.** New York Heidelberg Dordrecht London: Springer, 2013. 181-184 p.

JOHNSON, C. O.; NGUYEN, M.; ROTH, G. A.; NICHOLS, E. *et al.* **Global, regional, and national burden of stroke, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016.** The Lancet Neurology, 18, n. 5, p. 439-458, 2019.

JOHNSON, R.; ZHANG, T. Deep Pyramid Convolutional Neural Networks for Text Categorization. *In:* Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics, 2017, Vancouver, Canada. p. 567-570.

KATZ, M.; FRANKEN, M.; MAKDISSE, M. **Value-Based Health Care in Latin America: An Urgent Discussion.** *J Am Coll Cardiol*, 70, n. 7, p. 904-906, Aug 15 2017.

KLIE, J.-C.; BUGERT, M.; BOULLOSA, B.; ECKART DE CASTILHO, R. *et al.* **The INCEPTION Platform: Machine-Assisted and Knowledge-Oriented Interactive Annotation.** In Proceedings of System Demonstrations of the 27th International Conference on Computational Linguistics (COLING 2018): Santa Fe, New Mexico, USA 2018.

KNOX; RITA; EID, T.; WHITE, A. **“Management Update: Companies should align their structured and unstructured dat.** Gartner Research, 2005.

KOGAN, E.; TWYMAN, K.; HEAP, J.; MILENTIJEVIC, D. *et al.* **Assessing stroke severity using electronic health record data: a machine learning approach.** *BMC Med Inform Decis Mak*, 20, n. 1, p. 8, Jan 8 2020.

KOWSARI; JAFARI, M.; HEIDARYSAFA; MENDU *et al.* **Text Classification Algorithms: A Survey.** *Information*, 10, n. 4, 2019.

KREIMEYER, K.; FOSTER, M.; PANDEY, A.; ARYA, N. *et al.* **Natural language processing systems for capturing and standardizing unstructured clinical information: A systematic review.** *J Biomed Inform*, 73, p. 14-29, Sep 2017.

KUHN, M.; JOHNSON, K. **Applied Predictive Modeling.** Springer, 2013. 978-1-4614-6848-6.

LARSON, R. R. **Introduction to Information Retrieval.** *Journal of the American Society for Information Science and Technology*, p. n/a-n/a, 2009.

LAYEGHIAN JAVAN, S.; SEPEHRI, M. M.; AGHAJANI, H. **Toward analyzing and synthesizing previous research in early prediction of cardiac arrest using machine learning based on a multi-layered integrative framework.** *J Biomed Inform*, 88, p. 70-89, Dec 2018.

LECUN, Y.; BENGIO, Y.; HINTON, G. **Deep learning.** *Nature*, 521, n. 7553, p. 436-444, May 28 2015.

LEE, T. H. **Putting the Value Framework to Work.** *New England Journal of Medicine*, 363, n. 26, p. 2481-2483, 2010/12/23/ 2010.

LEUNG, T. I.; VAN MERODE, G. G. Value-Based Health Care Supported by Data Science. *In: KUBBEN, P.; DUMONTIER, M., et al (Ed.). Fundamentals of Clinical Data Science.* Cham (CH): Springer Copyright 2019, The Author(s). 2019. p. 193-212.

LI, L.; WEINBERG, C. R.; DARDEN, T. A.; PEDERSEN, L. G. **Gene selection for sample classification based on gene expression data: study of sensitivity to choice of parameters of the GA/KNN method.** *Bioinformatics*, 17, n. 12, p. 1131-1142, Dec 2001.

LING, A. Y.; KURIAN, A. W.; CASWELL-JIN, J. L.; SLEDGE, G. W., Jr. *et al.* **Using natural language processing to construct a metastatic breast cancer cohort from linked cancer registry and electronic medical records data.** *JAMIA Open*, 2, n. 4, p. 528-537, Dec 2019.

LOPES, L.; VIEIRA, R.; FINATTO, M. J.; MARTINS, D. *et al.* **Automatic extraction of composite terms for construction of ontologies: an experiment in the health care area.** *Reciis*, 3, n. 1, 2009.

LYNCH, S. M.; MOORE, J. H. **A call for biological data mining approaches in epidemiology.** *BioData Mining*, 9, n. 1, p. 1, 2016/01/04 2016.

Machine Learning for Natural Language Processing. Lexalytics, 2019/11/25/T07:00:53+00:00 2019.

MAKDISSE, M.; KATZ, M.; RAMOS, P.; PEREIRA, A. *et al.* **What Is a Value Management Office? An Implementation Experience in Latin America.** *Value in Health Regional Issues*, 17, n. (Makdisse

M., marcia.makdisse@einstein.br; Katz M.; Ramos P.; Pereira A.; Shiramizo S.; Neto M.C.; Klajner S.) Hospital Israelita Albert Einstein, São Paulo, Brazil, p. 71-73, 2018 2018.

MANEVITZ, L. M.; YOUSEF, M. **One-class SVMs for document classification**. J. Mach. Learn, 2, p. 139-154, 2001.

MANNING, C.; RAGHAWAN, P.; SCHUTZE, H. **An introduction to information retrieval**. Cambridge, U.K.: Cambridge University Press, 2009.

MANNING, C.; SCHUTZE, H. **Foundations of statistical natural language processing**. MIT Press: Cambridge, MA., 1999.

MESKÓ, B.; GÖRÖG, M. **A short guide for medical professionals in the era of artificial intelligence**. npj Digital Medicine, 3, n. 1, 2020.

MESKÓ, B.; RADÓ, N. **Data annotators: the unsung heroes of artificial intelligence development.**, 2019. Disponível em: <https://medicalfuturist.com/data-annotation/>.

MEULEN, M. v. d. **Artificial Intelligence as a Driver of Value in Value Based Health Care Systems**. 2019.

MEYSTRE, S. M.; KIM, Y.; GOBBEL, G. T.; MATHENY, M. E. *et al.* **Congestive heart failure information extraction framework for automated treatment performance measures assessment**. J Am Med Inform Assoc, 24, n. e1, p. e40-e46, Apr 1 2017.

MICHAEL E. PORTER; JAMES S MOUNT FORD; KAMALINI RAMDAS; TAKVORIAN, S. **Reconfiguring Stroke Care in North Central London**. 2012.

MICHAEL E. PORTER; KAPLAN, R. S. How to pay for Health Care. Harvard Business Review. : HBR Reprint 2016.

MINISTÉRIO DA SAÚDE. Manual de Rotinas Para Atenção do AVC. 2013a.

MINISTÉRIO DA SAÚDE. Protocolo Clínico e Diretrizes Terapêuticas: Trombólise No Acidente Vascular Cerebral Isquêmico Agudo. 2013b.

MINISTÉRIO DA SAÚDE. Uma Análise de Situação de Saúde e Das Doenças e Agravos Crônicos: Desafios e Perspectiva. MINISTÉRIO DA SAÚDE, B. Brasília: MS 2018.

MIOTTO, R.; WENG, C. **Case-based reasoning using electronic health records efficiently identifies eligible patients for clinical trials**. J Am Inform Assoc, 1, 22, p. 141-150, 2015.

NI, Y.; WRIGHT, J.; PERENTESIS, J.; LINGREN, T. *et al.* **Increasing the efficiency of trial-patient matching: automated clinical trial eligibility pre-screening for pediatric oncology patients**. BMC Med Inform Decis Mak, 15, p. 28, Apr 14 2015.

NORIA, S.; NEEDLEMAN, B.; MIKAMI, D.; LAYTON, K. *et al.* **The impact of care coaching on hospital length of stay, readmission rates and overall patient satisfaction post-bariatric surgery**. Surgery for Obesity and Related Diseases, 11, n. 6, p. S8-S9, 2015 2015.

OURIQUES MARTINS, S. C.; SACKS, C.; HACKE, W.; BRAININ, M. *et al.* **Priorities to reduce the burden of stroke in Latin American countries**. The Lancet Neurology, 18, n. 7, p. 674-683, 2019.

PBGH. **Variation in NTSV C-section rates among California hospitals**. Pacific Business Group on Health. LeapFrog Group. 2014.

PEDREGOSA; ET AL. **Scikit-learn: Machine Learning in Python.**, 2011. pp. 2825-2830.

PENNINGTON, J.; SOCHER, R.; MANNING, C. D.; DOHA, Q. Glove: Global Vectors for Word Representation. *In: In Proceedings of the 2014 Conference on Empirical Methods in Natural Language*

Processing (EMNLP), 2014, **14**. p. 1532–1543.

PESTIAN, J.; NASRALLAH, H.; MATYKIEWICZ, P.; BENNETT, A. *et al.* **Suicide Note Classification Using Natural Language Processing: A Content Analysis**. Biomedical Informatics Insights, 3, 2010.

PORTER, M. E. **What Is Value in Health Care?** New England Journal of Medicine, 363, n. 26, p. 2477-2481, 2010/12/23/ 2010.

PORTER, M. E., Alee, H., Witkowski, M.L. and Feeley, Thomas W. . **“Navy Medicine: Moving to Value-Based Care”** Harvard Business School,. 2019.

PORTER, M. E., and Elizabeth O. Teisberg. **"Cleveland Clinic: Transformation and Growth 2015."**. Harvard Business School Case: February 2009, p. 709-473. 2009.

PORTER, M. E., Jens Deerberg-Wittram, and Clifford Marks. **"Martini Klinik: Prostate Cancer Care."** Harvard Business School Case Collection, p. 714-471. 2014.

PORTER, M. E.; LARSSON, S.; LEE, T. H. **Standardizing Patient Outcomes Measurement**. N Engl J Med, 374, n. 6, p. 504-506, Feb 11 2016.

PORTER, M. E.; LEE, T. H. **The Strategy That Will Fix Health Care**. Harvard Business Review, 91, n. 10, p. 50-70, 2013/10// 2013.

PORTER, M. E.; TEISBERG, E. O. **Redefining health care: creating value-based competition on results**. Boston, Mass: Harvard Business School Press, 2006 2006. 506 p. 978-1-59139-778-6.

PORTER, M. E., Thomas H. Lee, and Meredith A. Alger. **"Oak Street Health: A New Model of Primary Care."** Harvard Business School Case : February 2017, p. 717-437. 2017.

RAGHAVAN, P.; CHEN, J. L.; LUSSIER, E. F.; LAI, A. M. How essential are unstructured clinical narratives and information fusion to clinical trial recruitment? . *In: AMIA Jt Summits Transl Sci Proc.*, 2014, p. 218-223. Disponível em: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4333685/pdf/1861388.pdf>.

REDE BRASIL AVC. **RedeBrasilAVC**. 2008. Disponível em: <http://www.redebrasilavc.org.br>.

ROSENBLOOM, S. T.; DENNY, J. C.; XU, H.; LORENZI, N. *et al.* **Data from clinical notes: a perspective on the tension between structure and flexible documentation**. Journal of the American Medical Informatics Association, 18, n. 2, p. 181-186, 2011.

RUIZ, E. E. S. **Mineração de texto em saúde**. 2016.

RUSSELL, S. J. **Artificial intelligence : a modern approach**. New Jersey: 1995. (Artificial intelligence. 013103805-2.

RUSSELL, S. J.; NORVIG, P. **Artificial Intelligence: A Modern Approach**. Englewood Cliffs, New Jersey: Alan Apt, 2016.

S., C.; X., S. D.; A., G. M.; C., R. T. **A Thorough Evaluation of Distance-Based Meta-Features for Automated Text Classification**. IEEE Transactions on Knowledge and Data Engineering, 30, n. 12, p. 2242-2256, 2018.

SALTON, G.; BUCKLEY, C. **Term-weighting approaches in automatic text retrieval**. . Inf. Process. Manag, 24, p. 513-523, 1988.

SHEIKHALISHAHI, S.; MIOTTO, R.; DUDLEY, J. T.; LAVELLI, A. *et al.* **Natural Language Processing of Clinical Notes on Chronic Diseases: Systematic Review**. JMIR Med Inform, 7, n. 2, p. e12239, Apr 27 2019.

SHEN, D.; SUN, J. T.; LI, H.; YANG, Q. *et al.* **Document Summarization Using Conditional Random Fields.** 7, p. 2862-2867, 2007.

SHEN, Y. **Selection incentives in a performance-based contracting system.** Health Serv Res, 38, n. 2, p. 535-552, Apr 2003.

SHIVADE, C.; HEBERT, C.; LOPETEGUI, M.; MARNEFFE *et al.* **Textual inference for eligibility criteria resolution in clinical trials.** J Biomed Inform, S211-S218, 58, 2015.

SILVA, A. M. P. d. Dissertação de Mestrado: Processamento de Linguagem Natural na Identificação de Critérios de Elegibilidade para a Pesquisa Clínica SAÚDE, P. d. P. G. d. T. e. : Pontifica Universidade Católica do Paraná: Escola Politécnica 2018.

SUN, W.; CAI, Z.; LI, Y.; LIU, F. *et al.* **Data Processing and Text Mining Technologies on Electronic Medical Records: A Review.** J Healthc Eng, 2018, p. 4302425, 2018.

THAKER, N. G.; PUGH, T. J.; MAHMOOD, U.; CHOI, S. *et al.* **Defining the value framework for prostate brachytherapy using patient-centered outcome metrics and time-driven activity-based costing.** Brachytherapy, 15, n. 3, p. 274-282, 2016/05// 2016.

TRANFIELD, D.; DENYER, D.; SMART, P. **Towards a Methodology for Developing Evidence-Informed Management Knowledge by Means of Systematic Review.** British Journal of Management, 14, n. 3, p. 207-222, 2004.

TSAI, M. H.; PORTER, J. C.; ADAMS, D. C. **The Denominator in Value-Based Health Care: Porter's Hidden Costs.** Anesth Analg, 127, n. 1, p. 317, Jul 2018.

UZUNER, Ö. e. a. **Challenge on concepts, assertions, and relations in clinical text.** Journal of the American Medical Informatics Association, 18, n. 5, p. 552-556, 2011.

VALUCK, T.; BLAISDELL, D.; DUGAN, D. P.; WESTRICH, K. *et al.* **Improving Oncology Quality Measurement in Accountable Care: Filling Gaps with Cross-Cutting Measures.** J Manag Care Spec Pharm, 23, n. 2, p. 174-181, Feb 2017.

VAN CALSTER, B.; VICKERS, A. J. **Calibration of risk prediction models: impact on decision-analytic performance.** Med Decis Making, 35, n. 2, p. 162-169, Feb 2015.

VAN DEN BERG, K.; DIJSMAN, L. M.; KEUS, S. H. J.; SCHEELE, F. *et al.* **Value-based health care in obstetrics.** Journal of Evaluation in Clinical Practice, 26, n. 1, p. 101-108, 2020 2020.

VAN LEIJEN-ZEELENBERG, J. E.; ELISSEN, A. M. J.; GRUBE, K.; VAN RAAK, A. J. A. *et al.* **The impact of redesigning care processes on quality of care: a systematic review.** BMC Health Serv Res, 16, p. 19, Jan 19 2016.

VAN VEGHEL, D.; SCHULZ, D. N.; VAN STRATEN, A. H. M.; SIMMERS, T. A. *et al.* **Health insurance outcome-based purchasing: The case of hospital contracting for cardiac interventions in the Netherlands.** International Journal of Healthcare Management, 11, n. 4, p. 371-378, 2018.

VANWINCKELEN, G. On Estimating Model Accuracy with Repeated Cross-Validation. *In:* Belgian-Dutch Conference on Machine Learning (BeneLearn), 2012, Ghent.

VELUPILLAI, S.; SUOMINEN, H.; LIAKATA, M.; ROBERTS, A. *et al.* **Using clinical Natural Language Processing for health outcomes research: Overview and actionable suggestions for future advances.** J Biomed Inform, 88, p. 11-19, Dec 2018.

VIEIRA, R.; LIMA, V. L. S. d. Linguística computacional: princípios e aplicações. IX Escola de Informática da SBC-Sul.: SBC-Sul 2001.

W. DAVID FREEMAN ; KEVIN M. BARRETT ; LISA NORDAN ; AARON C. SPAULDING *et al.* **Lessons from Mayo Clinic's Redesign of Stroke Care.** Harvard Business Review, n. Operations, 2018.

WANG, B.; MCKAY, R.; ABBASS, H.; BARLOW, M. **A Comparative Study for Domain Ontology Guided Feature Extraction.** 2003. 69-78 p.

WANG, S. Y.; PERSHING, S.; TRAN, E.; HERNANDEZ-BOUSSARD, T. **Automated extraction of ophthalmic surgery outcomes from the electronic health record.** International Journal of Medical Informatics, 133, p. 104007, 2020/01// 2020.

WANG, S. Y.; PERSHING, S.; TRAN, E.; HERNANDEZ-BOUSSARD, T. **Automated extraction of ophthalmic surgery outcomes from the electronic health record.** Int J Med Inform, 133, p. 104007, Jan 2020.

WANG, W.; KIIK, M.; PEEK, N.; CURCIN, V. *et al.* **A systematic review of machine learning models for predicting outcomes of stroke with structured data.** PLoS One, 15, n. 6, p. e0234722, 2020.

WANG, Y.; WANG, L.; RASTEGAR-MOJARAD, M.; MOON, S. *et al.* **Clinical information extraction applications: A literature review.** J Biomed Inform, 77, p. 34-49, Jan 2018.

XIAO, C.; CHOI, E.; SUN, J. **Opportunities and challenges in developing deep learning models using electronic health records data: a systematic review.** Journal of the American Medical Informatics Association, 25, n. 10, p. 1419-1428, 2018/10/01/ 2018.

XIAO, C.; CHOI, E.; SUN, J. **Opportunities and challenges in developing deep learning models using electronic health records data: a systematic review.** J Am Med Inform Assoc, 25, n. 10, p. 1419-1428, Oct 1 2018.

XU, B. G., X.; YE, Y.; CHENG, J. **An Improved Random Forest Classifier for Text Categorization.** . JCP 2012, 7, p. 2913-2920.

YANG, Y.; GOPAL, S. **Multilabel classification with meta-level features in a learning-to-rank framework.** JMLR, 88, p. 47-68, 2012.

YE, F., 2016, **Simultaneous Support Vector selection and parameter optimization using Support Vector Machines for sentiment classification.** 59-62.

YIN, R. K. **Case study research: design and methods.** Los Angeles London New Delhi Singapore Washington, DC: SAGE, 2014.

ZHANG, C. **Automatic keyword extraction from documents using conditional random fields.** J. Comput. Inf. Syst, 4, p. 1169-1180, 2008.

5. ARTIGO(S)

Os artigos a seguir estão no formato e normas da revista as quais serão submetidos.

6. CONCLUSÕES E CONSIDERAÇÕES FINAIS

A partir da contextualização sobre a importância do avanço em capacidade tecnológica informacional para suportar modelos de avaliação de valor em saúde, esta pesquisa se dispôs a avaliar a temática de modo exploratório.

A revisão da literatura realizada deste trabalho atendeu ao objetivo específico de compreender como os resultados em saúde estão sendo medidos em casos do mundo real, e analisar a origem dos dados usados para relatar a iniciativa de cuidado em saúde baseada em valor. Os resultados do primeiro artigo sugerem que ainda há uma lacuna na medição dos resultados que são importantes para os pacientes e na medição dos resultados financeiros por meio de métodos metodológicos rigorosos. Avanços na capacidade tecnológica e uma cultura de mudança na gestão parecem ser as principais barreiras para tornar a agenda de valor mais facilmente reproduzível, e ainda são requisitos necessários para que os itens acima ocorram.

O segundo artigo apresenta os melhores métodos a serem investidos na classificação de texto a partir de dados não estruturados de prontuários eletrônicos para medir de forma automatizada alguns desfechos de saúde de pacientes que sofreram AVC. O framework demonstra visualmente os modelos que apresentam melhor desempenho para avaliar cada tarefa. Com isso é possível orientar e impulsionar a adequação de parâmetros que vão ao encontro do que se busca na gestão por valor e a adoção de indicadores que permitam o seu controle.

A escrita deste trabalho e o interesse pela temática só foi possível devido a oportunidade de participar e auxiliar na coleta de dados de projeto “Proposição de um modelo de gestão de saúde baseada em valor para saúde suplementar e pública do Brasil”. As discussões e experimentos em avanços em capacidade tecnológica para estudos de microcusteio despertaram motivação para explorar o outro fator da equação de valor trabalhado nesta dissertação. Em parte, a construção desta parte da pesquisa foi desencadeada pelo interesse em inovação em saúde e tentativa da união de áreas de ensino e pesquisa que juntas podem atender a um interesse comum no campo da saúde ponderada pelo valor.

Em uma das primeiras reuniões sobre o projeto nos foi dito que “(..) *os modelos computacionais não podem trazer mais trabalho para o prestador, os algoritmos precisam aprender a interpretar os dados tal como eles estão sendo colocados no dia a dia, por isso*

que se chama 'inteligência artificial' (...)". A partir daí pessoas com formações e bagagens diferentes desenvolveram relação de equipe e propósito comum. Por isso finalizo as considerações finais sobre este trabalho com a reflexão do primeiro passo dado como grupo de pesquisa em prol de avanços de capacidade tecnológica para programas de valor em saúde, e que isso inove e reflita positivamente no nosso papel como pesquisador no Brasil.

7. LIMITAÇÕES DO ESTUDO

Existem limitações a serem consideradas. A primeira delas foi o número de pacientes que compreendeu a nossa amostra, embora seja uma característica do banco de dados ser um texto muito desbalanceado em termos de informações extraídas, um número maior de pacientes poderia gerar classes mais homogêneas. Em segundo lugar, embora em termos de sentença seja um grande volume, não testamos a performance dos modelos a nível de evolução e reconhecemos que algumas variáveis poderiam ter uma performance mais adequada se fossem avaliadas a nível de evolução e não de sentença. Uma terceira limitação, é que até o momento foram usados dados de apenas um hospital, sendo difícil fazer alguma inferência sobre a generalizabilidade dos resultados encontrados.

Técnicas de aprendizado de máquina prometem resultados satisfatórios na extração de informações clínicas, mas para isso faz-se necessário um corpus anotado para treinamento de algoritmos. *Corpus* para textos em português são raros, especialmente focados na área da saúde (BUTTE; WEINSTEIN; KOHANE, 2000; JAGANNATHA; LIU; LIU *et al.*, 2019; MIOTTO; WENG, 2015; RAGHAVAN; CHEN; LUSSIER *et al.*, 2014) A grande maioria das ferramentas de PLN são treinadas com corpus jornalístico o que dificulta a identificação de termos pela falta de similaridade entre eles.

A baixa disponibilidade de dados de registros clínicos também pode estar associada a falta de padrão na escrita dos atendimentos realizados pelos profissionais. Os textos utilizados para aplicação do método não seguiam nenhum padrão de escrita e de estrutura, apresentavam erros gramaticais e ortográficos, diferentemente dos trabalhos desenvolvidos por Friedlin e McDonald (2006) e Long (2007) que extraíram informações a partir de um cenário ideal de estrutura e escrita.

8. TRABALHOS FUTUROS

Como perspectivas de trabalhos futuros e maturidade da pesquisa, entende-se a necessidade de expandir o uso dos algoritmos com melhor performance para outras linhas de cuidado e hospitais. Melhorias contínuas de capacidade tecnológica informacional para garantir a disponibilidade de dados de resultados em todo o ciclo de atendimento, e a instituição com cultura baseada em valor entre os provedores são essenciais para conduzir a implementação do VBHC.

Os modelos gerados por esta pesquisa utilizaram dados de um hospital privado do sul do país, o uso de novos dados da mesma instituição pode ampliar a aplicação dos modelos para outras linhas de cuidado. Abre-se espaço, por exemplo, para pesquisa no âmbito do mapeamento epidemiológico da instituição com base nas morbidades da população, controle de escalas de risco, suporte para vigilância de eventos de risco, integração com dados de performance e desfechos reportados pelo paciente. E, dessa forma, contribuir para o desenvolvimento de novas pesquisas de VBHC em países de língua portuguesa.

Em paralelo, entende-se a aplicação dos algoritmos em outros hospitais irá contribuir com a generalizabilidade dos modelos de monitoramento de resultados para linha de cuidado do AVC, facilitando o uso de indicadores que corroborem com uma maior transparência do desempenho das instituições para a linha de cuidado específica. Para isso são necessários mais dados manualmente anotados para conferência em diferentes instituições. Para o aperfeiçoamento do algoritmo de expressão regular, o grupo de pesquisa em recuperação de informação prossegue os estudos na construção de um *corpus* anotado em português para auxiliar no processo de análise semântica de textos clínicos. Isso também abre espaço para uso de plataformas *open-source* de anotação de dados em larga escala, como *Inception* ou *Webanno*.

E finalmente, vê-se a oportunidade da criação de *dashboards* ou plataformas *Business Intelligence* que permitam a visualização desses dados a nível do gestor hospitalar. Visualizar os dados em tempo real da jornada de cuidado e a nível de paciente é um dos grandes objetivos de implementação de VBHC em uma instituição (PORTER; LEE, 2013), que permite maior transparência e objetividade nas comparações entre instituições. Por isso, esforços que permitam tornar a saída desses algoritmos passíveis de serem implementadas na prática é outra tarefa a seguir.

APÊNDICE A - DETALHES METODOLÓGICOS ADICIONAIS

Esta seção inicia-se com as considerações éticas da pesquisa, tipo de estudo, local onde a pesquisa foi desenvolvida e cenário da pesquisa. Posteriormente são apresentadas as etapas e design do experimento realizado.

1. Considerações Éticas da Pesquisa

A disponibilização da base de dados para o grupo de pesquisa de recuperação de informação em saúde ocorreu após aprovação do Comitê de Ética em Pesquisa (CEP) do Hospital Moinhos de Vento de Porto Alegre (HMV), sob parecer nº 29694720.0.0000.5330 obtido em 06 de maio de 2020 (ANEXO A)

Este projeto também atende aos preceitos éticos, diretrizes e normas regulamentadoras em pesquisa envolvendo seres humanos estabelecidos na Lei Geral de Proteção de Dados Pessoais 13.853 de 8 de julho de 2019 que prevê, entre outros preceitos, a de-identificação das informações confidenciais de pacientes, familiares e profissionais da base de dados utilizada. Pacientes não serão identificados individualmente em nenhuma análise que será realizada e os dados das evoluções serão extraídos dos sistemas de informação seguindo os processos internos estabelecidos.

O acesso à base completa de dados para análise dos resultados será apenas dos pesquisadores do HMV que assinaram os termos de compromisso com o uso de dados. Os dados coletados serão mantidos em sigilo a fim de assegurar a privacidade e o anonimato dos participantes. As informações obtidas foram utilizadas apenas para fins científicos vinculados a este projeto.

2. Tipo de estudo

Esta pesquisa, por estar orientada a propor soluções de um problema específico em ambiente real, apresenta natureza exploratória, a qual proporciona um aprofundamento na temática, ainda relativamente pouco explorada, de Gestão de Saúde Baseada em Valor, ao estudá-la no contexto de mensuração de desfechos (BOAVENTURA, 2004; TRANFIELD; DENYER; SMART, 2004; YIN, 2014). No que diz respeito à abordagem, tem-se uma discussão qualitativa e quantitativa pois, ao mesmo tempo em que se analisa de maneira qualitativa quais parâmetros e indicadores são relevantes para uso na linha de cuidado, propõe-se um modelo que permite automatização de análises em registros médicos eletrônicos e a eficácia deste modelo.

3. Local de estudo

Esta é uma pesquisa que faz parte do projeto intitulado “Proposição de modelo de gestão de saúde baseada em valor para os sistemas de saúde pública e suplementar do Brasil” proposto pelo Instituto de Avaliação de Tecnologias em Saúde (IATS/INCT) e Universidade Federal do Rio Grande do Sul (UFRGS), em parceria com alunos de grupos de pesquisa da Universidade Federal de Minas Gerais (UFMG) e da Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS).

4. Cenário da Pesquisa

Este estudo está consolidando dados do hospital proponente da pesquisa. O estudo foi desenvolvido em um hospital privado de alta qualidade do Brasil, credenciado pela *Joint Comission International*, sendo classificado como o 5º melhor hospital do país. Desde 2017, o hospital introduziu a coleta de dados dos conjuntos de padrões ICHOM em sua rotina para diferentes caminhos clínicos e criou o escritório de valor institucional. Para o caminho do AVC, a coleta de dados teve início em 15 de outubro de 2015 e, desde agosto de 2018, está estruturada no Redcap e integrado ao Escritório de Valor do hospital. Em 2019, o hospital passou a incorporar o *Angel Awards Program*, sendo certificado com a categoria Platinum ao final do primeiro ano.

Foram levantados os números de prontuário para recuperação dos dados de evolução texto de pacientes com AVC isquêmico internados no hospital de agosto de 2018 a março de 2020. Foram anotadas 44206 sentenças, que corresponde 1558 evoluções de 188 pacientes. A plataforma de prontuários eletrônicos da instituição usa o sistema de prontuários MV.

Antes do processamento dos dados de evolução para anotação das variáveis, foi necessário uma limpeza e preparação da base de dados, que consistiu na remoção dos registros que não fossem da internação correspondente ao AVCi do paciente – que eventualmente poderia ter internado com suspeita do diagnóstico mas não ter sido confirmado durante a internação, ou ainda paciente que internou mais de uma vez nesta instituição no mesmo período.

5. Pré-processamento

O pré-processamento ocorreu em três momentos: i) de-identificação das narrativas clínicas, ii) aleatorização das evoluções, e iii) quebra de sentença.

A de-identificação representa a retirada de informações confidenciais e sigilosas da base de dados e foi feita por meio de um algoritmo de expressão regular que permite a busca por padrões de escrita para substituição de trechos ou caracteres específicos nos textos. A equipe de pesquisadores responsável pela programação dos algoritmos obteve a relação de todos os pacientes e nome de profissionais para que fosse possível essa substituição. Entretanto, alguns problemas surgiram durante a interpretação do algoritmo em virtude da ambiguidade dos nomes. Pelo fato de alguns nomes da listagem coincidirem com pronomes e substantivos que não deveriam ser de-identificados como, por exemplo, Tala, Clara, Bem e Pelo. Desta forma, algumas informações como pronomes de tratamento, apelidos, abreviações de registros de classes profissionais não foram de-identificadas.

As expressões regulares utilizadas para a de-identificação dessas informações são substituídas no texto da seguinte maneira: “NOME_PACIENTE” e “NOME_COMPLETO_PACIENTE”. Os dados de evolução foram aleatorizados antes de serem entregues para anotação aos pesquisadores para abranger um maior número de pacientes durante as anotações, e para não viciar a marcação do anotador por já estar familiarizado com a descrição de um mesmo paciente.

Foi utilizada a estratégia de quebra das evoluções em sentenças baseados nos pressupostos de que as informações das variáveis estão presentes em um espaço curto do texto, o uso de grandes quantidades de dados pode facilitar aprendizado preciso da semântica (CAROPRESO; MATWIN, 2006), redução de ruídos (textos em branco, não informativos)(MANNING; RAGHAWAN; SCHUTZE, 2009; MANNING; SCHUTZE, 1999), e é uma estratégia comum em plataformas de anotação semântica (KLIE; BUGERT; BOULLOSA *et al.*, 2018). A quebra de sentenças pode ser feita de diversas formas, por pontuação, número de palavras, uso de *stop words*, entre outros, a nossa estratégia incluiu o uso da biblioteca spaCy, que é uma biblioteca de código aberto gratuita para PLN avançado em Python. Foi escolhida por apresentar a melhor performance em tarefas de quebra de sentença em textos longos e mal formatados (HONNIBAL; MONTANI, 2015), como é o caso

de evoluções clínicas, e por fazer uma quebra baseada em semântica e não pontuação, como seria uma humana.

Uma limitação importante da quebra de sentenças é a quebra errônea, por exemplo, quebras de sentença em locais que não fazem sentido gerando um dado de má qualidade. Dado de má qualidade é um dado incoerente e sem sentido, e que por vezes quando em formato de sentenças perdem seu significado, isso pode implicar em falhas nas anotações não pela ausência da informação e sim pela quebra que não foi informativa o suficiente.

5.1 Variação das Terminologias

A verificação da variação de terminologias em meio as narrativas clínicas são relevantes para este estudo, pois os termos relacionados as variáveis e suas classes podem ser redigidos de maneiras diferentes, porém com o mesmo significado. Os textos utilizados não seguem nenhum padrão formal de escrita, com alta variação de escrita, além de erros ortográficos e alguns textos inteiramente em letras maiúsculas. Exemplos de tais variações são as sinonímias e as variações da palavra como abreviaturas, acrônimos e estilos de escrita. Um exemplo é o termo “trombólise” que também pode ser encontrado como “terapia de reperfusão” e como acrônimo “rtPA”. Da mesma forma, o termo “hipertensão” e “diabetes” pode ser encontrado respectivamente como “HAS” e “DM”, entre outros.

O intuito de verificar o padrão de escrita resultou na criação de um manual (ANEXO B) para auxiliar os anotadores a manterem a mesma qualidade de anotação. Por vezes, quando se trata de variáveis subjetivas, o fato de ter o manual de classificação auxilia o entendimento das sentenças por parte dos anotadores. Além disso, este manual construído foi utilizado também para os algoritmos de ontologia que foram detalhados na seção 2.5.

6. Rotulação dos dados e validação do experimento

Anotação dos dois foi feita aos pares nos documentos em *Google Drive*. Para definição de cada variável e instrução as anotadoras tinham o manual de anotação para treinamento dos algoritmos (ANEXO B). A planilha de anotação continha as sentenças, e cada coluna representava uma variável. As anotadoras marcavam qual classe de cada variável aquela sentença representava e para as variáveis que não forneciam informação sobre a sentença era marcado -1. Após a anotação, os documentos foram cruzados, e as divergências eram resolvidas por terceiros a depender da especialidade (fisioterapia, neurologia, cardiologia, etc) (MESKÓ; RADÓ, 2019).

A validação cruzada é um método estatístico usado para estimar a habilidade dos

modelos de AP (JAMES; WITTEN; TIBSHIRANI, 2013; RUSSELL; NORVIG, 2016). O procedimento tem um único parâmetro chamado k que se refere ao número de grupos em que uma determinada amostra de dados deve ser dividida. Como tal, o procedimento é frequentemente denominado validação cruzada k -fold (RUSSELL; NORVIG, 2016). Quando um valor específico para k é escolhido, ele pode ser usado no lugar de k na referência ao modelo, por exemplo $k = 5$ tornando-se uma validação cruzada de 5 vezes (MEYSTRE; KIM; GOBBEL *et al.*, 2017). É comumente usado no aprendizado de máquina aplicado para comparar e selecionar um modelo para um determinado problema de modelagem preditiva porque é fácil de implementar e resulta em estimativas de habilidade que geralmente têm um viés menor do que outros métodos (BALYAN; CROSSLEY; BROWN *et al.*, 2019; CUNHA; CANUTO; VIEGAS *et al.*, 2020; GOLAS; SHIBAHARA; AGBOOLA *et al.*, 2018).

Neste estudo foi utilizado uma validação cruzada de 5 *folds*. Informações de um mesmo paciente foram atribuídas sempre ao mesmo *fold* com objetivo de testar a capacidade do modelo de prever novos dados que não foram usados na estimativa, sinalizando problemas como *overfitting* ou viés de seleção (CAWLEY; TALBOT, 2010). Apesar de conservadora, essa estratégia busca ser mais condizente com a aplicação na realidade, uma vez não se espera que a instituição tenha informações prévias de um novo paciente. Além disso, foi feita a estratificação de cada *fold*, ou seja, foi preservada porcentagem de amostras para cada classe de cada tarefa. No caso da classificação binária, isso significa que cada partição contém aproximadamente as mesmas proporções dos dois tipos de rótulos de classe (VANWINCKELEN, 2012). A implementação tem como objetivo gerar conjuntos de teste de forma que todos contenham a mesma distribuição de classes ou o mais próximo possível preserva as dependências de ordem na ordenação do conjunto de dados, gera conjuntos de teste onde o menor e o maior diferem em no máximo uma amostra (PEDREGOSA; ET AL, 2011).

7. Métricas de performance

Estudos de ferramentas de usando PLN são semelhantes a avaliação dos estudos de diagnóstico - testando se um histórico item, descoberta de exame ou resultado do teste está associado a um subsequente diagnóstico. A construção é uma tabela de contingência 2×2 onde o número de valores atribuídos correta e incorretamente para um dado resultado binário e da criação da curva ROC para mensuração da área sobre a curva. Essa tabela permite avaliação de métricas de desempenho, como precisão (Valor preditivo positivo), *recall*

(sensibilidade), precisão, *F-score*. *F-score* é a média ponderada de precisão e sensibilidade para cada classe (KOWSARI; JAFARI; HEIDARYSAFA *et al.*, 2019; KUHN; JOHNSON, 2013; VELUPILLAI; SUOMINEN; LIAKATA *et al.*, 2018). Portanto, essa pontuação leva em consideração tanto os falsos positivos quanto os falsos negativos, e é calculado pela seguinte fórmula:

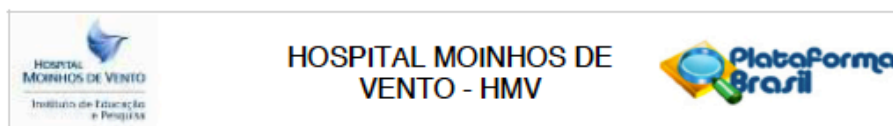
$$(1) \quad F1 \text{ Score} = 2 * (\text{Revocação} * \text{Precisão}) / (\text{Revocação} + \text{Precisão})$$

Fórmula do cálculo de F-score

Como estamos tratando de tarefas com multiníveis de classe, podemos fazer uso da Macro-F1 que retorna a média da métrica F1 para todas as classes de cada tarefa. Assim pra se ter um valor alto de Macro F1 é preciso ter um valor alto de F1 pra cada uma das classes, tanto as maiores (com muitos eventos) quanto as menores (com poucos eventos).

Embora as métricas de avaliação intrínseca sejam importantes e valorizem especialmente quando comparamos diferentes métodos de PLN para a mesma tarefa, para estabelecer um valor de referência é importante ter bem claro o que o modelo propõe responder. Não são necessariamente este valor é informativo quando a tarefa é aplicada em um problema de nível superior (por exemplo, nível do paciente) ou em novos dados. Por exemplo, se o estado da arte atual que é alcançado e a classificação de desfecho é $> 80\%$ F-score (UZUNER, 2011; VAN CALSTER; VICKERS, 2015), o que é próximo de acordo humano sobre a mesma tarefa, pode ser satisfatório para os objetivos do modelo. No entanto, se esse sistema tem por objetivo ser implantado na prática clínica e for essencial para tomada de decisão médica, qualquer taxa de erro $> 0\%$, como indicação de medicamento ou história de alergia grave, pode ser vista como inaceitável.

ANEXO A. CÓPIA DO PARECER DE APROVAÇÃO DO PROJETO NO COMITÊ DE ÉTICA



PARECER CONSUBSTANCIADO DO CEP

DADOS DA EMENDA

Título da Pesquisa: Proposição de modelo de gestão de saúde baseada em valor para os sistemas de saúde pública e suplementar do Brasil

Pesquisador: Carisi Anne Polanczyk

Área Temática:

Versão: 4

CAAE: 29894720.0.0000.5330

Instituição Proponente: Hospital Moinhos de Vento - HMV

Patrocinador Principal: Capes Coordenação Aperf Pessoal Nível Superior

DADOS DO PARECER

Número do Parecer: 4.199.370

Apresentação do Projeto:

Trata-se de uma Emenda referente a pesquisa: Proposição de modelo de gestão de saúde baseada em valor para os sistemas de saúde pública e suplementar do Brasil.

Pesquisador Responsável: Carisi Anne Polanczyk

Objetivo da Pesquisa:

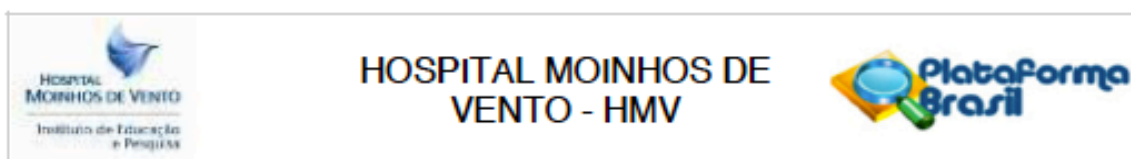
Objetivo Primário:

Propor um modelo de gestão de saúde baseada em valor para os sistemas de saúde suplementar e pública que permita orientar a gestão e alternativas de reembolso de saúde no Brasil.

Objetivos Secundários:

- I. Definir medidas de avaliação do desempenho clínico, operacional, de qualidade e segurança ao longo da Internação e de percepção do paciente que permitam avaliar os desfechos do serviço de saúde prestado.
- II. Propor cost tool kit baseadas em métodos acurados de microcusteio para aferir custos reais e compor as equações de valor.
- III. Propor framework de gestão de valor para a saúde suplementar.
- IV. Definir método de avaliação de valor para o sistema público de saúde através da proposição da equação de valor Incremental.

Endereço: Rua Tiradentes, 198 - Subsolo
Bairro: Floresta **CEP:** 90.560-030
UF: RS **Município:** PORTO ALEGRE
Telefone: (51)3314-3537 **E-mail:** cep.iep@hmv.org.br



Continuação do Parecer: 4.199.370

Avaliação dos Riscos e Benefícios:

A avaliação risco/benefício é favorável e justifica a pesquisa. Estão bem descritos e explicados no projeto de pesquisa.

Comentários e Considerações sobre a Pesquisa:

Nesta emenda, foram incluídas as seguintes informações:

- Cartas de autorização de desenvolvimento do estudo dentro dos serviços de neurologia, e documentações que foram solicitadas pelos centros co-participantes.
- Documentos na plataforma por solicitação dos CEPs dos demais centros.

Cabe destacar que nenhuma alteração foi feita no detalhamento do projeto.

Neste contexto, incluem-se os seguintes documentos:

- carta_justificativa_emendar_HMV.docx
- TCLEpara_pacientes_HospitalSaoJose.pdf
- TCLEpara_HospitalSaoJose.pdf
- anuencia_octavio_USP.pdf
- AUTORIZACAO_SAOJOSE_JOINVILLE.pdf

Considerações sobre os Termos de apresentação obrigatória:

Estão adequados.

Recomendações:

Vide conclusões.

Conclusões ou Pendências e Lista de Inadequações:

Emenda pode ser aprovadas, sem pendências.

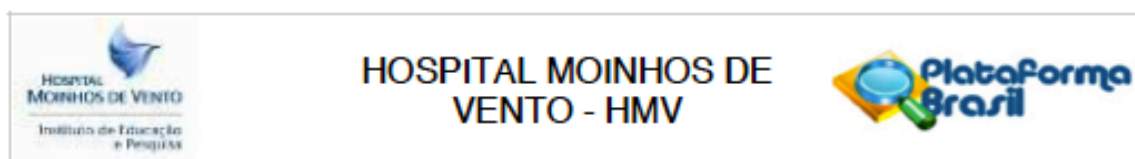
Considerações Finais a critério do CEP:

Diante do exposto, o Comitê de Ética em Pesquisa em Seres Humanos do Hospital Moínhos de Vento, de acordo com as atribuições definidas na Resolução 466/2012 do CNS e complementares, e pela Norma Operacional Nº 001/2013 do CNS, manifesta-se pela aprovação da Emenda.

Este parecer foi elaborado baseado nos documentos abaixo relacionados:

Tipo Documento	Arquivo	Postagem	Autor	Situação
----------------	---------	----------	-------	----------

Endereço: Rua Tiradentes, 198 - Subsolo
 Bairro: Floresta CEP: 90.560-030
 UF: RS Município: PORTO ALEGRE
 Telefone: (51)3314-3537 E-mail: cep.iep@hmv.org.br



Continuação do Parecer: 4.199.370

Informações Básicas do Projeto	PB_INFORMAÇÕES_BÁSICAS_1599698_E2.pdf	23/07/2020 08:17:10		Acelto
Outros	carta_justificativa_emendar_HMV.docx	23/07/2020 08:16:27	ANA PAULA BECK DA SILVA ETGES	Acelto
Outros	TCLEpara_pacientes_HospitalSaoJose.pdf	22/07/2020 09:01:56	ANA PAULA BECK DA SILVA ETGES	Acelto
Outros	TCLEpara_HospitalSaoJose.pdf	22/07/2020 09:00:15	ANA PAULA BECK DA SILVA ETGES	Acelto
Outros	anuencia_octavio_USP.pdf	22/07/2020 08:57:32	ANA PAULA BECK DA SILVA ETGES	Acelto
Outros	AUTORIZACAO_SAOJOSE_JOINVILLE.pdf	22/07/2020 08:56:36	ANA PAULA BECK DA SILVA ETGES	Acelto
Outros	EMENDA_JUSTIFICATIVA.docx	02/06/2020 09:03:00	ANA PAULA BECK DA SILVA ETGES	Acelto
Folha de Rosto	FolhaRostoAbril2020.pdf	17/04/2020 16:03:49	ANA PAULA BECK DA SILVA ETGES	Acelto
Recurso Anexado pelo Pesquisador	cartaResposta_parecr_HMV.docx	29/03/2020 09:38:49	ANA PAULA BECK DA SILVA ETGES	Acelto
TCLE / Termos de Assentimento / Justificativa de Ausência	AnexoIIconviteprofissionais.docx	29/03/2020 09:28:05	ANA PAULA BECK DA SILVA ETGES	Acelto
TCLE / Termos de Assentimento / Justificativa de Ausência	AnexoITCLEprofissionais.docx	29/03/2020 08:51:56	ANA PAULA BECK DA SILVA ETGES	Acelto
TCLE / Termos de Assentimento / Justificativa de Ausência	AnexoIIITCLEIhom.docx	29/03/2020 08:51:19	ANA PAULA BECK DA SILVA ETGES	Acelto
Projeto Detalhado / Brochura Investigador	projetoVbhchMV.docx	14/01/2020 10:09:30	ANA PAULA BECK DA SILVA ETGES	Acelto
Outros	ParecerdaComissaocientifica.pdf	14/01/2020 10:07:18	ANA PAULA BECK DA SILVA ETGES	Acelto
TCLE / Termos de Assentimento / Justificativa de Ausência	ANEXOIVdispensaTCLE.docx	14/01/2020 10:04:14	ANA PAULA BECK DA SILVA ETGES	Acelto
TCLE / Termos de Assentimento / Justificativa de Ausência	AnexoIIconvitelCHOM.docx	14/01/2020 10:03:59	ANA PAULA BECK DA SILVA ETGES	Acelto
Declaração de concordância	CarlisiPolanczyk130120pdf.pdf	14/01/2020 09:55:56	ANA PAULA BECK DA SILVA ETGES	Acelto
Declaração de Pesquisadores	termostodosassinados.pdf	14/01/2020 09:54:19	ANA PAULA BECK DA SILVA ETGES	Acelto

Endereço: Rua Tiradentes, 198 - Subsolo

Bairro: Floresta

CEP: 90.560-030

UF: RS

Município: PORTO ALEGRE

Telefone: (51)3314-3537

E-mail: cep.lep@hmv.org.br



Continuação do Parecer: 4.199.370

Situação do Parecer:

Aprovado

Necessita Apreciação da CONEP:

Não

PORTO ALEGRE, 07 de Agosto de 2020

Assinado por:

Guilherme Alcides Fiôres Soares Rollin
(Coordenador(a))

Endereço: Rua Tiradentes, 198 - Subsolo

Bairro: Floresta

CEP: 90.560-030

UF: RS

Município: PORTO ALEGRE

Telefone: (51)3314-3537

E-mail: cep.lep@hmv.org.br

ANEXO B. MANUAL DE TREINAMENTO PARA OS ALGORITMOS

A construção de algoritmos de classificação de textos baseados em processamento de linguagem natural requer que os dados de texto sejam anotados por especialistas que devem buscar um padrão de anotação. As anotações foram feitas por duas especialistas da área da saúde e conhecedoras do método aplicado, levando em consideração o manual de definição para cada variável. Este manual foi produzido e validado pela equipe do projeto, nas suas diferentes expertises.

Universidade Federal do Rio Grande do Sul
Instituto de Avaliação de Tecnologias em Saúde

Proposta de automatização da análise de desfecho para dimensionar programas de
gestão de saúde baseada em valor.

Manual de Orientações para Classificação de texto das Evoluções

Coordenação do projeto:

Profa Dra. Carisi Anne Polanczyk

Confecção do manual: Bruna Stella Zanotto, Renata Garcia Ruschel

Porto Alegre, 2020.

LISTA DE ABREVIATURAS

AVC	Acidente vascular cerebral
AVCi	Acidente vascular cerebral isquêmico
AVC ISQ	Acidente vascular cerebral isquêmico
AVCh	Acidente vascular cerebral hemorrágico
ATB	Antibiótico
CI	Cardiopatía isquêmica
IAM	Infarto agudo do miocárdio
DAC	Doença arterial coronariana
DCC	Doença cardíaca coronariana
SCA	Síndrome coronariana aguda
IC	Insuficiência cardíaca
CRM	Cirurgia de revascularização do miocárdio
MP	Marca passo
FA	Fibrilação atrial
ACFA	Fibrilação atrial
CA	Câncer
DM	Diabetes mellitus
DM2	Diabetes mellitus tipo II
HAS	Hipertensão
DLP	Dislipidemia
TC	Tomografia computadorizada
TCC*	Tomografia computadorizada cardíaca / controle
RM	Ressonância magnética
SV	Sinais vitais
AA	Ar ambiente
AO	Abertura ocular
JH	John Hopkins (escala de risco de queda)
GF	Grau de força (escala)
FG	Força grau (escala)
VO	Via oral
GTT	Gastrostomia

SNE	Sonda nasointestinal
BI	Bomba de infusão
ITU	Infecção do trato urinário
HMC	Hemocultura
EVA	Escala visual analógica
PAINAD	Pain Assessment in Advanced Dementia
BPS	Behavioral Pain Scale
MI	Membro inferior
MsIs	Membros inferiores
MIE	Membro inferior esquerdo
MID	Membro inferior direito
MS	Membro superior
MsSs	Membros superiores
MSE	Membro superior esquerdo
MSD	Membro superior direito
LOC	Lucido, orientado e coerente
BEG	Bom estado geral
REG	Regular estado geral
MEG	Mal estado geral
Tmax	Temperatura máxima
NIHSS	National Institute of Health Stroke Scale
CTI	Centro de tratamento intensivo
UTI	Unidade de tratamento intensivo
EMG	Emergência
UI	Unidade de internação

1. Apresentação

A construção de algoritmos de classificação de textos baseados em processamento de linguagem natural requer que os dados de texto sejam anotados por especialistas que devem buscar um padrão de anotação.

Este manual foi criado para orientar o(a) anotador (a) no contexto do projeto de pesquisa intitulado “Modelo de Gestão de Saúde Baseado em Valor para o Sistema Único de Saúde e Saúde Suplementar do Brasil”, aprovado pelo Comitê de Ética em Pesquisa (CAAE 29694720000005330).

2. População-Alvo

Esse material é destinado a pesquisadores das instituições hospitalares que venham a replicar em suas instituições as soluções computacionais desenvolvidas no projeto de pesquisa “Modelo de Gestão de Saúde Baseado em Valor para o Sistema Único de Saúde e Saúde Suplementar do Brasil”.

3. Orientações gerais

Ao anotar dados de sua instituição, algumas premissas são importantes:

- Solicite os códigos fonte aos pesquisadores escrevendo um e-mail para anabsetges@gmail.com.

- Ao receber os códigos, não altere a estrutura ou nome das variáveis, caso seja necessário, entre em contato com os pesquisadores do projeto nos e-mails: anabsetges@gmail.com ou brunaszanotto@gmail.com.

- A informação negativa de uma sentença sempre deve ser anotada, por exemplo, “nega hipertensão”, precisa ser classificada como uma sentença que refere a ausência de hipertensão. Diferentemente do caso de uma sentença não ter relação nenhuma com a variável avaliada, quando isso acontecer, deixe-a em branco

- Ao longo do processo de anotação, o objetivo é extrair dados de evoluções de texto com base na melhor evidência descrita. Por isso, desconsidera-se o dado que informa apenas resultados de exames ou hipóteses diagnósticas por mais que esses sejam passíveis de fazer

associações fortes com morbidades do nosso interesse. Sob hipótese alguma os modelos irão indicar ou predizer algum diagnóstico, apenas buscam estruturar informações para que os profissionais de saúde possam agir e diagnosticar.

- Considerar erro gramatical desde que fique claro a informação a ser anotada, por exemplo, “diabete melitos”, “isnfuciencia respiratoria”

- Certifique-se de que a base de dados texto que será anotada está anonimizada. Caso nome de pacientes e profissionais sejam muito frequentes nas evoluções, entre em contato com nossa equipe para juntos buscarmos uma forma de ocultá-los.

4. Apresentação do manual para as Tasks (Variáveis)

O formulário é composto por 30 variáveis, divididas em quatro grupos (características clínicas; manejo clínico e processo de cuidado; escalas de avaliação e eventos de risco; indicadores clínicos, desfechos e status do paciente).

Para cada variável é apresentada a definição acrescida de exemplos de sentenças que remetem a classificação do texto em cada variável.

4.1. Características Clínicas

Variável	Definição	Opções de Resposta	Orientações para resposta / Sentenças
Doença coronariana	CID10 I20 – Angina Pectoris	0- Doença ausente	# Sem histórico de CI
	CID10 I21 – Infarto Agudo do Miocárdio		# Não possui história progressiva de IAM
	CID10 I22 – Infarto do Miocárdio Recorrente		# Paciente sem histórico de DAC
			#Miocardiopatia hipertrófica
			# Angina estável # Angina instável
	CID10 I24 – Outras Doenças Isquêmicas Agudas do Coração	1- Doença presente	# ACTP
	CID10 I25 – Doença Isquêmica Crônica do Coração		# IAM # IAMSSST # IAMCSST # CRM # Cardiopatia isquêmica # CI # DAC # SCA # DCC #Aterosclerose coronariana/coronária #Estenose coronariana/coronária

Variável	Definição	Opções de Resposta	Orientações para resposta / Sentenças
Fibrilação Atrial	CID10 I48 - Flutter e fibrilação atrial	0- Doença ausente	# Sem história progressa de FA # Não apresenta fibrilação atrial
		1- Doença presente	#FA #Fibrilação atrial # Fibrilação atrial paroxística # ACFA
Diabetes	CID10-E10 – Diabetes Mellitus Insulino-dependente	0- Doença ausente	# Nega DM # Não possui diabetes # Sem história progressa de DM
	CID10-E11 – Diabetes Mellitus Não-insulino-dependente		
	CID10-E12 – Diabetes Mellitus Relacionado Com a Desnutrição	1- Doença presente	# DM # Diabetes mellitus
	CID10-E13 – Outros Tipos Especificados de Diabetes Mellitus		# Diabete mellitus # Diabetes
CID10-E14 – Diabetes Mellitus Não Especificado		# DM2	
AVC prévio	CID10-I64 Acidente vascular cerebral, não especificado como hemorrágico ou isquêmico	0- Não	# Sem história prévia de AVC
	CID10-I63 – Infarto Cerebral	1- Sim	# AVC isquêmico prévio # AVC ISQ em 2018 # História prévia: AVCi AVCh # Acidente vascular cerebral # Acidente vascular encefálico
Hipertensão arterial sistêmica	CID10-I10 – Hipertensão Essencial (primária)	0- Doença ausente	# Nega HAS. # Não possui história progressa de hipertensão # Sem histórico de HAS

CID10- I11 – Doença Cardíaca Hipertensiva	1- Doença presente	# HAS # Hipertensão arterial Hipertensa # História prévia de Has.
---	-----------------------	--

Variável	Definição	Opções de Resposta	Orientações para resposta / Sentenças
Obesidade	CID10- E66 – Obesidade	0- Doença ausente	-
		1- Doença presente	# Obesidade # Obeso # IMC > 30 Índice de massa corporal > 30
Dislipidemia	CID10-E78 - Distúrbios do metabolismo de lipoproteínas e outras lipidemias	0- Doença ausente	# Nega DLP # Sem história progressiva de dislipidemia
		1- Doença presente	# Dislipidemia # Dislipidêmico # DLP
Câncer	Registro de câncer em atividade, indicativo de realização de tratamentos para a doença (radioterapia, quimioterapia) ou indicativo de câncer recente. Quando indicar que o paciente já teve câncer curado ou no passado não deverá ser considerado.	0- Doença ausente	#2005 câncer de pele #Ca no passado
		1- Doença presente	# Ca em atividade # Câncer em tratamento #em tto quimioterápico # Neoplasia de próstata ativa
Tabagismo	Registro do consumo condicionado a dependência de cigarros ou outros produtos que contenham tabaco (possível relato do paciente).	0- Não tabagista	# Nega tabagismo # Não tabagista
		1- Tabagista	# Tabagismo # Tabagista # Fumante
		2- Ex-tabagista	# Ex-tabagista # Tabagista no passado

Variável	Definição	Opções de Resposta	Orientações para resposta / Sentenças
Etilista	Registro do consumo condicionado a dependência de álcool (possível relato do paciente). O relato de negação ou ex-alcoolista também deve ser considerado	0- Não etilista	# Nega etilismo # Não etilista
		1- Etilista	# Etilismo # Alcoolismo # Abuso de álcool # Alcoolista
		2- Ex etilista	# História prévia de alcoolismo # Ex alcoólatra #Ex elitista

4.2. Manejo Clínico e Processo de cuidado

Variável	Definição	Opções de Resposta	Orientações para resposta / Sentenças
Localização	Indica o local de presente momento do paciente. ***Sentenças que indicam deslocamento do paciente sem informativo do destino são desconsideradas	1- Emergência	# Evolução diária- emergência- enfermagem
		3- CTI	## CTIA ## # Paciente proveniente da emergência por volta das 15:50h pós trombólise por AVCi. # Interna na UTI após procedimento. # Fisioterapia terapia intensiva
		4-Unidade Internação	# Enfermagem C1. # Paciente deu entrada na UI
Trombectomia	Refere ao possível manejo clínico que o paciente AVCi recebeu.	0- Não realizou ou sem janela para terapia	# Sem indicação de trombólise ou trombectomia. # Sem janela para trombectomia
	*A contraindicação, sem delta ou sem janela também deve ser considerada como não realização do procedimento	1- Trombectomia	# Trombectomia mecânica – # Trombectomia mecânica as 12 horas
Trombólise	Refere ao possível manejo clínico que o paciente AVCi recebeu.	0- Não realizou ou sem janela para terapia	# Não trombolizada - delta # Contraindicação a trombólise (hipodensidade >1/3). # Sem janela trombolítica
	*A contraindicação, sem delta ou sem janela também deve ser considerada como não realização do procedimento.	1- Trombólise	# Trombolise endovenosa # ALTEPLASE # AVCi trombolizado

4.3. Escalas de Avaliação e Eventos de Risco

Variável	Definição	Opções de Resposta	Orientações para resposta/ Sentenças
Hemorragia intracraniana	CID10 - I61 Hemorragia intracerebral.	0- Hemorragia ausente	# TC controle ontem sem sangramento # TCC sem alterações # Sem indícios de sangramento
	CID10 - I62 Outras hemorragias intracranianas não-traumáticas.	1- Hemorragia presente	# Hemorragia cerebelar # Hemorragia cerebral com transformação hemorrágica # Hemorragia em fossa posterior próximo ao IV ventrículo.
Queda	Registro de queda do paciente durante período de internação.	1- Sim, queda	# Paciente relata insegurança de usar andador após queda # Paciente perdeu o equilíbrio e apresentou queda da própria altura. # Caiu no quarto a noite
Escala Braden	Escala de Braden é um recurso utilizado nas Unidades de Terapia Intensiva para medir o risco dos pacientes críticos de desenvolverem lesões por pressão.	1- Risco baixo	Baixo risco: escore =>17
		2-Risco moderado	Risco moderado: escore 16-13
		3-Risco alto	Risco alto: escore =<12
Risco de queda	Variável com objetivo de identificar o risco de queda do paciente de acordo com a escala John Hopkins. Atenta-se também a descritivos da escala que podem ser ajustados de acordo com o risco definido por consenso dos profissionais.	1- Baixo risco	Baixo risco: escore de 0-5 pontos ou atentar para medidas preventivas de queda. Ex.: # Risco para Quedas (Escala John Hopkins 02); # Manter medidas de prevenção de queda;
		2- Moderado risco	Risco moderado: escore de 6-13 pontos. Ex.: #Reforço orientações com a familiar sobre o risco de quedas. # risco de queda. *Sentenças que apresentam apenas "risco de queda", sem score ou contexto, podem ser consideradas como risco moderado.

	3-Alto risco	Alto risco: escore > 13 pontos. Implementação do protocolo de risco de queda. Ex.: i) Alto risco de queda (14). ii) Protocolo de alto risco para queda;
--	--------------	--

Variável	Definição	Opções de Resposta	Orientações para resposta/ Sentenças
Indicativo de Infecção	<p>Sentenças que indiquem possível infecção viral, bacteriana ou fúngica no paciente.</p> <p>Observar: sinais flogísticos (Calor (febre), rubor ou hiperemia (vermelhidão), edema (inchaço), hiperestesia (dor ao toque), perda de função), uso de antibióticos, exames de investigação como hemocultura, início de secreção por infecção</p> <p>**desconsiderar resultados de exames sem indicativo explícito de infecção *** em anexo alguns antimicrobianos comuns na prática clínica</p>	0- Indicativo negativo de infecção	<p>#Sem sinais flogísticos</p> <p>#Afebril, sinais vitais estáveis</p> <p>#Temp máx 37,2°C</p> <p># Culturais HMCs 09/08: negativas ATQ 09/08: negativo HMCs 12/08: negativas ATQ</p>
		1- Indicativo alerta de infecção	<p># Pneumonia base direita-< BACTRIN</p> <p># EQU: urocultura em andamento.</p> <p>#NOVA ITU - MEROPENEN -> E.COLI</p> <p>#sepse</p>

4.4. Indicadores Clínicos Desfechos e Status do Paciente

Variável	Definição	Opções de Resposta	Orientações para resposta/ Sentenças
Óbito	Sentenças que indiquem nota de óbito/falecimento. Indicativos de não-óbito podem ser marcados e são considerados: sinais vitais estáveis, registros de paciente presente no leito, bem, respirando espontaneamente, previsão de alta do paciente.	0- Indicativos de presença e sinais vitais	# No leito, tranquilo, acompanhado. # Condição Ventilatória: ar ambiente, eupneica # Eupneica, ventilando espontâneo em ar ambiente, sem esforço ventilatório; # Hemodinamicamente estável # SV estáveis
		1- Nota de óbito	# Nota de falecimento # Paciente faleceu hoje pela manhã
Dor	Indica se o paciente apresenta sinais clínicos de dor. **observar: escala EVA de dor, escala PAINAD, escala BPS e composições de palavras que justifiquem marcação.	0- Sem dor	# Paciente sem queixas # Paciente bem, sem dor #Paciente não refere dor # Sem queixas algicas # Nega dor.
		1- Dor leve/moderada	Escores (EVA/PAINAD/BPS) 1 a 6 Ex.: #Dor indefinida #Dor leve #Queixa de dor.
		2- Dor intensa	Escores (EVA/PAINAD/BPS) 7 a 10 Ex.: #Dor difusa #Dor grave #Dor intensa
Alimentação	Indica se o paciente precisa de sonda ou gastrostomia para se alimentar.	1- Alimentação por via oral	#Paciente alimentando-se por via oral. #Dieta para semi pastosa #Boa aceitação VO.
		2- Uso de sondas ou gastrostomia	#Paciente alimentando-se por via alternativa # Cuidados com SNE;

		#Dieta em BI; #Cuidados com gastrostomia #GTT #Dieta polimérica
--	--	--

Variável	Definição	Opções de Resposta	Orientações para resposta/ Sentenças
Força	<p>Indica a capacidade de superar ou opor-se a uma resistência por meio da atividade muscular, conforme escala de grau de força, composições descritivas ou nível de mobilidade associado.</p> <p>Levar em conta apenas membros superiores e inferiores (facial não), no caso de uma sentença apresentar vários graus de força, priorizar força de Membros Inferiores e proximais e menor grau apresentado.</p> <p>*Verificar relação com escala NIH</p>	Escore de 0 a 5	<p>Apresentação da escala:</p> <p>0 – Não percebe contração. Plegia/paciente plégico é marcado com força 0.</p> <p>1 – Traço de contração sem produção de movimento. *Passível de correlação com escala NIH domínio Membros Superiores (5) e Membros Inferiores (6) pontuação 4.</p> <p>2 – Contração fraca, elimina gravidade. *Passível de correlação com escala NIH domínio Membros Superiores (5) e Membros Inferiores (6) pontuação 3.</p> <p>3 – Realiza movimento contra a gravidade, porém sem resistência adicional. *Passível de correlação com escala NIH domínio Membros Superiores (5) e Membros Inferiores (6) pontuação 2 ou 1.</p> <p>4 – Realiza movimento contra a gravidade e resistência externa. *Passível de correlação com escala NIH domínio Membros Superiores (5) e Membros Inferiores (6) pontuação 0.</p> <p>5 - Supera maior quantidade de resistência. Sem alteração em força.</p>
			<p>Movimento limitado ou fraco, motilidade num padrão abaixo do normal. No que se refere à força muscular, precisão do movimento, amplitude do movimento e a resistência muscular localizada, ou seja, refere-se a um comprometimento parcial, uma perda de força.</p> <p>*Levar em conta apenas membros superiores e inferiores (paresia facial não)</p>

Variável	Definição	Opções de Resposta	Orientações para resposta/ Sentenças
Mobilidade	<p>Variável indica se paciente é capaz de andar com autonomia.</p> <p>Associada com escala de nível de mobilidade utilizada pela instituição (Callen, BL., et al. <i>Medsurg nursing, 2004</i>) ou a descrição dos seus níveis, podendo ser adaptada conforme necessário.</p> <p>***Usualmente associada com grau de força e escala de nível de mobilidade da instituição (se aplicável) conforme orientações de resposta.</p>	0 - Incapaz de andar	<p>Incapacidade de andar podem ser relacionados aos níveis de 1 a 8 na escala de mobilidade, ou descrições como:</p> <p>i) paciente dependente: equipe promove trocas de decúbito, posicionamento e ADM;</p> <p>ii) paciente participa com a equipe das trocas de decúbito, posicionamento e ADM;</p> <p>iii) paciente é independente no leito;</p> <p>iv) transferido para a cadeira; v) auxílio mecânico ou 3 pessoas para a cadeira /cadeira de rodas;</p> <p>vi) transferência para a cadeira/de rodas com auxílio de duas pessoas;</p> <p>vii) auxílio de duas pessoas, ortostase e pivô para a cadeira;</p> <p>viii) auxílio de uma pessoa, ortostase e pivô para a cadeira;</p>
		1 - Capaz de andar sem ajuda de outra pessoa ou dispositivo	<p>Capaz de andar sem ajuda de outra pessoa ou dispositivo pode ser relacionados aos níveis de mobilidade de 12 a 15 na escala de mobilidade, ou descrições como:</p> <p>i) deambula com um auxiliar de prontidão;</p> <p>ii) deambula independente somente no quarto;</p> <p>iii) deambula fora do quarto, distância menor que um corredor;</p> <p>iv) deambula fora do quarto, distância maior que um corredor.</p> <p>*** Quando mobilidade 1, marcar força 5.</p>
		2 - Capaz de andar com ajuda de outra pessoa ou dispositivo	<p>Capaz de andar com a ajuda de outra pessoa ou dispositivo podem ser relacionados aos níveis de 9 a 11 na escala de mobilidade, ou descrições como:</p> <p>i) uma pessoa ao lado de prontidão para transferir para a cadeira;</p> <p>ii) deambula com dois auxiliares;</p> <p>iii) deambula com um auxiliar.</p> <p>*** Quando mobilidade 2, marcar força 4.</p>

Variável	Definição	Opções de Resposta	Orientações para resposta/ Sentenças
Nível de mobilidade	<p>Facultativo: variável dependente do uso, pela instituição, de escalas de rastreamento para níveis de mobilidade do paciente.</p> <p>Caso não haja uma escala padrão proposta, podem ser feitas de acordo com as descrições indicadas.</p> <p>Escala de nível de mobilidade utilizada pela instituição (<i>Callen, BL., et al. Medsurg nursing, 2004</i>)</p>	Escore de 1 a 15	<p>1 - Paciente dependente</p> <p>2- Paciente participa com a equipe das trocas de decúbito, posicionamento e ADM</p> <p>3- Paciente é independente no leito</p> <p>4- Transferido para a cadeira</p> <p>5-auxílio mecânico ou 3 pessoas para a cadeira /cadeira de rodas</p> <p>6- Transferência para a cadeira/de rodas com auxílio de duas pessoas</p> <p>7-Auxílio de duas pessoas, ortostase e pivô para a cadeira</p> <p>8-Auxílio de uma pessoa, ortostase e pivô para a cadeira</p> <p>9-Uma pessoa ao lado de prontidão para transferir para a cadeira</p> <p>10-Deambula com dois auxiliares</p> <p>11-Deambula com um auxiliar</p> <p>12-Deambula com um auxiliar de prontidão</p> <p>13-Deambula independente somente no quarto</p> <p>14-Deambula fora do quarto, distância menor que um corredor</p> <p>15-Deambula fora do quarto, distância maior que um corredor</p>

Variável	Definição	Opções de Resposta	Orientações para resposta / Sentenças
Comunicação	<p>Variável indica se paciente tem problemas de comunicação verbal.</p> <p>*Verificar relação com escala NIH.</p>	0- Sem comunicação verbal	<p>Impossibilidade de efetuar comunicação verbal.</p> <p>Afasia grave, toda a comunicação é feita através de expressões fragmentadas; necessidade de interferência, a quantidade de informação que pode ser trocada é limitada.</p> <p>Ex.:</p> <p># Paciente comunica por gestos.</p> <p>*Passível de correlação com escala NIH domínio Linguagem (9) e Disartria (10) pontuação 2 ou 3.</p>
		1- Comunica bem	<p>Comunicação sem dificuldades/assintomáticas.</p> <p>Ex.:</p> <p># Paciente dá bom dia;</p> <p># Fala ok;</p> <p># Recuperado dos sintomas de fala;</p> <p># Sem afasia, normal;</p> <p># Comunicativo</p> <p>*Passível de correlação com escala NIH domínio Linguagem (9) e Disartria (10) pontuação 0.</p>
		2- Comunica pouco ou mal	<p>Comunicação verbal com dificuldade/problemas.</p> <p>Afasia leve a moderada; perda óbvia de alguma fluência, sem limitação significativa das ideias expressas ou formas de expressão.</p> <p>Ex.:</p> <p># Paciente com afasia de expressão e agora disartria.</p> <p># Fala enrolada</p> <p>*Passível de correlação com escala NIH domínio Linguagem (9) e Disartria (10) pontuação 1.</p>

Variável	Definição	Opções de Resposta	Orientações para resposta/ Sentenças
Capacidade cognitiva	Variável sobre o entendimento, situação de lucidez e orientação do paciente.	0- Não estado de clareza	<p>Paciente sem clareza ou em estado de sedação.</p> <p>Ex.:</p> <p># Paciente sedado, não responsivo,</p> <p># Paciente no leito, abertura ocular espontânea, não interage, sem atender comandos.</p> <p>*Passível de correlação com escala NIH domínio Nível de Consciência (1,1a,1b) pontuação 2.</p>
	Deve-se sempre avaliar o contexto da sentença, pois palavras como "colaborativo", "bem", "interagindo", "atendem comandos" quando sozinhas não indicam entendimento pleno, porém quando agrupadas em contexto podem traduzir algum discernimento do paciente.	1- Lúcido, orientado, coerente	<p>Situação de lucidez e orientação do paciente.</p> <p>Ex.:</p> <p># LOC</p> <p># Escala glasgow 15</p> <p># Paciente lúcido, responsivo, tranquilo, obedece a comandos</p> <p>*Passível de correlação com escala NIH domínio Nível de Consciência (1,1a,1b) pontuação 0.</p>
	*Verificar relação com escala NIH.	2- Confuso, compreende mal	<p>Paciente confuso, compreendendo mal, delirium.</p> <p>Ex.:</p> <p># Paciente relata confusão em pensamentos, desorientado em espaço e tempo.</p> <p>*Passível de correlação com escala NIH domínio Nível de Consciência (1,1a,1b) pontuação 1.</p>

Variável	Definição	Opções de Resposta	Orientações para resposta / Sentenças
Escala Rankin	<p>Escala de avaliação funcional pós-AVC. Preenche a variável conforme indicação do score na evolução do paciente (de 0 a 6).</p> <p>*** Quando rankin 4 ou 5, marcar auto-cuidado = 0</p>	Escore de 0 a 6	<p>Apresentação da escala:</p> <p>0 - Sem sintomas</p> <p>1 - Nenhuma deficiência significativa, a despeito de sintomas</p> <p>2 - Leve deficiência - Incapaz conduzir todas as atividades de antes, mas é capaz de cuidar dos próprios interesses sem assistência</p> <p>3 - Deficiência moderada -Requer alguma ajuda, mas é capaz de caminhar sem assistência (pode usar bengala ou andador)</p> <p>*4 - Deficiência moderadamente grave - Incapaz de caminhar sem assistência e incapaz de atender às próprias necessidades fisiológicas sem assistência</p> <p>*5 - Deficiência grave - Confinado à cama, incontinente, requerendo cuidados e atenção constante de enfermagem</p>
			6 – Óbito
Auto cuidado	<p>Capacidade de autonomia em atividades básicas como vestir-se, ir sozinho ao banheiro.</p> <p>*A não autonomia também é indicada, inclusive pela relação com rankin 4 e 5.</p>	0- Dependente para atividades básicas/de higiene	<p># Eliminações espontâneas, em fraldas</p> <p># Rankin 4</p> <p># Restrito ao leito no momento</p> <p># Precisa de auxílio para as atividades da vida diária.</p>
		1- Independente para atividades básicas/de higiene	<p># Tem vida de relação, conversa, caminha, alimenta-se.</p> <p># Paciente faz uso banheiro/WC</p>
NIH	<p>A National Institute of Health Stroke Scale (NIHSS) é uma escala padrão, validada, segura, quantitativa da severidade e magnitude do déficit neurológico após o AVC.</p> <p>Seu escore varia de 1 a 42, porém não é incomum esta escala</p>	Escore de 1 a 42	<p># NIH 14 (sonolento 1, errou 1 pergunta, paresia face 2, desvio do olhar 1, paresia MSE 3, MIE3, disartria 1, Heminegligência 2)</p> <p># NIH: face: 1, 2, negligência 2, MSE 3, MIE 3 - Total: 11.</p> <p># NIHSS 5</p>

aparecer de forma a indicar o nível de cada domínio considerado, quando isso acontece os valores podem ser traduzidas a outras variáveis que temos presente no manual, conforme instruções acima.

*** em anexo a escala NIH

A variável “status_chegada” é uma variável-suporte nos casos em que a sentença diz respeito a uma informação claramente no passado do paciente (por exemplo, situação do paciente na chegada ao hospital), e não refere seu quadro atual; isso não desonera o fato da marcação da variável correspondente.

4.4.1. ANEXO A1 – Informação de apoio para antimicrobianos comuns na prática clínica

Antimicrobianos comuns na prática clínica	Exemplos
Penicilinas	Amoxicilina, Ampicilina, Azlocilina, Carbenicilina, Cloxacilina
Cefalosporinas	Cefaclor, Cefadroxilo, Cefazolina, Cefixime., Cefoperazona, Cefotaxima, Cefotetan, Cefoxitina, Ceftazidima, Ceftriaxona, Cefuroxima, Cefalexina, Cefalotina, Loracarbef.
Quinolonas	Ciprofloxacina, Enoxacina, Levofloxacina, Lomefloxacina, Moxifloxacina, Norfloxacina, Ofloxacina.
Aminoglicosídeos	Amicacina, Gentamicina, Canamicina, Neomicina, Estreptomicina, Tobramicina.
Macrolídios	Azitromicina, Claritromicina, Eritromicina.
Tetraciclínas	Doxiciclina, Minociclina, Tetraciclínas.
Outros exemplos	Aztreonam, Clindamicina, Etambutol, Fosfomicina, Imipenem, Isoniazida, Meropenem, Metronidazol, Nitrofurantoina, Pirazinamida, Rifampicina, Trimetoprima, Fluconazol

4.4.2. ANEXO A2 – Escala NIHSS e definição de cada domínio

Instrução	Definição da escala
<p>1a. Nível de Consciência</p> <p>O investigador deve escolher uma resposta mesmo se uma avaliação completa é prejudicada por obstáculos como um tubo orotraqueal, barreiras de linguagem, trauma ou curativo orotraqueal. Um 3 é dado apenas se o paciente não faz nenhum movimento (outro além de postura reflexa) em resposta à estimulação dolorosa.</p>	<p>0 = Alerta; reponde com entusiasmo.</p> <p>1= Não alerta, mas ao ser acordado por mínima estimulação obedece, responde ou reage.</p> <p>2= Não alerta, requer repetida estimulação ou estimulação dolorosa para realizar movimentos (não estereotipados).</p> <p>3= Responde somente com reflexo motor ou reações autonômicas, ou totalmente irresponsivo, flácido e irreflexo.</p>
<p>1b. Perguntas de Nível de Consciência</p> <p>O paciente é questionado sobre o mês e sua idade. A resposta deve ser correta – não há nota parcial por chegar perto. Pacientes com afasia ou esturpor que não compreendem as perguntas irão receber 2. Pacientes incapacitados de falar devido a intubação orotraqueal, trauma orotraqueal, disartria grave de qualquer causa, barreiras de linguagem ou qualquer outro problema não secundário a afasia receberá um 1. É importante que somente a resposta inicial seja considerada e que o examinador não “ajude” o paciente com dicas verbais ou não verbais.</p>	<p>0= Responde ambas as questões corretamente.</p> <p>1= Responde uma questão corretamente.</p> <p>2= Não responde nenhuma questão corretamente.</p>
<p>1c. Comandos de Nível de Consciência</p> <p>O paciente é solicitado a abrir e fechar os olhos e então abrir e fechar a mão não parética. Substitua por outro comando de um único passo se as mãos não podem ser utilizadas. É dado crédito se uma tentativa inequívoca é feita, mas não completada devido à fraqueza. Se o paciente não responde ao comando, a tarefa deve ser demonstrada a ele (pantomima) e o resultado registrado (i.e., segue um, nenhum ou ambos os comandos). Aos pacientes com trauma, amputação ou outro impedimento físico devem ser dados comandos únicos compatíveis. Somente a primeira tentativa é registrada.</p>	<p>0= Realiza ambas as tarefas corretamente.</p> <p>1= Realiza uma tarefa corretamente.</p> <p>2= Não realiza nenhuma tarefa corretamente.</p>

<p>2. Melhor olhar conjugado</p> <p>Somente os movimentos oculares horizontais são testados. Movimentos oculares voluntários ou reflexos (óculo-cefálico) recebem nota, mas a prova calórica não é usada.</p> <p>Se o paciente tem um desvio conjugado do olhar, que pode ser sobreposto por atividade voluntária ou reflexa, o escore será 1. Se o paciente tem uma paresia de nervo periférica isolada (NC III, IV ou VI), marque 1. O olhar é testado em todos os pacientes afásicos. Os pacientes com trauma ocular, curativos, cegueira preexistente ou outro distúrbio de acuidade ou campo visual devem ser testados com movimentos reflexos e a escolha feita pelo investigador. Estabelecer contato visual e, então, mover-se perto do paciente de um lado para outro, pode esclarecer a presença de paralisia do olhar.</p>	<p>0= Normal.</p> <p>1= Paralisia parcial do olhar. Este escore é dado quando o olhar é anormal em um ou ambos os olhos, mas não há desvio forçado ou paresia total do olhar.</p> <p>2= Desvio forçado ou paralisia total do olhar que não podem ser vencidos pela manobra óculo-cefálica.</p>
<p>3. Visual</p> <p>OS campos visuais (quadrantes superiores e inferiores) são testados por confrontação, utilizando contagem de dedos ou ameaça visual, conforme apropriado. O paciente deve ser encorajado, mas se olha para o lado do movimento dos dedos, deve ser considerado como normal. Se houver cegueira unilateral ou enucleação, os campos visuais no olho restante são avaliados. Marque 1 somente se uma clara assimetria, incluindo quadrantanopsia, for encontrada. Se o paciente é cego por qualquer causa, marque 3. Estimulação dupla simultânea é realizada neste momento. Se houver uma extinção, o paciente recebe 1 e os resultados são usados para responder a questão 11.</p>	<p>0= Sem perda visual.</p> <p>1= Hemianopsia parcial.</p> <p>2= Hemianopsia completa.</p> <p>3= Hemianopsia bilateral (cego, incluindo cegueira cortical).</p>
<p>4. Paralisia Facial</p> <p>Pergunte ou use pantomima para encorajar o paciente a mostrar os dentes ou sorrir e fechar os olhos. Considere a simetria de contração facial em resposta a estímulo doloroso em paciente pouco responsivo ou incapaz de compreender. Na presença de trauma /curativo facial, tubo orotraqueal, esparadrapo ou outra barreira física que obscureça a face, estes devem ser removidos, tanto quanto possível.</p>	<p>0= Movimentos normais simétricos.</p> <p>1= Paralisia facial leve (apagamento de prega nasolabial, assimetria no sorriso).</p> <p>2= Paralisia facial central evidente (paralisia facial total ou quase total da região inferior da face).</p> <p>3= Paralisia facial completa (ausência de movimentos faciais das regiões superior e inferior da face).</p>
<p>5. Motor para braços</p> <p>O braço é colocado na posição apropriada: extensão dos braços (palmas para baixo) a 90° (se sentado) ou a 45° (se deitado). É valorizada queda do braço se esta ocorre antes de 10 segundos. O paciente afásico é encorajado através de firmeza na voz e de pantomima, mas não com estimulação dolorosa. Cada membro é testado isoladamente, iniciando pelo braço não-parético. Somente em</p>	<p>0= Sem queda; mantém o braço 90° (ou 45°) por 10 segundos completos.</p> <p>1= Queda; mantém o braço a 90° (ou 45°), porém este apresenta queda antes dos 10 segundos completos; não toca a cama ou outro suporte.</p> <p>2= Algum esforço contra a gravidade; o braço não atinge ou não mantém 90° (ou 45°), cai na cama, mas tem alguma força contra a gravidade.</p> <p>3= Nenhum esforço contra a gravidade; braço</p>

<p>caso de amputação ou de fusão de articulação no ombro, o item deve ser considerado não-testável (NT), e uma explicação deve ser escrita para esta escolha.</p>	<p>despenca. 4= Nenhum movimento. NT = Amputação ou fusão articular, explique: _____</p> <p>5a. Braço esquerdo 5b. Braço direito</p>
<p>6. Motor para pernas A perna é colocada na posição apropriada: extensão a 30° (sempre na posição supina). É valorizada queda do braço se esta ocorre antes de 5 segundos. O paciente afásico é encorajado através de firmeza na voz e de pantomima, mas não com estimulação dolorosa. Cada membro é testado isoladamente, iniciando pela perna não-parética. Somente em caso de amputação ou de fusão de articulação no quadril, o item deve ser considerado não-testável (NT), e uma explicação deve ser escrita para esta escolha.</p>	<p>0= Sem queda; mantém a perna a 30° por 5 segundos completos. 1= Queda; mantém a perna a 30°, porém esta apresenta queda antes dos 5 segundos completos; não toca a cama ou outro suporte. 2= Algum esforço contra a gravidade; a perna não atinge ou não mantém 30°, cai na cama, mas tem alguma força contra a gravidade. 3= Nenhum esforço contra a gravidade; perna despenca. 4= Nenhum movimento. NT = Amputação ou fusão articular, explique: _____</p> <p>6a. Perna esquerda 6b. Perna direita</p>
<p>7. Ataxia de membros Este item é avaliado se existe evidência de uma lesão cerebelar unilateral. Teste com os olhos abertos. Em caso de defeito visual, assegure-se que o teste é feito no campo visual intacto. Os testes índex-nariz e calcanhar-joelho são realizados em ambos os lados e a ataxia é valorizada, somente, se for desproporcional à fraqueza. A ataxia é considerada ausente no paciente que não pode entender ou está hemiplégico. Somente em caso de amputação ou de fusão de articulações, o item deve ser considerado não-testável (NT), e uma explicação deve ser escrita para esta escolha. Em caso de cegueira, teste tocando o nariz, a partir de uma posição com os braços estendidos.</p>	<p>0= Ausente. 1= Presente em 1 membro. 2= Presente em dois membros. NT = Amputação ou fusão articular, explique: _____</p>

<p>8. Sensibilidade</p> <p>Avalie sensibilidade ou mímica facial ao beliscar ou retirada do estímulo doloroso em paciente torporoso ou afásico. Somente a perda de sensibilidade atribuída ao AVC é registrada como anormal e o examinador deve testar tantas áreas do corpo (braços [exceto mãos], pernas, tronco e face) quantas forem necessárias para checar acuradamente uma perda hemisensitiva. Um escore de 2, “grave ou total” devem ser dados somente quando uma perda grave ou total da sensibilidade pode ser claramente demonstrada. Portanto, pacientes em esturpor e afásicos irão receber provavelmente 1 ou 0. O paciente com AVC de tronco que tem perda de sensibilidade bilateral recebe 2. Se o paciente não responde e está quadriplégico, marque 2. Pacientes em coma (item 1a=3) recebem arbitrariamente 2 neste item.</p>	<p>0= Normal; nenhuma perda.</p> <p>1= Perda sensitiva leve a moderada; a sensibilidade ao beliscar é menos aguda ou diminuída do lado afetado, ou há uma perda da dor superficial ao beliscar, mas o paciente está ciente de que está sendo tocado.</p> <p>2= Perda da sensibilidade grave ou total; o paciente não sente que está sendo tocado.</p>
<p>9. Melhor linguagem</p> <p>Uma grande quantidade de informações acerca da compreensão pode obtida durante a aplicação dos itens precedentes do exame. O paciente é solicitado a descrever o que está acontecendo no quadro em anexo, a nomear os itens na lista de identificação anexa e a ler da lista de sentença anexa. A compreensão é julgada a partir destas respostas assim como das de todos os comandos no exame neurológico geral precedente. Se a perda visual interfere com os testes, peça ao paciente que identifique objetos colocados em sua mão, repita e produza falas. O paciente intubado deve ser incentivado a escrever. O paciente em coma (Item 1A=3) receberá automaticamente 3 neste item. O examinador deve escolher um escore para pacientes em estupor ou pouco cooperativos, mas a pontuação 3 deve ser reservada ao paciente que está mudo e que não segue nenhum comando simples.</p>	<p>0= Sem afasia; normal.</p> <p>1= Afasia leve a moderada; alguma perda óbvia da fluência ou dificuldade de compreensão, sem limitação significativa das ideias expressão ou forma de expressão. A redução do discurso e/ou compreensão, entretanto, dificultam ou impossibilitam a conversação sobre o material fornecido. Por exemplo, na conversa sobre o material fornecido, o examinador pode identificar figuras ou item da lista de nomeação a partir da resposta do paciente.</p> <p>2= Afasia grave; toda a comunicação é feita através de expressões fragmentadas; grande necessidade de interferência, questionamento e adivinhação por parte do ouvinte. A quantidade de informação que pode ser trocada é limitada; o ouvinte carrega o fardo da comunicação. O examinador não consegue identificar itens do material fornecido a partir da resposta do paciente.</p> <p>3= Mudo, afasia global; nenhuma fala útil ou compreensão auditiva.</p>
<p>10. Disartria</p> <p>Se acredita que o paciente é normal, uma avaliação mais adequada é obtida, pedindo-se ao paciente que leia ou repita palavras da lista anexa. Se o paciente tem afasia grave, a clareza da articulação da fala espontânea pode ser graduada. Somente se o paciente estiver intubado ou tiver outras barreiras físicas a produção da fala, este item deverá ser considerado não testável (NT). Não diga ao paciente por que ele está sendo testado.</p>	<p>0= Normal.</p> <p>1= Disartria leve a moderada; paciente arrasta pelo menos algumas palavras, e na pior das hipóteses, pode ser entendido, com alguma dificuldade.</p> <p>2= Disartria grave; fala do paciente é tão empastada que chega a ser ininteligível, na ausência de disfasia ou com disfasia desproporcional, ou é mudo/anártrico.</p> <p>NT = Intubado ou outra barreira física; explique _____</p>

<p>11. Extinção ou Desatenção (antiga negligência) Informação suficiente para a identificação de negligência pode ter sido obtida durante os testes anteriores. Se o paciente tem perda visual grave, que impede o teste da estimulação visual dupla simultânea, e os estímulos cutâneos são normais, o escore é normal. Se o paciente tem afasia, mas parece atentar para ambos os lados, o escore é normal. A presença de negligência espacial visual ou anosagnosia pode também ser considerada como evidência de negligência. Como a anormalidade só é pontuada se presente, o item nunca é considerado não testável.</p>	<p>0= Nenhuma anormalidade. 1= Desatenção visual, tátil, auditiva, espacial ou pessoal, ou extinção à estimulação simultânea em uma das modalidades sensoriais. 2= Profunda hemi-desatenção ou hemidesatenção para mais de uma modalidade; não reconhece a própria mão e se orienta somente para um lado do espaço.</p>
---	---

